

Copyright 2008 Society of Photo-Optical Instrumentation Engineers.

This paper was (will be) published in Proc. SPIE 7013 and is made available as an electronic reprint (preprint) with permission of SPIE. One print or electronic copy may be made for personal use only. Systematic or multiple reproduction, distribution to multiple locations via electronic or other means, duplication of any material in this paper for a fee or for commercial purposes, or modification of the content of the paper are prohibited.

Polarization fidelity in an optical interferometer

David F. Buscher, Fabien Baron, Christopher A. Haniff

University of Cambridge, Cavendish Laboratory, JJ Thomson Avenue, Cambridge CB3 0HE,
United Kingdom

ABSTRACT

The optical trains of interferometers invariably contain oblique reflections that alter the polarization state of the light from the source. Even for arrays with symmetric optical paths, large systematic visibility errors can be introduced when observing sources with intrinsic polarization. We have identified the key metric for polarization fidelity in an optical interferometer - the diattenuation of the optical train - and we evaluate the visibility penalties incurred by an interferometer that is not optimized for polarimetric purity for a number of different types of polarized source.

Keywords: Interferometry, Polarization, Polarimetry

1. INTRODUCTION

In an optical or infrared interferometer the polarisation states of the interfering beams affect the observed fringes. This paper examines to what extent making accurate interferometric images is possible when the quality of the measurements is degraded by polarization effects.

The polarization inherent to an astronomical source indicates a physical phenomenon at work. At optical wavelengths, the polarization structure of most astronomical targets generally arises from unpolarized thermal radiation (perhaps from a star) scattering off gas molecules or dust grains, for example in a stellar atmosphere or in an extended envelope around a star, to give linearly polarized light. In some objects, multiple scattering gives rise to circularly polarized light, though we will ignore this case in this paper. If the radiation source and the scattering medium have spherical symmetry then the observed polarization pattern will have circular symmetry. When observed with conventional monolithic telescopes the typical intrinsic polarization shows levels of less than a few percent.¹ This is due to an averaging effect: the polarized radiation is integrated over a large area, and the resulting total flux will be only weakly polarized. However interferometers make measurements on milliarcsecond scales where the scattering environment is likely to be resolved. High levels polarizations can then be observed. For instance the polarization from scattering in the envelopes of Mira-like variable stars has been measured at levels of order 15%.²

If we assume that there is no significant polarization of the signal during propagation to the Earth, then the polarization of the beams of an inteferometer is just the inherent polarization of the astronomical source modified by its propagation through the interferometer. In most existing interferometers, the dominant polarization effects arise from the reflections in the optical train. A quasi-monochromatic electromagnetic wave is conventionally modelled as the superposition of two perpendicular S and P vibrations. Reflections and partial absorption of the waves by the optics of an interferometer arm may attenuate the amplitudes of the S or P vibrations by different amounts, rotate those directions, or introduce a differential phase delay between them. For example, the reflection coefficient for light at a wavelength of 632nm incident at an angle of incidence of 45° on a bare silver mirror is about 1% greater for S -polarized light than for P -polarized light, and the P -polarized light is phase-retarded by about 160° when compared with the S -polarized light. If one considers that an interferometric optical train may contain several dozen such reflections in each “arm” of the interferometer, it is clear that the polarization effects of the optical train can be significant. Moreover in some interferometers, single-mode optical fibres are used for beam transport and/or spatial filtering.³ These are typically strongly birefringent devices which can therefore also affect the state of polarization of the starlight strongly.

D.F.B: E-mail: dfb@mrao.cam.ac.uk, Telephone: +44 (0)1223 337302

An additional complication is that the polarization properties of the interferometer optics can change as a function of time. The angles of incidence of the starlight on the mirrors of the individual light-collecting telescopes (we designate these as “unit telescopes” to distinguish them from the aperture-synthesis telescope formed by the interferometric array as a whole) will vary with the telescope pointing direction and this will cause the polarization effects of each reflection to vary from both from object to object, and as any given object tracks across the sky due to sidereal motion. Furthermore, when a given object is observed at different hour angles, the characteristic S and P polarization planes of the unit telescope mirrors will also typically rotate with respect to the plane of polarization of the object. On longer timescales, aging of optical coatings in the unit telescopes and elsewhere in the optical train can give rise to additional changes in the instrumental polarization properties.

The effect of this instrumental polarization on the fringes depends on the type of interferometer, in particular on the symmetry of the optical trains in different arms of the interferometer. In some interferometers the arms are not symmetric, and consequently the beams from different arms come into the beam combiner with different polarizations. One example is the configuration used in the GI2T. Without correction the fringe pattern would be strongly perturbed, so this is generally corrected within the beam combiner. The corresponding problems have already been studied thoroughly elsewhere⁴ and so we will not treat them here.

In other interferometers the arms are designed to be identical in the sense that all the beam paths are symmetric with respect to the number of surfaces and their incidence angles. In this case, the states of polarization of the beams arriving from different unit telescopes will be perturbed by the same amount. Without any additional correction, high-contrast fringes will be observed.⁵ Among interferometers with symmetric arms are the VLTI,^{6,7} CHARA⁸ and the MROI.^{9,10}

If the polarization properties are the same for all arms, one might be led to believe polarization does not constitute a problem. However, we show later that this is only true if the source being observed is unpolarized. Since many sources are polarized on precisely the angular scales being investigated interferometrically, we need to consider the interaction between the polarization properties of the instrument and the polarization properties of the observed object. We show that, in general, sources with the same total intensity distribution but different inherent polarizations would give rise to different visibility measurements, even in a symmetric interferometer. As a consequence, given some visibility measurements of a source, there is an ambiguity between source morphology and source polarization structure. This will make scientific interpretation of the results difficult, especially if, as described above, the polarization properties of the interferometer optics changes with the hour angle of the object.

