Ground-based optical interferometry: a practical primer

Chris Haniff

Astrophysics Group, Cavendish Laboratory, University of Cambridge, J.J. Thomson Avenue, Cambridge, CB3 0HE, UK

Abstract

In our earlier lectures we have been introduced to some of the theoretical bases of ground-based optical/IR interferometry. In this paper, we take a look at the subject again, but with a focus on how the conceptual ideas described before can be implemented in practice. For astronomical users of the VLTI, some familiarity with these ideas is beneficial from the point of view of planning interferometric observations and in understanding what limitations a specific implementation may imply. As in my earlier chapter, this treatment will be brief, but will draw attention to a number of the key underlying principles.

Key words: Interferometry, synthesis telescopes, array design, delay lines, beam combiners
PACS: 07.60.Ly, 42.15.Eq, 42.25.Hz, 42.68.Bz, 95.55.-n, 95.55.Br, 95.75.Kk

1 Introduction

As we have seen in previous chapters, the fundamental principles associated with interferometric imaging are reasonably straightforward. The radiation from a source is sampled at different locations, these samples are used to estimate the spatial coherence function of the radiation, and these data are inverse Fourier transformed so as to recover the source brightness distribution. In this chapter, we will look at some of the more practical issues related to performing these tasks. Since the detailed design of an array such as the VLTI represents a huge technical task, we will only have time here to touch the surface of many of the relevant issues. However, as in my earlier chapter, my goals will be to identify the key physical principles underlying these challenges

Email address: cah@mrao.cam.ac.uk (Chris Haniff).

so that the reader will, in her own time, be able to assess the relative merits of different technical solutions, and appreciate the origins of the limitations that constrain any given implementation.

As before, space constraints have meant that I have not been able to address the specific technical challenges associated with the search for extra-solar planets and with astrometric interferometry. There is a broad consensus that both of these science goals demand that exquisite care and attention be paid to the interferometer design: the reader should not assume that the omission of these topics here implies that these challenges are beyond the capability of the VLTI, but simply that a helpful treatment would have been beyond the scope of this lecture. By way of introduction, readers interested in the implementation of the astrometric and planet-hunting mode of the VLTI are referred to the recent review of the PRIMA instrument by Delplancke et al. (2006) and the references therein.

2 A smorgasbord of tasks

From the point of view of someone tasked with designing one, the functional requirements of the various sub-systems comprising an interferometric array can be treated as a linear sequence. Broadly speaking, once the radiation from the source has been sampled, each of the interferometric sub-systems performs some activity on the light and then "hands it over" to the next sub-system. This can be most easily appreciated by looking at Fig. 1. This shows a cartoon of the VLTI optical train with the signals from the target following a path from source to detector that can be roughly described as follows:

(1) Sampling of the light by two telescopes
(2) Relay of the light beams to a central beam-combining laboratory
(3) Correction for the geometric delay between the beams
(4) Combination of the signals to form a fringe pattern
(5) Detection of the fringe pattern and estimation of the fringe amplitude and phase

In each case, the sub-system charged with performing the function can operate quasi-autonomously from all the others, and its design can, to first-order, be optimised independently too.

In the following sections we will examine a number of these key tasks in more detail, and comment on three further aspects of interferometric practice that deserve mention: first, the impact of spatial perturbations on the wavefronts hitting the telescopes, second, the impact of temporal atmospheric perturbations, and finally, how we can usefully characterise the sensitivity of

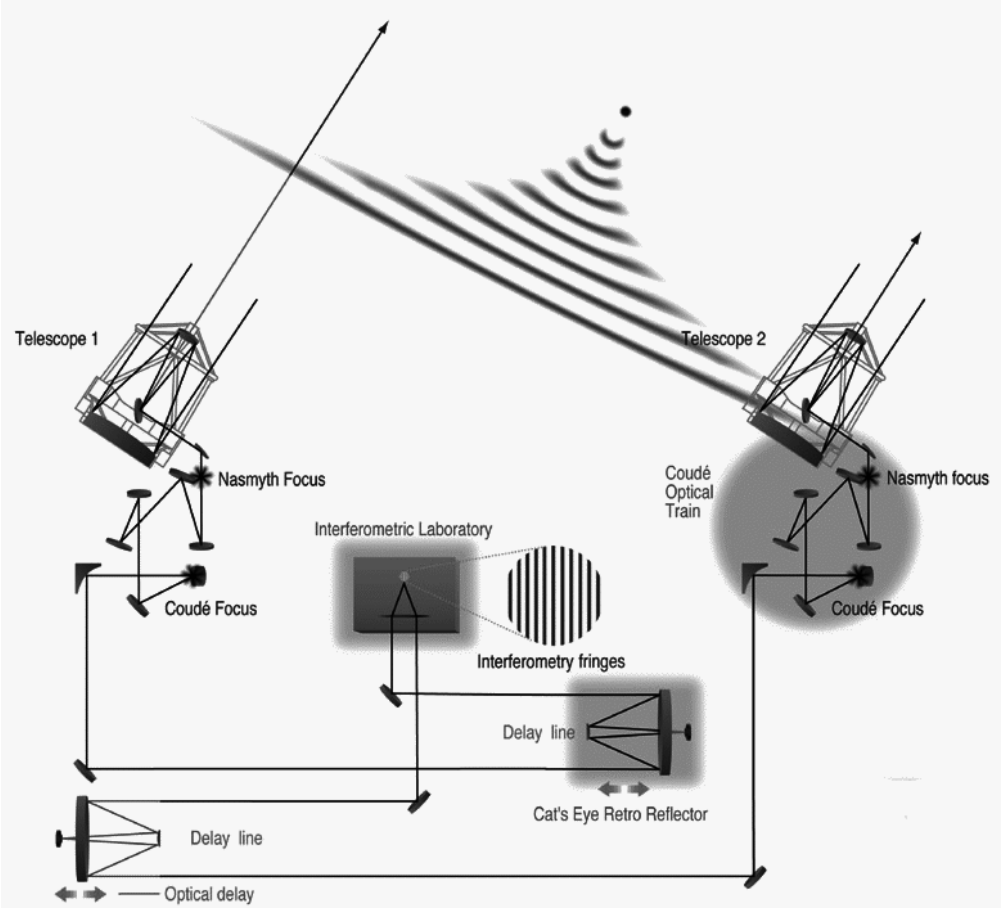


Fig. 1. A schematic cartoon showing the principal sub-systems associated with the VLTI. Light from the target is intercepted by the two telescopes, is relayed to the delay lines, and thereafter is mixed in a focal plane to form the fringes which encode information on the source structure.

an interferometric array.

3 Where should the radiation field be sampled?

The basic "rules-of-thumb" relevant to choosing how best to sample the radiation at ground level have been described in my earlier chapter. The critical parameters we discussed there were the projected separation vectors between the array elements (as seen from the target), since these determine the samples of the Fourier transform of the source brightness distribution that the interferometer measures. Of particular importance were (i) the overall number of data secured — this must be sufficiently large if the structure of the source under study is complicated, (ii) the uniformity of the distribution of samples in the Fourier plane, and (iii) the matching of the shortest and longest projected

baselines to the range of angular scales of interest.

In many practical situations, the ability to optimise the Fourier plane sampling is often restricted by rather simple considerations. I have listed a few of these below to provide a flavour of what is meant; the reader will I hope appreciate that this list is not comprehensive:

- Can the array elements be relocated appropriately without interfering with each other mechanically, or shadowing the light paths from the target?
- Can the delay lines introduce the appropriate geometric delay for the give combination of telescope location and instantaneous target position?
- Will the zenith distance of the target be sufficiently small such that the effects of a large air mass not compromise the fringe contrast?
- Will the source be too resolved on the shorter baselines to permit on-source fringe-tracking (see later)?

Many contemporary interferometers such as the CHARA, NPOI and COAST arrays have exploited "Y"-shaped configurations of telescopes. These give good instantaneous Fourier plane coverage, and permit the use of shallow angle reflections to redirect the light back to the central beam combining laboratory — this latter feature can help in maintaining the polarization state of the radiation as it propagates through the interferometer. However, this type of telescope layout is certainly not obligatory, and so whatever "works" in terms of satisfying the rules of thumb, for example the grid-like arrangement of Auxiliary Telescope pads at the VLTI, will usually suffice.

4 Methods for beam relay

Two mechanisms have been favoured for relaying the light from the interferometric collectors to the central beam combining laboratory in most contemporary arrays. By far the most popular strategy is to send the light along a shielded duct in the form of a collimated beam. By making the beam diameter, D, sufficiently large, i.e. by setting $D > (\lambda z)^{1/2}$, where z is the propagation distance, and λ the wavelength, diffractive spreading of the beams can be minimized, and little light loss is incurred. More sophisticated approaches (see, e.g. Horton et al., 2001) use slightly narrower beam profiles and allow modal filtering of the light to take place as the beams propagate. In many cases free-space propagation is also combined with an evacuated lightpath. Vacuum propagation ensures that both longitudinal dispersion, i.e. the mismatch between the optical paths for different colours, and the effects of optical turbulence can be wholly eliminated within the relay train and leads to fully polychromatic performance.

An alternative approach, which has been receiving more attention of late, is the use of optical fibres, i.e. waveguide technology. This has the potential for more flexible "reconfiguration" of the relay system — one could in principle imagine a "plug-and-play"interferometer — but the management of multi-wavelength beam injection and dispersion compensation over long relay paths remain challenging (see, e.g. Vergnole et al., 2005).

Whichever of these approaches is adopted, polarization control is an important concern. There are at least two issues to address:

(1) Can the polarization states of the interfering beams be matched so as to allow for interference to take place?
(2) Can the beam transport sub-system leave the polarization state of the radiation unchanged as it is delivered to the beam combining laboratory?

In free-space beam relay systems the first of these issues can be mitigated by ensuring that that the beams of light from all telescopes hit their respective relay mirrors at the same angles of incidence, i.e. that the relay trains are geometrically matched (see, e.g. Traub, 1988). The second, is generally realised by using a relay geometry that utilises mirrors at near-normal incidence, thereby limiting any diattenuation of the beams. For fibre-based beam transport, the use of polarization-preserving fibres, explicit polarization compensation and environmental control can all be exploited. Initial results for the OHANA project appear to suggest that any polarization-induced reduction in apparent fringe contrast can be managed adequately (Kotani et al., 2005) so that fibre beam transport may become increasingly popular in the future.

5 Equalizing the optical path

In any interference experiment there is always need to match the optical paths of the interfering beams to better than the coherence length of the light being mixed, i.e. the optical path difference (OPD) should be $\leq l_{\rm coh} \sim \lambda^2/\Delta\lambda$, and this is no different at an interferometer such as the VLTI. The major contributor to this OPD will be the geometric delay associated with relative orientation of the target and interferometer baseline. This will be equal to $\vec{\hat{s}} \cdot \vec{B}$, where $\vec{\hat{s}}$ is a unit vector in the pointing direction and $\vec{B}$ is the baseline vector. So, in principle, delays of as large as the longest baseline may need to be introduced for targets close to the horizon.

The optical path needed to correct the geometric delay is usually introduced using a movable carriage carrying retro-reflecting optics, the whole mechanism being called a delay line. The position of the carriage is typically measured using off-the-shelf laser metrology hardware. These commercial metrology sys-

Fig. 2. A schematic cartoon showing of a cat's-eye based delay line. In this design, light enters at the top right, is focused by a parabolic primary onto a flat secondary located at the focus, and reappears to the right as a collimated beam beneath the input collimated beam. The drive wheels for the carriage can be seen at lower right and left, as well as the vacuum vessel in which the whole assembly runs. (Figure courtesy of D. Sun)

tems can routinely monitor positions to a few tens of nano-metres over many tens or even hundreds of metres, and so the art of delay line design is really in controlling the dynamical behaviour of the opto-mechanical components rather than monitoring them. It is interesting to consider how fast the geometric delay can change: a rough estimate is given by $|\vec{B}|\cos(\theta)\frac{d\theta}{dt}$, where θ is the angle between the $\vec{\hat{s}}$ and the zenith. At the VLTI, this can be as fast as $\sim 0.5\,\mathrm{cm\ s^{-1}}$, although the delay-line carriages can move much faster than this when slewing between targets.

Two delay lines are shown in fig. 1, allowing for variable extra optical paths to be introduced into each telescope beam separately. The optical elements are normally vibrationally decoupled from the moving carriage using some type of compliant flexures, and the motion of the whole assembly is typically controlled using a combination of a coarse motor stage and more precisely commanded electromagnetic actuators. Fig. 2 shows a schematic view of a modern delay line that uses a parabolic primary/flat secondary mirror optical design. The input beam enters from the right and is returned to the right, in identical form, but displaced beneath its incoming partner.

The precision with which the OPD must be controlled will depend on the maximum permissible reduction in apparent fringe contrast that is allowed. At the VLTI the OPD jitter introduced by the delay lines is only a few tens of nm (i.e. less than one fiftieth of a wave at $1.25\,\mu$m) and so since the coherence length in even the lowest spectral resolution modes of AMBER and MIDI are several waves, the loss in fringe contrast due to finite bandwidth effects will be negligible if the delay lines can be commanded to move to the correct locations.

There is one final aspect of delay compensation that is worth mentioning here. This is associated with the fact that the geometric delay occurs in vacuum, and so its correction should ideally take place using evacuated delay lines. While

this is realised at a number of interferometers, e.g. the NPOI in Flagstaff, this is not the case at the VLTI. As a result, since the delay is compensated for in a medium with a wavelength-dependent refractive index, for any given pointing direction the compensation will only be correct for a single wavelength. At long wavelengths, the effect of this so called "longitudinal dispersion" (see, e.g. Lawson & Davis (1996); Tango (1990)) is, generally speaking, small. However, for long baselines and at large zenith angles with broad bandpasses it must be corrected for with additional optical elements. To get a feel for the magnitude of this effect at the VLTI we can consider measurements of a source 50° from the zenith with a 100 m baseline. Between the edges of the near-infrared K-band, from 2.0 – 2.5 μm, the dispersion-induced differential delay will be $\sim$ 10 μm and hence comparable to the coherence length of the light. From a practical point of view this implies that keeping the visibility losses due to this effect below 10% will need a spectral resolution, $R \geq 5$, or a value of $R > 12$ for measurements made in the J band at 1.25 μm.

We can close this section with a reminder of a result we introduced in an earlier chapter, i.e. that the use of delay lines as described almost always enforces a narrow field of view. This is simply a consequence of the fact that the geometric delay is a function of the source location and so a delay line can only correct for the geometric delay for one part of a source at a time. This quantitative result says that the field of view can be no larger than approximately $[\lambda/B_{max}] \times [\lambda/\Delta\lambda]$, i.e. the product of the spatial and spectral resolutions. A work-around to this problem does exist for two independed targets separated by some angle θ. In this case, if the light from each can be propagated separately from the telescopes, the use of one primary delay line together with a smaller differential delay line can be exploited to correct for the two unequal geometric delays.