One way around this is to make several measurements with different polarizers in front of the beam combiner optics. Then, given some knowledge of the interferometer polarization properties, these measurements can be used to recover both the source intensity structure and the source polarization structure. The first successful use of this approach, called interferometric polarimetry,^{11,12} has been recently reported with the study of dust scattering in Mira-like variables.² However to perform such a measurement over a range of angles on the sky requires complex and potentially lossy interferometer optics¹³ and impacts the total time required to make a given observation.

1.1 Nomenclature

In this paper we consider an alternative strategy of constructing a “polarization fidelity” interferometer, by which we mean interferometer sensitive only to the intensity distribution and not to the polarization structure of the source. We will now briefly review the nomenclature of polarization in the context of an imaging system in order to precisely define what we mean by a polarization fidelity interferometer.

For a single-pixel detector, the Stokes vector S is defined from intensity measurement with different types of polarizers in front of the detector:

$$S = \begin{pmatrix} S_0 \\ S_1 \\ S_2 \\ S_3 \end{pmatrix} = \begin{pmatrix} I_0 \\ I_1 - I_0 \\ I_2 - I_0 \\ I_3 - I_0 \end{pmatrix}, \quad (1)$$

where I_0 measures the intensity with no polarizer present, I_1 is the intensity measured through an ideal horizontal linear polarizer, I_2 through a linear polarizer rotated by 45° , and I_3 through an ideal polarizer which lets through only left-handed circular polarization:

The single-pixel Stokes parameters can readily be generalized to a spatially-varying intensity distribution, called here an “image”, where $I_n(x, y)$ represents the intensity I_n measured at the pixel with coordinate (x, y) . Any incoherent partially polarized intensity distribution is then fully determined by the set of four images $[S_0(x, y), S_1(x, y), S_2(x, y), S_3(x, y)]$. An ideal polarimetric interferometer would allow measurement of any or all of these four images.

An alternative is to construct an interferometer to measure only $S_0(x, y)$, but to measure it independent of the values of $S_1(x, y)$, $S_2(x, y)$ and $S_3(x, y)$. This interferometer we choose to call a “polarization fidelity” interferometer. The problem explored here is what is required in the interferometer design to approach as closely as possible this ideal, i.e. to accurately measure visibilities corresponding to the unpolarized image while minimizing “crosstalk” from the polarized image structure. We use a simple model for the polarization properties of the interferometer optics and investigate how it interacts with the source polarization structure. We derive a key metric to assess the polarization fidelity of an interferometer and compare this metric in a few example designs for the unit telescopes. We assess how the differences in this metric between different designs would affect observations of a number of illustrative objects.

2. POLARIZATION FIDELITY IN A SYMMETRIC INTERFEROMETER

Let us consider an interferometer for which all interferometer beam paths are identical in terms of their polarization properties, i.e. all optical surfaces are of identical construction and all angles of incidence are identical. In the following we will not consider the effects of atmospheric turbulence (atmospheric piston and atmospheric speckle), so that the complex fringe visibility constitutes a good observable. We will also suppose that the minor atmospheric polarization effect induced by seeing¹⁴ is negligible.

A beam entering one unit telescope, travelling through the interferometer optics, and landing on the fringe detector undergoes a polarization state transformation as it passes through or reflects off each surface. The instantaneous polarization state of the beam can be described by a simple Jones vector $J = [E_x, E_y]$ and the polarization transformation at any surface can be modelled by a Jones matrix M_i for surface i which acts on the input Jones vector to produce an output Jones vector. The resulting Jones matrix for any given interferometric arm of the interferometer is $M = \prod_{i=1}^n M_i$ which describes the transformation of the light through the entire optical train to the detector.

For any optical system with Jones matrix M there will be a corresponding pair of orthonormal Jones vectors J_+ and J_- which represent characteristic polarization states. These states are eigenstates of the optical system, passing through the optics without any change in their state of polarization. In the general case they are elliptical polarization states. Let us call a_+ and a_- the corresponding scalar complex eigenvalues, which can be understood as polarization transfer coefficients ($|a_+|, |a_-| < 1$ as most systems are lossy) of the two eigenstates. As the different arms of the interferometer are identical, they share the same Jones matrix (we could multiply the Jones matrices by different complex scalars to encode any optical path differences between the arms, but choose not to do so here for simplicity) and hence the same characteristic polarization states.

We model the interferometer beam combiner as an idealized polarization-neutral device: it simply superposes the beams arriving from two or more arms of the interferometer without introducing any change in the state of polarisation of the beams. This ideal is not difficult to approach in practice, providing care is taken when designing the beam combiner to minimize the angles of incidence of the light beams on all optical surfaces. In order to produce a fringe pattern, the beam combiner incorporates a method of introducing a variable but polarization-independent phase difference between the beams. This phase difference θ is made to vary either spatially (for example using the geometrical variation in pathlength across a focal plane) or temporally (for example using a piezo-electrically actuated mirror) and the resultant fringe intensity I is measured as a function of θ .

We consider here only two-beam combination, but the results can readily be generalized to multiple-baseline beam combiners. The detected intensity can be written in terms of the amplitudes of the characteristic polarization states of the interferometer optics. Writing the Jones vectors of the instantaneous electric fields incident on each of the two unit telescopes in each of the two characteristic polarization states as $\vec{E}_+(1)$ and $\vec{E}_-(1)$ at telescope 1, and $\vec{E}_+(2)$ and $\vec{E}_-(2)$ at telescope 2 respectively, and assuming the arms are identical such that $a_+(1) = a_+(2) \equiv a_+$ and $a_-(1) = a_-(2) \equiv a_-$, we have the intensity detected at the beam combiner being given by:

$$\begin{aligned} I(\theta) &= \left\langle \left| a_+ \vec{E}_+(1) + a_- \vec{E}_-(1) + \exp(-i\theta) \left[a_+ \vec{E}_+(2) + a_- \vec{E}_-(2) \right] \right|^2 \right\rangle \\ &= |a_+|^2 \left[\langle |E_+(1)|^2 \rangle + \langle |E_+(2)|^2 \rangle + 2\Re \{ \langle E_+(1)E_+(2)^* \rangle \exp(i\theta) \} \right] \\ &\quad + |a_-|^2 \left[\langle |E_-(1)|^2 \rangle + \langle |E_-(2)|^2 \rangle + 2\Re \{ \langle E_-(1)E_-(2)^* \rangle \exp(i\theta) \} \right] , \end{aligned} \quad (2)$$

where $\langle \rangle$ indicates averaging over periods much longer than the coherence time of the radiation. We have made use of the fact that the two characteristic polarisation states are orthogonal to one another and therefore all the cross-polarization interference terms are zero. We note the important result that, with a symmetric optical system, only the moduli of the polarization transfer coefficients a_+ and a_- have any effect on the interference pattern, while the retardations between polarizations have no effect.

We can see that the detected interference pattern is simply a weighted superposition of the two fringe patterns which would be seen by an ideal interferometer (i.e. one which does not perturb the polarization states of the beams) when a (potentially very large!) polarizer selecting either the J_+ or the J_- state is inserted between the source and the interferometer. It is therefore helpful to split the object being observed into the two images as seen through these polarizers, denoted as $S_+(x, y)$ and $S_-(x, y)$ respectively. In general these are linear combinations of the S_0, S_1, S_2 and S_3 images.

The van Cittert-Zernike theorem relates the fringe pattern seen in an interferometer (or equivalently the spatial coherence function of the radiation) to the Fourier transform of the apparent object intensity distribution on the sky, giving:

$$\langle E_+(1)E_+(2)^* \rangle = \tilde{S}_+(u, v), \quad (3)$$

$$\langle E_-(1)E_-(2)^* \rangle = \tilde{S}_-(u, v), \quad (4)$$

where $\tilde{S}(u, v)$ is the Fourier component of $S(x, y)$ at the spatial frequency (u, v) corresponding to the vector baseline between the two unit telescopes. We note that the zero-spatial-frequency Fourier component corresponds to the total flux received by any one unit telescope:

$$\langle |E_+(1)|^2 \rangle = \langle |E_+(2)|^2 \rangle = \tilde{S}_+(0, 0) \equiv \tilde{S}_+(0), \quad (5)$$

$$\langle |E_-(1)|^2 \rangle = \langle |E_-(2)|^2 \rangle = \tilde{S}_-(0, 0) \equiv \tilde{S}_-(0) . \quad (6)$$

Combining equation 2 with equations 3, 4, 5, and 6, we can write the expression for the observed fringe pattern as

$$I(\theta) = I_0 + \Re \{ A(u, v) \exp(i\theta) \}, \quad (7)$$

where the constant ‘‘DC’’ component is given by

$$I_0 = 2 \left[|a_+|^2 \tilde{S}_+(0) + |a_-|^2 \tilde{S}_-(0) \right], \quad (8)$$

and the component oscillating sinusoidally with θ has a complex amplitude given by:

$$A(u, v) = 2 \left[|a_+|^2 \tilde{S}_+(u, v) + |a_-|^2 \tilde{S}_-(u, v) \right]. \quad (9)$$

It is conventional in optical interferometry to measure the complex visibility of the fringes, defined as the normalised complex fringe amplitude

$$V(u, v) \equiv A(u, v)/I_0. \quad (10)$$

Combining equations 10, 8, and 9, we get

$$V(u, v) = \frac{|a_+|^2 \tilde{S}_+(u, v) + |a_-|^2 \tilde{S}_-(u, v)}{|a_+|^2 \tilde{S}_+(0) + |a_-|^2 \tilde{S}_-(0)}. \quad (11)$$

In order to quantify the polarization fidelity of a given interferometer we express the measured visibility $V(u, v)$ relative to the visibility $V_0(u, v)$ which would be measured for the same source by an ideal (i.e. perfect polarization fidelity) interferometer:

$$\frac{V(u, v)}{V_0(u, v)} = \frac{|a_+|^2 \tilde{S}_+(u, v) + |a_-|^2 \tilde{S}_-(u, v)}{|a_+|^2 \tilde{S}_+(0) + |a_-|^2 \tilde{S}_-(0)} \cdot \frac{\tilde{S}_+(0) + \tilde{S}_-(0)}{\tilde{S}_+(u, v) + \tilde{S}_-(u, v)}. \quad (12)$$