6 Beam combination at optical/IR wavelengths

We have already seen that the essential principle underlying beam combination at optical and near-IR wavelengths is the *addition* of the electric fields from the interferometer arms, and the visualisation of the resulting intensity as a function of the OPD between the interfering beams. The contrast and location of these fringes then encode the amplitude and phase of the Fourier transform of the source brightness distribution.

The two most popular implementations of this process are shown schematically in Fig. 3. The first of these is usually referred to as "pupil plane" beam combination (left hand panel) and involves superposing afocal beams from each telescope at a beamsplitter plate (or its equivalent fibre- or integrated-optics component). This is exemplified by the MIDI instrument at the VLTI.

It is important to realise that it is at the beamsplitter plate that the fields are added and that any optics thereafter serve only to deliver the resulting intensity to a suitable detector. Usually a 50 : 50 splitter is employed and the complementary outputs are focused onto two separate single pixel detectors. If it is desired, each of these beams can be dispersed to give a 1-dimensional array of spectro-interferometric measurements. The interferometric fringes are visualised by deliberately introducing an OPD between the beams as a function of time. In multi-beam pupil plane combiners (see, e.g. Mozurkewich, 1994) non-redundant modulation of the optical paths in each beam is used to give rise to a modulated intensity output that is separable into different temporal frequency components. Each of these can then be attributed to a unique interferometer baseline.

An alternative approach, used in the AMBER instrument at the VLTI, involves adding the fields from each telescope in a focal plane. This is conventionally called "image plane" combination and is depicted schematically in the right hand panel of Fig. 3. This is exactly what Michelson implemented in his stellar interferometer, with the focal plane image being crossed by fringes. The angular size of the image will be governed by the diameters of the incoming beams, and will thus set the minimum detector array size, while the fringe period, and hence the necessary pixel size, will depend on the physical separation of the incoming beams at the focusing lens/mirror. This last point is important to stress because it underlies the method by which multi-way beam combination can be realised: if many beams are being interfered then the input pupil configuration must be chosen so that every pair of beams corresponds to a uniquely identifiable vector separation and hence spatial fringe period on the detector. In other words, the fringe encoding relies upon using a non-redundant input pupil. This can often be arranged with a 1-dimensional configuration of beams which then allows the other spatial dimension of an array detector to be used for wavelength dispersion as in AMBER (Robbe-Dubois et al., 2007).

The fact that the VLTI has exploited both pupil- and image-plane combiners highlights an important lesson: while the two approaches of temporal and spatial fringe encoding are formally equivalent, the choice as to which to use will actually depend on considerations such as the availability of low-noise array detectors, access to linear and fast path modulators, and the choice of suitable optical components.

In the near future it is likely that the use of integrated optics components may become more common in interferometric beam combiners. Results with first-generation devices have been very successful (see, e.g. Lebouquin et al., 2006) but what remains to be seen is how well they can be adapted to cater for multi-beam and multi-wavelength operation.

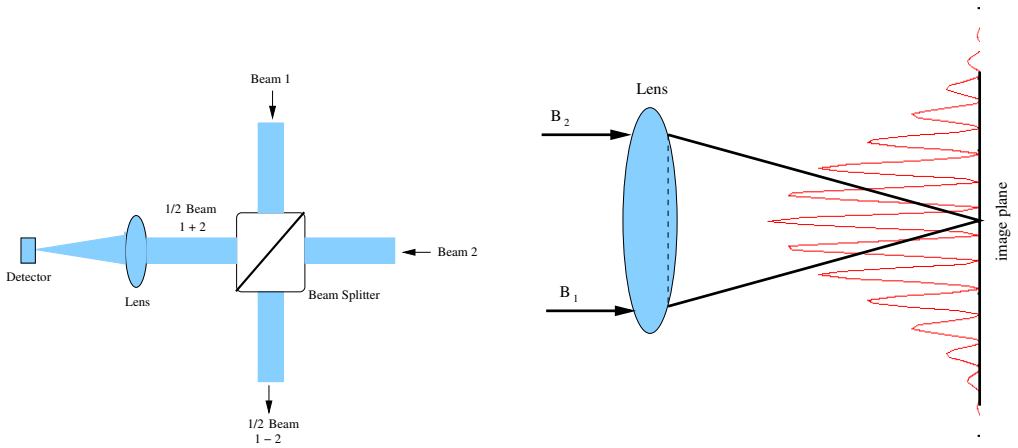


Fig. 3. Schematic diagrams of the optical set-ups required for two-beam pupil plane (left) and image plane (right) beam combination. In both panels a pair of collimated beams enters the beam combiner and is focused onto a detector. In the case of pupil-plane combination, each of the two complementary outputs from the beamsplitter is focused onto a single pixel (this is only shown for one output in the figure), and the intensity is measured as a function of a deliberately introduced OPD. In the image plane configuration, the intensity as a function of OPD is directly visible as a spatial fringe pattern on an array detectors. (Figures courtesy of D. Pearson.)

7 Dealing with atmospheric perturbations: problems and solutions

A lecture on the practical aspects of ground based optical/IR interferometry would not be complete without a significant digression on the negative impact of the atmosphere on the capabilities of the method. This has arguably been (and continues to be) the most important challenge for ground-based interferometers, i.e. how to overcome the seeing. The fluctuations the atmosphere introduces into the wavefronts arriving from astronomical sources are the most important factors limiting the astrophysical exploitation of optical/IR interferometry, and so an understanding of these perturbations is of particular importance for us today.

A comprehensive treatment of the spatio-temporal fluctuations produced by the atmosphere is beyond the scope of this paper, but interested readers are encouraged to consult Roddier (1981) for an excellent overview of the subject. To first-order, however, a useful model for the effects of the atmosphere is to associate them with a time-varying corrugation of the wavefronts from the source. So, we can ignore any fluctuations in the amplitude of the received radiation but we do need to consider the impact of phase perturbations. The questions I want to concentrate on here then are wholly pragmatic: exactly how do these perturbations limit the ability of interferometers to secure high quality and astronomically interesting Fourier data, and how can these limitations be overcome?

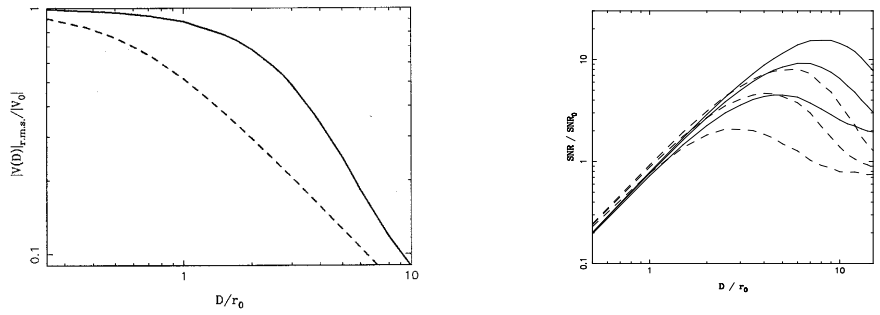


Fig. 4. The effect of spatial wavefront corrugations on interferometric measurements. The left hand panel shows the rms visibility amplitude expected for an unresolved star as a function of the ratio D/r_o. Results for uncorrected (dashed line) and tip-tilt corrected (solid line) optical systems are given. The right hand panel shows the signal-to-noise ratio for power spectrum measurements for telescopes with different levels of modal adaptive correction. Results are show for perfect 2-, 5- and 9-Zernike mode correction, with (solid line) and without (dashed line) the use of an optical fibre spatial filter. Note the beneficial effects of adaptive optics and spatial filtering.