We define the diattenuation D of an optical system as the fractional difference in transmission of the two states:

$$D = \frac{|a_+|^2 - |a_-|^2}{|a_+|^2 + |a_-|^2}, \quad (13)$$

where we have adopted the convention that $|a_+| \geq |a_-|$. A little algebraic manipulation yields the visibility relative to a perfect interferometer in terms of the diattenuation:

$$\frac{V(u, v)}{V_0(u, v)} = \frac{1 + D \left[\frac{\tilde{S}_+(u, v) - \tilde{S}_-(u, v)}{\tilde{S}_+(u, v) + \tilde{S}_-(u, v)} \right]}{1 + D \left[\frac{\tilde{S}_+(0) - \tilde{S}_-(0)}{\tilde{S}_+(0) + \tilde{S}_-(0)} \right]}. \quad (14)$$

The measured complex visibility can then be written as:

$$V(u, v) = \left(\frac{\tilde{S}_+(u, v) + \tilde{S}_-(u, v)}{\tilde{S}_+(0) + \tilde{S}_-(0)} \right) \left(\frac{1 + D \left[\frac{\tilde{S}_+(u, v) - \tilde{S}_-(u, v)}{\tilde{S}_+(u, v) + \tilde{S}_-(u, v)} \right]}{1 + D \left[\frac{\tilde{S}_+(0) - \tilde{S}_-(0)}{\tilde{S}_+(0) + \tilde{S}_-(0)} \right]} \right), \quad (15)$$

which simplifies into:

$$V(u, v) = \frac{(1 + D)\tilde{S}_+(u, v) + (1 - D)\tilde{S}_-(u, v)}{(1 + D)\tilde{S}_+(0) + (1 - D)\tilde{S}_-(0)}. \quad (16)$$

In this case, a convenient approximation of the absolute error due to polarization effects is given by:

$$V(u, v) - V_0(u, v) \simeq D \left[\frac{\tilde{S}_+(u, v) - \tilde{S}_-(u, v)}{\tilde{S}_+(0) + \tilde{S}_-(0)} \right], \quad (17)$$

where in this case we have assumed that $\tilde{S}_+(0) \simeq \tilde{S}_-(0)$, a condition which holds for most cases of astronomical interest.

Eq. (14) and Eq. (16) show that the polarization fidelity depends *both* on the source structure via the Stokes \tilde{S} terms, and on the instrumental and observational conditions via the diattenuation D . Moreover, the approximation given by Eq. (17) underlines that in most cases the absolute error introduced by polarization effects is roughly proportional to the diattenuation. In the following section we discuss these source-dependent and observation-dependent terms in turn.

3. APPLICATIONS

3.1 The source-dependent terms

In Eq. (14), the term $\left[\frac{\tilde{S}_+(u, v) - \tilde{S}_-(u, v)}{\tilde{S}_+(u, v) + \tilde{S}_-(u, v)} \right]$ can be seen as a Fourier component of the polarization difference image $[S_+(x, y) - S_-(x, y)]$, while $\left[\frac{\tilde{S}_+(u, v) + \tilde{S}_-(u, v)}{\tilde{S}_+(0) + \tilde{S}_-(0)} \right]$ is the Fourier component at the same frequency of the

unpolarized flux $S_0(x, y)$. Thus the ratio of the two, which appears in the numerator, can be thought of as a percentage polarization in the Fourier plane, i.e. at the angular resolution being probed by the interferometer baseline. The same expression, but evaluated for zero spatial frequency (i.e. the total flux), also appears in the denominator. This ratio, however, will typically be much smaller than that in the numerator because it corresponds to the polarization averaged over a much larger region of the sky.

For an unpolarized source we have $S_+(x, y) = S_-(x, y) = S_0(x, y)/2$ and so Eq. (14) shows that a visibility measurement gives $V(u, v) = V_0(u, v)$. This leads to the important result that for an unpolarized source, we measure the same visibility independent of the interferometer optics polarization properties (i.e. D), provided the optics in different interferometer arms have identical polarization properties. This is of particular relevance for a critical step of the interferometric measurement, the calibration of visibility measurements. To calibrate systematic errors in the interferometer, one generally relies on the measurement of sources of known coherence functions. Most calibrators will be normal stars and will not be resolved, so the calibrator polarization will be low (typically $\ll 1\%$). As a result, when deriving the effects of the interferometer polarization on any calibrated visibility measurements, we need only consider the polarization properties of the optics when observing the source (which potentially is significantly polarized) and not those when observing the calibrator.

The term in the denominator refers to the zero-spatial-frequency polarized flux, i.e. the polarization integrated over the seeing disk. For most astronomical sources, the polarization is at most a few percent at visible wavelengths. For a diattenuation of a few percent (see later), the denominator will be unity to better than a few parts in 10^4 , and in the following we will assume it to be unity. A few types of objects may however exhibit high degrees of polarization, such as dust enshrouded nebulae or stars. For example the polarization of IRC+10216, an AGB star with an large circumstellar envelope, reaches up to 50% far from the stellar core.¹⁵ As the diattenuation should probably stay below 20% for any well designed interferometer, the denominator remains lower than 1.1. In any case, the numerator will be the main factor determining the polarization fidelity of the instrument.

The percentage linear polarization term in the numerator depends on the polarization at high spatial frequencies. This is less easy to estimate *a priori* since such properties have not been measured extensively, as most polarimetric results have so far been obtained with non-interferometric instruments. Sources are expected to be more polarized on small angular scales than they are on larger scales, but little theoretical work has been done on the expected polarization structure on milliarcsecond scales. Several examples of sources will be discussed further in section 3.3.