7.1 Spatial fluctuations

The spatial fluctuations introduced by the atmosphere are conventionally characterised by a spatial scale, Fried's parameter, r_o. This is roughly equal to the diameter of the circular aperture over which the root mean-square wavefront perturbation is 1 radian. Interestingly, the strength of the wavefront perturbations is a strongly decreasing function of linear scale, r: the power spectrum of the fluctuations scales as $\kappa^{-11/3}$, where $\kappa \propto 1/r$. This spectrum has a pole at the origin (corresponding to the largest linear scales) and so, on physical grounds, it is usually assumed that on scales larger than some maximum, L_0, the so-called "outer scale", the fluctuation strength saturates.

Fried's parameter varies with wavelength like $\lambda^{6/5}$, and at the best astronomical sites takes values of order 15 cm at 500 nm. Broadly speaking telescope diameters, D, smaller or greater than r_o will give instantaneous images that are either diffraction limited ($D < r_o$) or highly distorted and speckled ($D > r_o$). It is interesting to compare the value of the ratio D/r_o at radio and optical wavelengths. At centimetric radio wavelengths, r_o is of order 30 km and so at the Very Large Array in New Mexico $D/r_o \sim 10^{-3}$. At the VLTI on the other hand, for the 8 m UTs at 2.2 μm, $D/r_o \sim 11$: this difference — a factor of 10^4 — highlights how much more of a problem the atmosphere is for us as compared to our radio colleagues.

The impact of atmospherically-induced spatial wavefront fluctuations on interferometric measurements is summarized in the left-hand panel of Fig. 4. This shows the expected root-mean-square visibility amplitude for observations of an unresolved source ($V_{\rm intrinsic} = 1.0$) as a function of D/r_o, both with

and without the use of a fast autoguiding system (Buscher, 1988). The major effects of the wavefront corrugations are a rapid drop in $V_{\rm rms}$, and a rapid increase in the fluctuations of V, as the telescope size increases. Together these lead to a loss in sensitivity (see section 8) and an increased difficulty in calibration. The situation is improved with the use of a fast autoguider, but this is only effective for aperture sizes of approximately $\leq 3r_o$. Beyond this, higher order adaptive optics correction becomes necessary to limit the precipitous drop in $V_{\rm rms}$ as the aperture size is increased.

Most optical/IR interferometric arrays use one (or both) of the following mitigation strategies to help moderate these performance hits:

- In principle one can measure the spatial fluctuations over each element in the array and correct for them in real time. This basically means fitting-out each telescope with its own adaptive optics (AO) system. The utility of AO for interferometry is a strong function of the ratio D/r_o (see Fig. 4 and table 1). For small telescopes correction of only the tip and tilt components of the perturbations may be adequate. However, once D/r_o exceeds three, higher order systems become desirable.

  The only real problem with this approach is the need to find a AO guide star close enough and bright enough to drive the wavefront sensor satisfactorily. At the time of writing (Spring 2007) the MACAO units at the VLTI can operate using reference stars as faint at $m_v = 17$ up to $57.5''$ away from the science target. For bright on-axis reference stars Strehl ratios as high as 60% at $2.2\,\mu$m can be achieved, but at $m_v = 16.5$ this is typically reduced to of order 10%. The Strehl's realised with off-axis reference stars will be similarly compromised, and so unless the seeing conditions are favorable, it may be more efficient to stop down the telescopes to reduce the value of D/r_o.
- A completely different — and passive — approach is to spatially filter the light being delivered to the beam combiners (see, e.g. Shaklan & Roddier, 1988). This exchanges the reduced and fluctuating visibility signal provided by a non-adaptively corrected array for a signal where the fringe visibility remains constant but where the overall optical throughput varies as a

λ/μm	1.25	1.65	2.2	10.0	20.0
ATs	4.8	3.4	2.4	0.4	0.2
UTs	21.2	15.2	10.7	1.7	0.8

Table 1
The ratio of D/r_o for the auxiliary (AT) and unit telescopes (UT) of the VLTI at different near and mid-infrared wavelengths. A median value (based on recent statistics) of 12.6 cm for r_o at 0.5μm has been assumed. Values of $D/r_o > 3$ imply the need for some form of moderate order adaptive optics correction to allow the telescopes to be used effectively for interferometry.

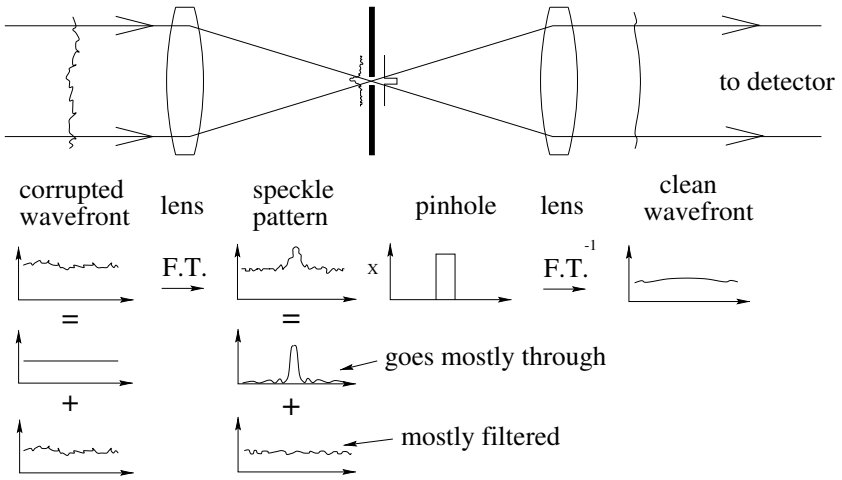


Fig. 5. A schematic illustration of how spatial filtering with a pinhole works. In the bottom panels three key ideas are presented from left to right: (i) the decomposition of the incident wavefront into a perfectly flat and a disturbed component (ii) the origin of the focal plane fields as the Fourier transform of the fields in the aperture (iii) the action of the pinhole as a "top-hat" filter for the focal plane field distribution. The behaviour of a piece of single-mode optical fibre is similar but with a transfer function described by the electric field profile within the fibre core.

function of time.

The physical basis for this method is the use of an optical component which only accepts a particular spatial mode of the incident radiation field. If this can be chosen so as to be "orthogonal" to components of the field associated with the atmospheric perturbations, then that part of the light can be rejected (this is depicted schematically in Fig. 5). Both pinholes and pieces of single-mode optical fibre can act as suitable spatial filters though to date most interferometric implementations have used fibre components. For a critical comparison of these two implementations the interested reader is referred to the paper by Keen et al, (2001).