3.2 The diattenuation term

In general the Jones matrix M of an interferometer will be rather complex to analyse. It would in principle be necessary to derive the two orthogonal – and in general, elliptical – eigenstates of polarization as a function of the pointing direction in the sky. However we can restrict analysis to a few simple cases indicative of the general behaviour. The most obvious simplification is to only analyse interferometer geometries for which linear polarization states are the eigenstates. This is particularly easy to interpret in term of the source polarization: for most astronomical observations at optical wavelengths (with a few rare exceptions such as magnetically active stars), the thermal radiation emitted by most sources is not appreciably polarized and so processes giving rise to linear polarization are dominant over processes giving rise to circular polarization.

If we suppose the beam relay system (by which we mean the arrangement of mirrors which directs the light from the unit telescopes to the point at which interference takes place) to be arranged such that all the mirror normals lie in a horizontal plane, then the beam relay eigenstates will be the horizontal and vertical linear polarization states. If in addition the unit telescopes are pointing in directions such that the S and P directions for any oblique reflections within the telescope are the same as (or perpendicular to) the S and P directions for the rest of the beam train, then the horizontal and vertical polarizations will be eigenstates for the entire optical system.

Once the eigenstates of the design are defined, the diattenuation can be evaluated or measured and eventually optimised. Several factors influence the diattenuation, namely the coatings of the optics, the wavelength of observation and the angles of incidence of the beams on the optics. In the discussion that follows we will assume the use of typical commercial mirror coatings, namely overcoated protected silver with three layers of SiO_2 , Al_2O_3

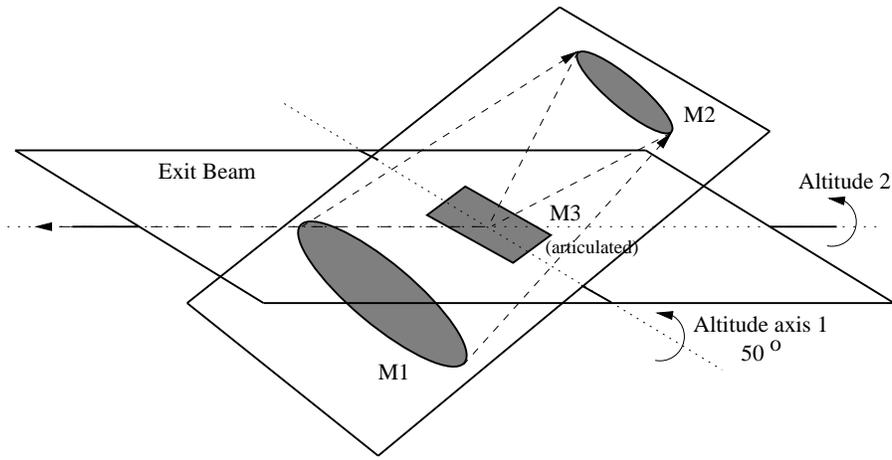


Figure 1. Cartoon of the three-mirror unit telescope optical train using an articulating tertiary referred to in the text. The output beam exits horizontally to the left. The geometry presenting the highest diattenuation is shown here.

and Ag in turn. For this coating the diattenuation decreases with increasing wavelength: broadly speaking at wavelengths longer than $1\mu\text{m}$, diattenuations are smaller than 1%, whereas at visible wavelengths values as large as several percent can be experienced. However this does not mean that polarization fidelity is less critical in the infrared: many important astrophysical processes inducing polarization take place in the infrared, and selection effects often favor their observable polarimetric signatures over the visible ones.

In general the induced diattenuation in an optical train will arise from any non-normal angle of incidence at reflecting surfaces in the train. As the angle of incidence is increased, the diattenuation will increase. Typically for the coatings considered here, the angle of incidence needs to be greater than approximately 45° for the diattenuation to rise above 1% in the visible. If the interferometer beam relay and combination optics have been designed to minimize such effects, then the angle of incidence of the starlight beam on the optics will be less or equal to 30° . For such angles the diattenuation of coated surfaces is small (less than 0.5%). In this case, only the oblique reflections in the unit telescopes (where the angles of incidence can easily exceed 30°) will contribute significantly to polarization effects.

We consider here two typical examples of optical trains in unit telescopes which output horizontal exit beams to the beam relay system. First let us consider an altitude over altitude telescope mount such as envisioned for the MROI^{9,10} which uses an optical train comprising three mirrors including an articulating tertiary. This is shown schematically in Fig. 1. This will give us an estimate of the polarization fidelity obtained with modern, efficient, optical trains. While more complex optical trains can be specifically designed to enhance polarimetric fidelity, these usually require more mirrors. In this case, other considerations not related to polarimetric performance have to be taken into account, including overall throughput, wavefront quality and mechanical design constraints.

Our second example is a more conventional seven-mirror ‘‘Coudé’’ train typical of that used in interferometric altitude over azimuth telescope mounts. This is shown schematically in Fig. 2. This is presented as an example of an optical train where polarization fidelity has not been prioritised during the design, and where several mirrors receive light at angles of incidence of 45° . Both optical layouts deliver horizontally oriented output beams fixed in space, and so the configuration of the remainder of the interferometer optics maintains the same geometric configuration for any pointing angle in the sky.

Simulations were carried with Zemax to determine the expected diattenuations for both the three-mirror and the seven-mirror Coudé design. We first determined the set of pointing angles for which the polarization eigenstates of the trains were linear, as well as horizontal and vertical in order to match the eigenstates of the beam relay. For the 3-mirror design this implies that the outer rotation axis must be rotated such that the inner axis lies horizontal, and the tertiary can be rotated by any angle about the inner axis. For the seven-mirror Coudé train, the altitude axis must be rotated such that the telescope elevation is 0 or 90° and the azimuth angle must be a multiple of 90° . Among those remaining possibilities, we selected the worst-case diattenuation figures as described below.

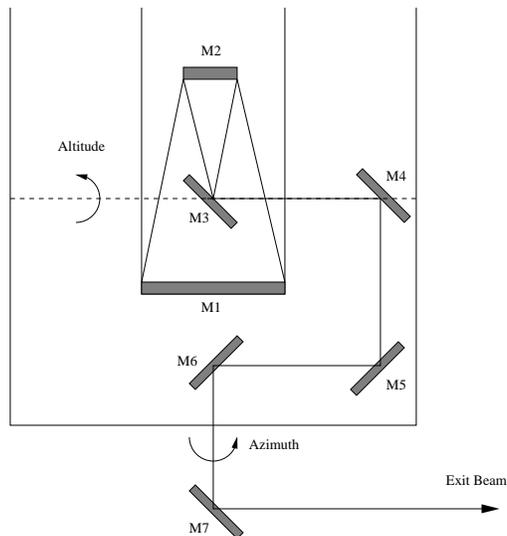


Figure 2. Cartoon of a seven-mirror Coudé unit telescope optical train, typical of several current interferometer implementations. The collimated output beam exits horizontally to the right. The configuration shown exhibits the highest diattenuation, and corresponds to an elevation angle of 90° and an azimuth angle of 0° .

For the three-mirror train, the elevation angle of the inner axis maximizes the diattenuation when the normal to the tertiary forms the maximum angle with regards to the exit beam, e.g. 40° for the typical case of the MROI. For wavelengths longer than about 800 nm, the diattenuation is still less than 0.5%. However for observations in the visible around 600 nm, the diattenuation rises to 1.5%. This number will be used as a worst-case diattenuation for illustrative purposes.

For the seven-mirror Coudé design, we can discard the configuration at 0° elevation as not being representative of real observations, and select an elevation angle of 90° . The highest diattenuation is then obtained for an azimuth angle of 0° . For sources observed at 600 nm each of the five mirrors of the Coudé train increases the diattenuation by approximately 1.5%, and so the total resulting diattenuation is about 8%.

3.3 Observation scenario

To examine the worst-case effects on the polarization fidelity let us use two simple astronomical models.

The first model is a simple star plus hotspot model, consisting of a uniform unpolarized stellar disk and an unresolved spot which emits 10% of the flux of the star. We assume the hotspot is linearly polarized in a direction which happens to match the J_+ characteristic polarization of the interferometer. The relative and absolute visibility errors given by Eq. 14 and Eq. 16 are shown as a function of the baseline length in Fig. 3. On short baselines, which do not resolve the disk, the error remains small and the measurement is virtually unperturbed by the polarization effects. On longer baselines where the disk is fully resolved, the visibility of the spot dominates, and so $S_-(u, v) \simeq 0$. The relative error fluctuates significantly and displays peaks when the source visibility is low. The level of diattenuation is clearly critical for low-visibility measurements, as the relative error rises up to 18% for the Coudé optical train, but only to 4% for the three-mirror one. On the baselines resolving the disk, the average relative errors in the measured visibility are about 1.5% (three-mirror) and 8% (Coudé). The fringe visibility there is of order 10%, so the absolute visibility errors are of order 0.15% and 0.80%, as shown in Fig. 3 (right). For the three-mirror optical train, the polarization effect is negligible here compared to other potential sources of visibility errors affecting the measurements. For the Coudé optical train these effects becomes significant when compared to the typical science goal of calibrating visibilities to better than 1%. They may even be the dominant error compared to atmospheric visibility calibration errors (about 1 – 5%).

A second, more extreme, source model consists of a binary star in which one star is linearly polarized in the J_+ direction and the other component is equally bright but polarized in the J_- direction. The evolution of the

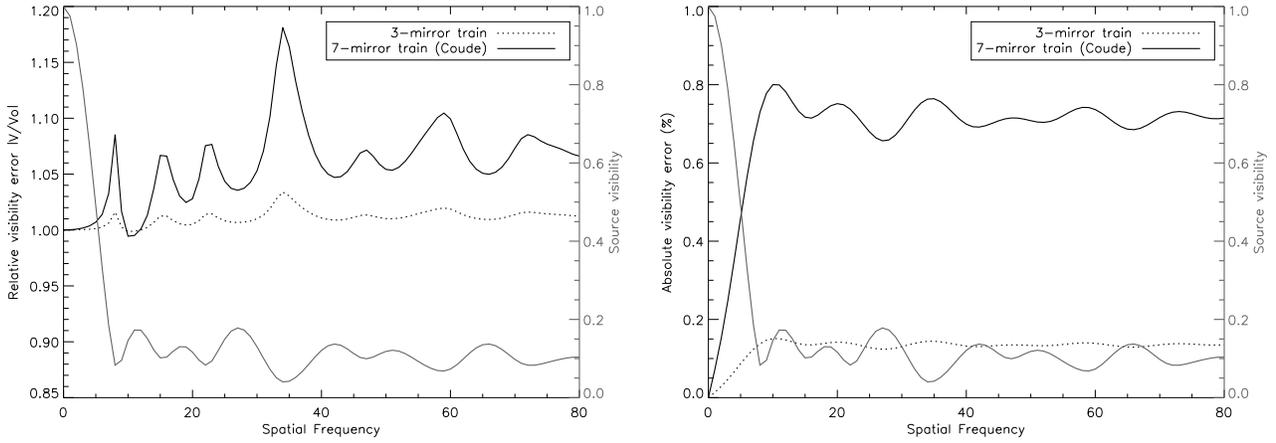


Figure 3. Simulations of visibility errors for a stellar disk plus hotspot model when observed with the three-mirror and seven-mirror designs. The visibility $|V_0|$ of the source is plotted along with the fractional error $|V/V_0|$ (left) and the absolute visibility errors (right) as a function of the spatial frequency. Note the reduction in fractional and absolute error in going from the seven-mirror to the three-mirror optical train design.