A interesting trick is to take advantage of both of these methods. The effect of this mixed strategy can be quantified by examining the signal-to-noise ratio (S/N) for low-light-level measurements of fringe amplitudes. In this regime a useful metric of performance is the power spectrum S/N. This is plotted, for the photon-limited case, in the right-hand panel of Fig. 4 (Buscher & Shaklan, 1994). The different curves show how the signal-to-noise varies with D/r_o for varying levels of adaptive correction, both with (solid lines) and without (dashed lines) single-mode fibre spatial filtering. Several interesting results are immediately obvious:

- It is *always* advantageous to use a spatial filter unless the telescope size is smaller than or comparable to r_o. However, the merit of the resulting S/N enhancement needs to be balanced against the cost of its implementation.

- Higher order adaptive correction is *always* more beneficial than lower order correction, though once again, the cost of implementing it may outweigh its benefit.
- For any given order of adaptive correction there will be an optimum telescope size, beyond which the S/N will *decrease*. This will perhaps be the most surprising result to the novice reader: what the computations reveal is that in certain instances *decreasing* the telescope diameter can enhance the sensitivity of the array.

To summarize then, reductions in fringe visibility and signal-to-noise ratio arising from spatial perturbations in the incoming wavefronts can be usefully moderated by exploiting adaptive optics and spatial filtering. The latter can be used for any target, while the success of the former will rely upon the source being bright enough to act as a wavefront reference or the presence of a suitable natural or laser guide star. In this sense, an appropriate way of viewing adaptive optics for interferometry is not primarily as a route for enhanced limiting sensitivity, but rather as a means of improving the sky coverage of interferometers and in allowing observations of moderately bright targets to be executed more rapidly or at high spectral resolution.

7.2 Temporal fluctuations

The properties of temporal wavefront fluctuations can be characterised in a similar fashion to their spatial counterparts, i.e. one can define a coherence time, t_o, which measures the time over which the rms variation of the wavefront phase at a fixed point reaches one radian. The precise relationship between the spatial scale r_o and the coherence time is complex. One scenario pictures the temporal fluctuations arising from the wind-driven motion of a "frozen" layer of turbulence past the interferometer, while another has some type of in-situ "boiling" of the wavefronts taking place (see, e.g. St.-Jacques & Baldwin, 2000, for a discussion of this point).

For a frozen screen, where the wavefront evolution time is assumed to be much greater than the time for the screen to blow past one of the collectors, the coherence time can be written as $t_o = 0.314 r_o/v$, where v is a characteristic wind velocity. In this picture, large scale spatial perturbations will be associated with timescales longer than the coherence time. Values of t_o typical of modern observatory sites are between 2 and 20 ms at 500 nm and scale with wavelength in the same manner as r_o. For the VLTI, recent data from Paranal give a median value of approximately 20 ms at 2.2μm, or equivalently 3 ms at 0.5μm. This small value of the coherence time, as compared to the exposure times use for conventional astronomical measurements, is one of the fundamental problems that ground-based interferometrists have to grapple with.

The effect of these temporal fluctuations on interferometric measurements can be best appreciated by remembering that they arise from changes in the optical paths along the lines of sight from the interferometric collectors to the source. Hence, their immediate and essential affect will to add an additional random and fluctuating component to the geometric delay. This will have three quite distinct implications:

- Fluctuations occurring on the shortest timescales will move the interferometric fringes past the detectors and reduce their measured contrast. To limit this systematic reduction in source visibility amplitude, the basic exposure time must be kept very short, of order or less than t_o.
- Fluctuations on longer timescales will lead to large offsets in the position of the center of the coherence envelope. For Kolmogorov turbulence the rms optical path difference, $\sigma_{\rm opd}$, for a baseline of length B is equal to $0.417\lambda(B/r_o)^{5/6}$ (see, e.g. Davis et al., 1995). On all but the very shortest baselines this fringe motion will likely exceed the coherence length of the radiation being observed and so some form of slow dynamic tracking of the white-light fringe motion will be required if fringes are to be observed.
- Unless the absolute value of the fringe motion can be monitored, measurements of the phase of the interferometric fringes will no longer characterise the phase of Fourier transform of the source brightness distribution. In this case it becomes necessary to focus attention on phase-type observables that remain resistant to this type of corruption if the phase data are needed.

It will be convenient to deal with these issues below by separating the solutions, as before, into active and passive techniques. We will first examine techniques for active fringe tracking.

7.2.1 Active fringe tracking

From the discussion above it should be clear that a hierarchy of fringe-tracking solutions must exist. In order of increasing difficulty these are:

(1) Monitoring the OPD fluctuations and re-positioning the delay lines so as to reduce the delay error to of order the coherence length, Λ_{coh}. Fringe data can then be secured for the fraction of the time that the random atmospheric OPD errors compensate the delay line positioning error.
(2) Monitoring the OPD fluctuations and re-positioning the delay lines so as to reduce the delay error to a small fraction of Λ_{coh}. Fringe data can then be secured continuously. This approach is usually referred to as "coherencing" the array.
(3) Monitoring the OPD fluctuations and correcting the OPD to a small fraction of a wavelength. Fringe data can then be secured continuously and using long exposure times. This approach is usually referred to as

"co-phasing" the array.

The reader should note that only the last of these possibilities allows for direct Fourier inversion of the visibility data and so the first two approaches still need to be combined with measurement strategies that focus on observables that are robust to any residual delay errors.

Coherencing usually involves dispersing the fringes, mapping them onto a 2-dimensional space of wavenumber vs fringe phase, and examining the 2-dimensional "spatial frequency" of the resulting fringes (Basden & Buscher, 2005). The value of this spatial frequency then measures the group delay, i.e. the location of the centre of the coherence envelope, hence the alternative name "group-delay" tracking. Group delay tracking represents a trade off between OPD sensing resolution and sensitivity. More precisely, the resolution in delay will be equal to the coherence length of the total bandpass, $\lambda^2/\Delta\lambda$, e.g. $\sim 12\,\mu$m for the near-IR K-band. However, since it will take the atmosphere some time to introduce such a large differential OPD, is will be possible to incoherently integrate the delay signal for many coherence times. For the 20% bandpass of the K-band an integration time as long as $30t_o$ might be suitable. Any science data, however, must be collected using much shorter integration times so as to limit any fringe smearing. Coherencing is the only technique that can be utilised successfully for the faintest targets: in the photon-limited regime it typically is at least two magnitudes more sensitive that co-phasing.

Coherencing can be contrasted with co-phasing, which has as its target "precision" correction of the atmospherically induced OPD. The basic idea is to monitor and control any OPD errors to better than a small fraction of a wavelength, so that the effective value of t_o is extended indefinitely. For the VLTI, the goal is to limit any resulting visibility loss to less than 1%, implying an rms OPD jitter of order 100 nm in the near-IR. This demands very rapid estimation of the instantaneous OPD (i.e. on timescales much shorter than t_o) and very low-latency control of the correcting element, hence the much poorer limiting sensitivity as compared to coherencing.

One can consider co-phasing as "adaptive optics" for the whole interferometer, stabilising it against temporal atmospheric fluctuations. Of course, perturbations to the OPD may arise from non-astronomical sources as well and so the success of co-phasing places quite strict demands on the optical path stability of every element in the interferometric optical train including the interferometric collectors.