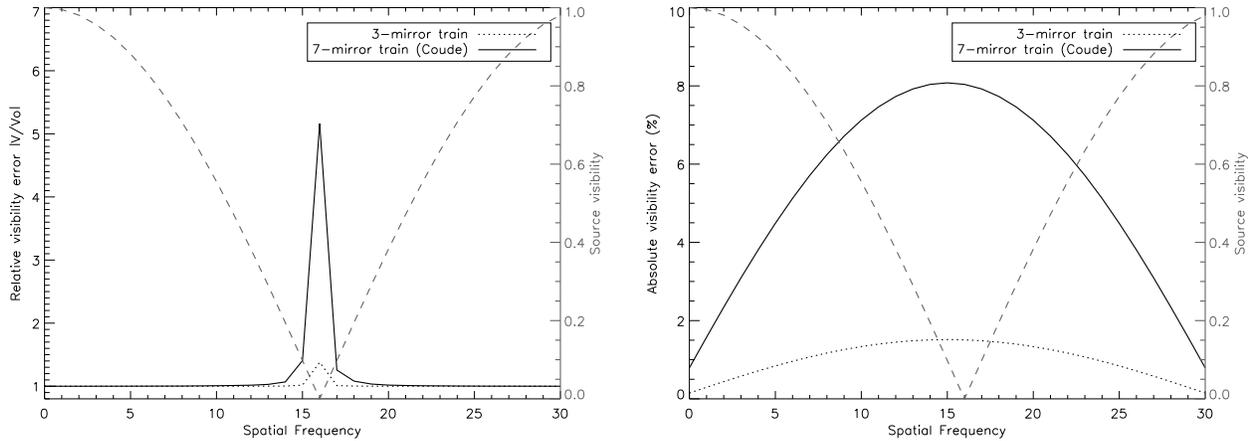


Figure 4. Simulations of visibility errors for a binary model when observed with the three-mirror and seven-mirror designs. The visibility $|V_0|$ of the source is plotted along with the fractional error $|V/V_0|$ (left) and the absolute visibility errors (right) as a function of the spatial frequency. Typical noise levels of 1% on visibilities have been added. Note the reduction in fractional and absolute error in going from the seven-mirror to the three-mirror optical train design.

relative visibility is presented in Fig. 4 (left). For each polarization direction, J_+ and J_- , a fringe pattern forms on the detector. The superposition of the two results in a blurred fringe pattern and a drop in visibility. In such a case the fringe contrast measured by a perfect polarization-fidelity interferometer will go to zero when the projected baseline is such that $\tilde{S}_+(u, v) = -\tilde{S}_-(u, v)$. On this baseline, the fractional visibility error of any imperfect interferometer would in theory then be infinite. In practise however, any error due to polarization leakage must be compared to other potential sources of error. Error sources such as photon noise give rise to measurement errors which are finite even when the fringe visibility is zero. In Fig. 4 (left), typical errors of 1% on the measured visibilities have been assumed and the peaks in the central parts of the curves show the difficulty of making accurate measurements in the presence of polarization-induced effects.

In contrast to the previous star plus hotspot model, here the absolute visibility errors due to polarization vary smoothly with the baseline. Fig. 4 (right) shows that the errors follow bell-shaped curves. As the assumption on which Eq. (17) is based (i.e. $\tilde{S}_+(0) = \tilde{S}_-(0)$) holds exactly in this case, the level of error is directly proportional to the diattenuation. The absolute visibility error rises up to a maximum of 1.5% for the three-mirror optical train, and 8% for the seven-mirror Coudé optical train. In practise this could amount to a huge difference in

scientific capability. While measurements with an interferometer based on the three-mirror design might remain usable (as stressed previously this level of random noise/error is not uncommon), the same measurements with the seven-mirror train would be susceptible to systematic errors that would probably render their scientific analysis extremely difficult. Admittedly, astronomical sources with this contrived polarization structure are rather unlikely to occur in nature, and so this second example can be seen as identifying a pessimistic lower bound to performance. Nevertheless both examples serve to demonstrate that the study of sources with unknown polarization structure can become difficult if the polarization fidelity of the interferometric train has not been optimized.

4. CONCLUSION

In an optical or infrared interferometer, the polarization states of the interfering beams are initially determined by the source but mainly affected by the oblique reflections within the optical train of each interferometer arm. As a discrepancy in the polarization states of the combined beams may result in a significant degradation of the visibility signal, most interferometers are now designed with symmetric arms/trains to allow all beams to be recombined in the same state.

We have demonstrated in this paper that this alone is not sufficient to resolve an additional problem: the emergence of an ambiguity between the source morphology and the source polarization structure in the scientific interpretation of the measurement.