There is a close relationship between our earlier discussion of adaptive optics for spatial wavefront control and the pros and cons of phase tracking. Both are active methods, and so both demand the presence of a suitable reference source to provide real-time feedback on the atmosphere. This signal can come

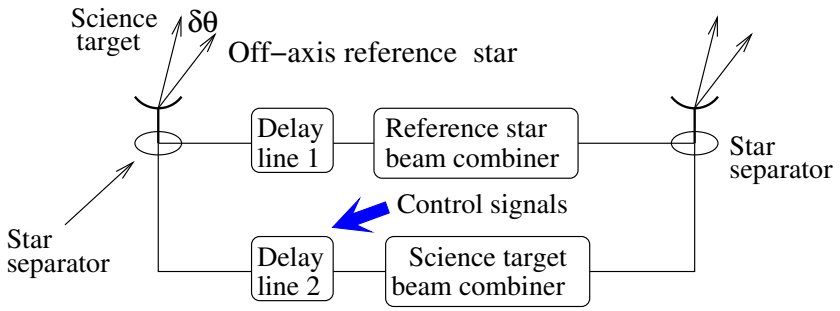
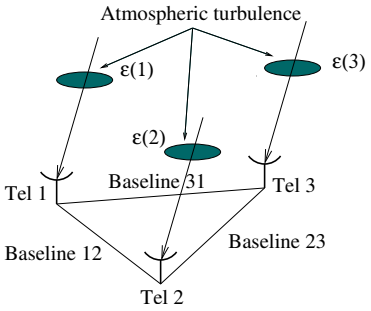


Fig. 6. A schematic layout of a "dual-beam" phase-tracking system such as the PRIMA instrument at the VLTI (Delplancke et al., 2006). The telescopes simultaneously observe two sources, separated by a small angle $\delta\theta$. The fringes from the reference source are monitored in real-time and corrections fed to a separate delay-line which matches the optical paths for the beams from the science target. A high-precision internal metrology system, not shown in the diagram, is used to tie the optical paths of the source and reference signals together.

from the science target itself, perhaps using light in a broadband channel adjacent to the science bandpass, or from a nearby reference star. If the latter approach is taken, this reference needs to be close (typically within a few tens of arc-seconds at $2.2\,\mu$m) to the science target.

The major difference between adaptive optics and phase-tracking becomes apparent once an off-axis reference source is being used. In AO systems the off axis beam is usually easy to deal with, because the telescope/instrument field-of-view is usually much greater than the off-axis offset angle; this is not the case for an interferometer. Here, the light from the reference target must be separated from that of the science object at the telescopes and be delivered to a separate beam combiner in the optical laboratory using an additional differential delay line (see Fig. 6). Moreover, the optical paths of the target and reference beams must be monitored with high precision so as to allow correction for any internal differential OPD fluctuations. The cost and complexity of this extra hardware, together with an additional requirement that the reference source not be resolved by the interferometer baseline (hence precluding the use of laser guide stars) means that the implementation of off-axis phase tracking is a non-trivial exercise.

We have seen then, that in any given situation observers will have to choose between a number of possible options for coping with the temporal fluctuations introduced by the atmosphere. For the brightest targets on-source co-phasing should be possible, while for fainter objects either self-referenced coherencing or off-axis group-delay tracking or co-phasing may be available. The presence, or not, of a suitable off-axis reference source will be a critical factor.



$$\Phi(12) = \phi(12) + \varepsilon(1) - \varepsilon(2)$$
$$\Phi(23) = \phi(23) + \varepsilon(2) - \varepsilon(3)$$
$$\Phi(31) = \phi(31) + \varepsilon(3) - \varepsilon(1)$$
$$\Phi(12) + \Phi(23) + \Phi(31) = \phi(12) + \phi(23) + \phi(31)$$

Fig. 7. The basic principle of closure phase measurement as applied to a simple 3-element interferometer. Each measured visibility phase, $\Phi(ij)$, is equal to the unperturbed visibility phase, $\phi(ij)$, to which have been added phase errors, $\epsilon(i)$ and $\epsilon(j)$, associated with the unknown optical paths above telescopes i and j. All of these terms cancel out when the measured visibility phases are summed "round the loop". This produces the closure phase, a phase-type quantity that only depends on the source structure, and is independent of the atmospheric phase errors.

7.2.2 Passive work-arounds for temporal fluctuations

As was the case for spatial wavefront fluctuations, it is also possible to exploit passive methods to mitigate against the temporal perturbations the atmosphere introduces. These take advantage of the fact that, for a variety of reasons, there are quantities — so called "good observables" — whose values are resistant to the presence of wavefront phase errors.

Consider, for example, making simultaneous measurements of the visibility function at two similar wavelengths. In the absence of a co-phasing sub-system, the instantaneous fringe phase at each wavelength will be equal to the true visibility phase plus an unknown phase error associated with the atmospherically-induced OPD. However, because the atmospheric optical path fluctuations are to first order achromatic — the refractive index of air varies only slowly with wavelength — the errors at one wavelength can be related those at any other. Hence, the phase error measured at one wavelength can be used to correct for the phase error at another. This method is most useful when the source is known to be unresolved at some "reference" wavelength so that its visibility phase can be assumed to be zero there. In this case, the difference in the measured fringe phases becomes a direct proxy for the true visibility phase at the science wavelength. This "differential phase" technique can in fact be used whatever the source structure is: all that is required is that the visibility phase be known *a priori* at some reference wavelength. The reader should note that the success of this method will be limited by the precision with which the dispersive effects of the atmosphere can be modelled.

Similar "robustness" to the instantaneous atmospheric fluctuations can be achieved by measuring visibility phases simultaneously on different baselines, where the baselines in question form a closed loop connecting at least three

telescopes (see Fig. 7). As the figure shows, in this case, the sum of the observed fringe phases, the "closure phase", is independent of the atmospheric error terms and is hence another good observable.

It is worth commenting on the fact that the number of independent closure triangles, N_c, that can be measured with an N-element interferometer is equal to $\frac{(N-1)(N-2)}{2}$ so that the fraction of the visibility phase information retained in the closure phases ($f = 1 - \frac{2}{N}$) rises rapidly as N increases. The increasing priority given to the number of array elements in modern interferometer designs, in part, stems from this result.

Before concluding this section, it will perhaps be useful to compare and contrast the differential and closure phase methods. Both are self-referenced techniques, both rely upon simultaneous measurements of different perturbed fringe phases, and both provide a signal that can be coherently integrated over many integration times. However, while differential phase methods are most useful when the source visibility function is known at some wavelength, the closure phase technique is wholly independent of the source morphology. Closure methods thus provide a powerful model-independent way of eliminating certain classes of perturbations, so-called antenna-dependent gain errors. The temporal fluctuations of the atmosphere give rise to just such errors but other types of error can, and do, exist and can only be overcome using other methods. Exactly how differential- and closure-phases are interpreted and used for interferometric science is covered in more detail elsewhere in this volume.

8 How can we quantify the sensitivity of an array?

The proper assessment of the sensitivity of an optical interferometric array is a complicated task because, as we have seen in an earlier chapter, it can depend on many different things. In particular, any of the following three considerations may be the limiting factor:

- Does the target (or a suitable reference source) provide enough photons so as to stabilize the array against temporal, and possibly spatial, atmospheric wavefront perturbations?
- Is the science target bright enough that the integrated signal-to-noise ratio on the visibility amplitudes and differential/closure/visibility phases is large enough after some moderate integration time (perhaps 5 minutes or so) to be useful?
- What is the faintest structure that can be reliably detected given the total number and quality of the Fourier data that have been secured?

For the purpose of our treatment today, I want to focus on only the first of these criteria. For ground-based arrays, which have to operate in the presence of rapidly varying atmospheric fluctuations, the ability to fringe-track will almost always be what limits the interferometric sensitivity. Furthermore, for most targets, the probability of finding a suitably bright and proximate reference star will be much less than 50% and so the brightness of the science target in some suitable waveband, which need not be the science bandpass of course, is what will be important.

It is reasonable to assume that at the sensitivity limit of the array it will be operating in a coherencing mode, and in this regime an appropriate metric for success will be the signal-to-noise ratio for group-delay fringe tracking. This takes the following form, which is characteristic of the S/N for a power-spectrum estimator:

$$(S/N) \simeq \frac{[VN]^2}{\sqrt{[(N+N_{\rm dark})^2 + 2(N+N_{\rm dark})N^2V^2 + 2(N_{\rm pix})^2(\sigma_{\rm read})^4]}} \ , \quad (1)$$

where V is the apparent fringe visibility, ranging between 0 and 1, N is the total number of photon counts detected from the source in the integration time, $N_{\rm dark}$ is the number of dark or background counts detected in the integration time, $N_{\rm pix}$ is the number of detector pixels over which the fringe signal is spread out, and $\sigma_{\rm read}$ is the readout noise per pixel. For useful tracking we will clearly require a signal-to-noise of order unity or above.

From the point of view of understanding what really impacts the sensitivity of an optical interferometric array the most important fact we have to remind ourselves of is that it is the *instantaneous* S/N that matters. i.e. the signal-to-noise that is realised in a time comparable to the coherence time. This is very different from the case where the total integration time matters, as for example in conventional imaging and spectroscopy, and this highlights the parallel between interferometry and adaptive optics where both demand that light from the target be used to drive real-time active control systems. In this sense, then, we should expect the limiting sensitivities of interferometric and adaptive optics systems to be comparable. The fact that most ground-based arrays struggle to observe targets of magnitude 10 in the near-infrared while many AO wavefront sensors operate up to 5 magnitudes fainter suggests that one can realistically expect the sensitivity of ground based arrays to improve by at least a factor of 50 in the near term, as optimized implementations mature.

In view of its perhaps unfamiliar nature, Eq. 1 deserves some careful scrutiny. The numerator is simply the power spectrum signal while the denominator is the quadrature sum of three different noise terms. From left to right these are

Poissonian photon noise, a noise contribution associated with the fringe signal itself, and a final term arising from readout noise in the detector. Depending on the particular implementation selected the relative contributions of each of these terms can vary dramatically, and so having some feel for this can be valuable when designing a "fringe tracking engine". In particular:

(1) Because the visibility, V, referred to in Eq. 1 characterises the visibility of the *detected* fringes, it incorporates all contributions from atmospheric and instrumental coherence losses including, e.g. the effect of the lowering of the apparent fringe contrast when more than two beams are interfered together at the same point.
(2) The contribution of readout noise can become dominant, especially when the fringes are detected by a large number of pixels, e.g. when they are being spectrally dispersed.
(3) Depending on the relative magnitudes of the terms in the denominator, maintaining high fringe visibility, V, may become less or more important relative to the maintenance of high throughput, i.e. N, from the point of view of keeping the S/N high.
(4) For sources that are significantly resolved ($V \ll 1$) it may be impossible to track fringes at all unless the target is extremely bright.

This last point is perhaps best demonstrated through an numerical example. Consider the measurement of the diameter of a low mass star with a two-element interferometer equipped with a photon-limited fringe sensor. Let us assume that on a short interferometer baseline, where the star is only moderately resolved ($V_{\rm source} = 0.75$) coherencing can be performed with a S/N of 3. Then, on a longer baseline where the fringe visibility has reduced to, say, 0.15 the signal-to-noise ratio for coherencing will have dropped by a factor of approximately 6 and so fringe tracking may cease to be possible. This corresponds to an effective reduction in the source brightness by a factor of 25 or equivalently 3.5 magnitudes.

This variation in sensitivity with source resolution may seem unusual to the novice interferometrist, but is a real challenge for ground-based interferometry: how can we observe faint and *resolved* targets? The obvious solutions are to either use off-axis guide stars for fringe stabilisation or to decompose all long baselines into shorter contiguous ones, so that monitoring of the atmospheric fluctuations can take place on baselines on which the target will not be as resolved. Once again we can see interesting parallels with strategies used in adaptive optics: there the use of off-axis guide stars is commonplace, as is the use of Shack-Hartmann wavefront sensors which break up the pupil into smaller zones across which the spatial wavefront perturbations can be relatively straightforwardly measured and then "stitched" together to recover the wavefront across the whole pupil.

Perhaps we can console ourselves with the fact that sources that are well matched in angular size to existing interferometers cannot be too faint: e.g., a 0.5 milli-arcsecond diameter blackbody with a temperature of 2500 K will have an apparent magnitude in the near-infrared of between 7 and 8. This is ~ 5 magnitudes brighter than the group-delay tracking limit for an interferometer employing 2 m-class collectors, and so sources with surface brightnesses 10^2 times lower should still be observable. By the same token, thermal sources that are much fainter than this are likely to subtend very small angles (i.e. much less than half a milli-arcsecond) and so will need to wait for a next generation of kilometric baseline arrays for their study.

9 Calibration

The calibration of interferometric measurements is a critical element of the practice of interferometry, and will be dealt with in detail elsewhere in this volume. However, it will be valuable to mention in passing just a few of the strategies that can help reduce the impact of calibration errors.

The fundamental basis for needing sophisticated calibration procedures in optical/IR interferometry is the sensitivity of the interferometric response to numerous time-variable quantities. A useful way of characterising this behaviour is in terms of a transfer function describing how a measured quantity — usually a visibility amplitude or phase — is related to its true value. The key problem at optical/IR wavelengths is that these transfer functions usually change on timescales of fractions of a second to hours due to variations in the local seeing conditions which feed through to all of the real-time control systems. In a very real sense then, the art of optical/IR interferometry is to design both an array and an observing strategy that delivers reliable interferometric data in the presence of fluctuating observing conditions.

There are numerous approaches that can be exploited to assist in mitigating the effects of variations in the interferometric response. In the following paragraphs I outline four of these. This is clearly not a comprehensive list, but it will hopefully give a flavour of the range of techniques that can be successfully applied.

- The most common strategy for handling calibration issues is to interleave observations of the science target with sources with known visibility functions. This is most commonly performed by observing unresolved sources close in both time and space to the science target. This permits statistical calibration of the interferometric response, most frequently for visibility amplitude calibration, but is of course dependent on the average seeing and instrumental conditions remaining constant between observations. Suffice

to say, the difficulties associated with finding calibrator stars of comparable brightness and sufficient compactness close enough to the science target, should not be underestimated (see, e.g. Daigne & Lestrade, 2003).

This type of statistical calibration can be usefully augmented by using real-time diagnostic data from the interferometer sub-systems to characterise the instrumental response. Hardware components such as the fast tip-tilt system and the delay-line metrology system can provide independent estimators for the perturbations sensed by the array, and can therefore be used to provide a first-cut calibration of the raw measurements even in the absence of any calibrator star data.

- A quite different approach to eliminating calibration uncertainties has been to design interferometers so as to be as insensitive as possible to the fluctuating ambient conditions. We have already seen an example of this in the use of spatial filtering. This dramatically reduces the impact of changes in the seeing by allowing the fluctuations in seeing to impact an observable that is only weakly related to the fringe contrast, i.e. the photometric coupling into the beam combiner. This type of strategy is perhaps best described as "ingenious experimental design" and is the mark of an experienced scientist who knows what she wants to measure and designs her apparatus to measure that alone.
- A somewhat similar technique, that has found most utility in allowing Fourier phase information to be recovered, has been to focus on measuring quantities that are largely unaffected by the seeing conditions. The measurement of good observables such as the closure phase is the classic example of this, but other quantities, such as the differential phase can be equally robust to seeing perturbations. Many of these good observables can be understood as examples of differential measurements, where the effects of the atmosphere essentially cancel out. They are thus much more robust against certain atmospherically induced systematics, but they can be just as easily compromised by instrumental biases, e.g. baseline dependent correlator errors, and so they should not be seen a panacea for all calibration problems.
- A final approach has simply been to measure interferometric observables whose mean properties are not systematically biased by the instrument and atmosphere. A comparison of the effects of atmospheric phase fluctuations on visibility amplitude and phase measurements is perhaps the clearest way to explain this. While random atmospheric fluctuations will only ever *reduce* the measured visibility amplitude of a target, they give rise to random *zero-mean* perturbations of the fringe phase. In principle then, ground-based fringe phase measurements ought to require little calibration but only averaging over many realizations of the atmosphere to produce unbiased estimators. In practice, however, other practical problems preclude this possibility, but it remains broadly speaking true that phase-type observables are generally less difficult to calibrate than amplitude data.

The enumeration of a long list of possible solution strategies always engenders

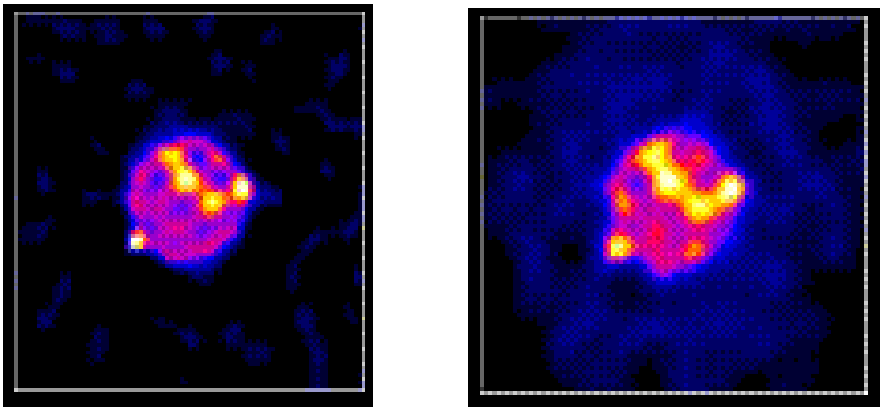


Fig. 8. Image reconstructions of a model spotted star made using visibility amplitudes and closure phases with two different levels of noise on the Fourier amplitudes. In the left hand panel a mean additive error of 0.008 has been combined with the data while in the right hand panel the mean visibility error is 0.100. In each case the zero-baseline visibility is 1.00 and the closure phases have typical errors of 20°. The reader should note that despite the poor visibility data used for the right hand reconstruction, the main features of the source are still recovered, albeit with some loss of photometric fidelity and an increase in background noise. The small change in image size is an artifact of the software used to prepare the images. (Figures courtesy of F. Baron.)

a risk that the earnest scientist will ignore asking an important question: "What level of calibration is actually needed for my science?" Fig. 8 may help to focus minds on this topic. The two panels in the figure shows image reconstructions of a resolved stellar disk model, the differences between the two being due to the differing level of noise added to the visibility amplitudes and closure phases used to recover them. Both reconstructions utilized relatively inaccurate closure phase data ($\sigma_{cphase} = 20°$), but the visibility data used in the left hand panel had additive noise levels equal to 0.8% of the total flux, while the equivalent errors for the right hand panel were 12.5 times higher.

What is interesting is that in both panels most of the coarse morphological and photometric data for the source have been correctly recovered. So, at least for this particular situation, calibration at the 1% level would likely not have been necessary for the purposes of determining what the target looked like. However, the lesson to take home here is not that calibration is unimportant — it is, principally because calibration errors may ultimately limit what science can be delivered — but that it pays to understand in advance how small the calibration errors really *need* to be.

10 Summary

This brief review has tried to present some helpful insights into the practice of optical/IR interferometry. It should be obvious that arrays such as the VLTI are a testament to the efforts of numerous dedicated engineers and technicians, who have risen to meet the many technical challenges that building an interferometer entails. While I may have enumerated some of these challenges are requirements, I hope the reader will not have been led to believe that interferometry is difficult, but rather that all of these implementational issues can, and are being, solved.

For those who stand to participate in our next steps on this path, I suggest only a few words of advice: ask what is really needed, design only that, and remember that almost every element of the system will depend on some other part that you will have temporarily forgotten about!

Acknowledgements

It is a pleasure to thank David Sun, Fabien Baron and Debbie Pearson for use of some of their graphics and all of colleagues who have helped better my understanding of the technical issues relevant to realizing working interferometric arrays. I would also like to acknowledge the generosity of the school organisers for inviting me and the participants for making the school so enjoyable.

References

Delplancke, F., Derie, F., Levque, S., Menardi, S., Abuter, R., Andolfato, L., Ballester, P., de Jong, J., Di Lieto, N., Duhoux, P., Frahm, R., Gitton, P., Glindemann, A., Palsa, R., Puech, F., Sahlmann, J., Schuhler, N., Duc, T.P., Valat, B., Wallander, A. 2006, Proc. SPIE., 6268, 62680U.

Horton, A.J., Buscher, D.F. & Haniff, C.A. 2001, MNRAS, 327, 217.

Vergnole, S., Kotani, T., Perrin, G., Delage, L. & Reynaud, F. 2005, Opt. Comm., 251, 115.

Traub, W.A. 1988, in *"High Resolution Imaging by Interferometry"*, ed. F. Merkle (European Southern Observatory: Garching, Germany), p. 1029.

Kotani, T., Perrin, G., Vergnole, S., Woillez, J. & Guerin, J. 2005, Appl. Opt., 44, 5029.

Lawson, P.R. & Davis, J. 1996, Appl. Opt., 35, 612.

Tango, W.J. 1990, Appl. Opt., 29, 516.

Mozurkewich, D. 1994, Proc. SPIE., 2200, 76.

Robbe-Dubois, S. et al. 2007, Astron. Astrophys., 464, 13.

Lebouquin, J.-B., Labeye, P., Malbet, F., Jocou, L., Zabihian, F., Rousselet-Perraut, K., Berger, J.-P., Delboulb, A., Kern, P., Glindemann, A., Schller, M. 2006, Astron. Astrophys., 450, 1259
Roddier, F. 1981, Prog. Optics 19, 281.
Buscher, D.F. 1988, MNRAS, 235, 1203.
Keen, J.W., Buscher, D.F. & Warner, P.J. 2001, MNRAS, 326, 1381.
Shaklan, S.B. & Roddier, F. 1988, Appl. Opt., 27, 2334.
Buscher, D.F. & Shaklan, S.B. 1994, Proc. SPIE., 2201, 980.
St.-Jacques, D. & Baldwin, J.E. 2000, Proc. SPIE., 4006, 951.
Davis, J., Lawson, P.R., Booth, A.J., Tango, W.J. & Thorvaldson, E.D. 1995, MNRAS, 273, L53.
Basden, A.G. & Buscher, D.F. 2005, MNRAS, 357, 656.
Daigne, G. & Lestrade, J.-F. 2003 Astron. Astrophys., 406, 1167.