Furthermore we have identified the diattenuation as the key metric for polarization fidelity in an interferometer, and we believe this should be considered critical for the design of modern interferometric optical trains. Limiting all angles of incidence to less than 30° is generally sufficient to keep the diattenuation small and hence minimise the final fringe visibility errors. However, this condition is usually violated in the beam relay trains conveying the light out of the unit telescopes in most modern interferometric implementations. Our simulations demonstrate that an optical train where polarization fidelity has not been prioritized may incur severe visibility errors due to polarization effects, rendering observation of polarized sources subject to significant ambiguity.

REFERENCES

- [1] Fosalba, P., Lazarian, A., Prunet, S., and Tauber, J. A., “Statistical Properties of Galactic Starlight Polarization,” *Astrophysical Journal* **564**, 762–772 (Jan. 2002).
- [2] Ireland, M. J., Tuthill, P. G., Davis, J., and Tango, W., “Dust scattering in the Miras R Car and RR Sco resolved by optical interferometric polarimetry,” *Monthly Notices of the Royal Astronomical Society* **361**, 337–344 (July 2005).
- [3] Perrin, G., Woillez, J., Lai, O., Guérin, J., Kotani, T., Wizinowich, P. L., Le Mignant, D., Hrynevych, M., Gathright, J., Léna, P., Chaffee, F., Vergnole, S., Delage, L., Reynaud, F., Adamson, A. J., Berthod, C., Brient, B., Collin, C., Crétenet, J., Dauny, F., Deléglise, C., Fédou, P., Goeltzenlichter, T., Guyon, O., Hulin, R., Marlot, C., Marteaud, M., Melse, B.-T., Nishikawa, J., Reess, J.-M., Ridgway, S. T., Rigaut, F., Roth, K., Tokunaga, A. T., and Ziegler, D., “Interferometric coupling of the Keck telescopes with single-mode fibers,” *Science* **311**, 194–+ (Jan. 2006).
- [4] Rousselet-Perraut, K., Vakili, F., and Mourard, D., “Polarization effects in stellar interferometry.,” *Optical Engineering* **35**, 2943–2955 (Oct. 1996).
- [5] Traub, W. A., “Polarization effects in stellar interferometers.,” in [*European Southern Observatory Astrophysics Symposia*], *European Southern Observatory Astrophysics Symposia* **29**, 1029–1038 (1988).
- [6] Beckers, J. M., “Polarization effects in astronomical spatial interferometry,” in [*Polarization considerations for optical systems II; Proceedings of the Meeting, San Diego, CA, Aug. 9-11, 1989 (A91-21262 07-74)*]. Bellingham, WA, *Society of Photo-Optical Instrumentation Engineers, 1990, p. 380-390.*], Chipman, R. A., ed., *Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference* **1166**, 380–390 (Jan. 1990).

- [7] Beckers, J. M., “The VLT interferometer. II - Factors affecting on-axis operation,” in [*Advanced technology optical telescopes IV; Proceedings of the Meeting, Tucson, AZ, Feb. 12-16, 1990. Part 1 (A91-23201 08-89)*]. Bellingham, WA, Society of Photo-Optical Instrumentation Engineers, 1990, p. 364-371.], Barr, L. D., ed., Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference **1236**, 364–371 (July 1990).
- [8] Sturmman, J., ten Brummelaar, T. A., Ridgway, S. T., Shure, M. A., Safizadeh, N., Sturmman, L., Turner, N. H., and McAlister, H. A., “Infrared beam combination at the CHARA array,” in [*Interferometry for Optical Astronomy II. Edited by Wesley A. Traub . Proceedings of the SPIE, Volume 4838, pp. 1208-1215 (2003)*]. Traub, W. A., ed., Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference **4838**, 1208–1215 (Feb. 2003).
- [9] Buscher, D. F., Bakker, E. J., Coleman, T. A., Creech-Eakman, M. J., Haniff, C. A., Jurgenson, C. A., KlingleSmith, III, D. A., Parameswariah, C. B., and Young, J. S., “The Magdalena Ridge Observatory Interferometer: a high-sensitivity imaging array,” in [*Unconventional Imaging II. Edited by Gamiz, Victor L.; Idell, Paul S.; Strojnik, Marija S.. Proceedings of the SPIE, Volume 6307, pp. 63070B (2006)*]. Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference **6307** (Aug. 2006).
- [10] Buscher, D. F., Boysen, R. C., Dace, R., Fisher, M., Haniff, C. A., Seneta, E. B., Sun, X., Wilson, D. M. A., and Young, J. S., “Design and testing of an innovative delay line for the MROI,” in [*Advances in Stellar Interferometry. Edited by Monnier, John D.; Schöller, Markus; Danchi, William C.. Proceedings of the SPIE, Volume 6268, pp. 62682I (2006)*]. Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference **6268** (July 2006).
- [11] Elias, II, N. M., “Optical Interferometric Polarimetry. I. Foundation,” *Astrophysical Journal* **549**, 647–668 (Mar. 2001).
- [12] Elias, II, N. M., “Optical Interferometric Polarimetry. II. Theory,” *Astrophysical Journal* **611**, 1175–1199 (Aug. 2004).
- [13] Tinbergen, J., “Polarization and optical aperture synthesis: the problem and a solution,” in [*Polarimetry in Astronomy. Edited by Silvano Fineschi . Proceedings of the SPIE, Volume 4843, pp. 122-136 (2003)*]. Fineschi, S., ed., Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference **4843**, 122–136 (Feb. 2003).
- [14] Sanchez Almeida, J. and Martinez Pillet, V., “Instrumental polarization in the focal plane of telescopes,” *Astronomy and Astrophysics* **260**, 543–555 (July 1992).
- [15] Ridgway, S. and Keady, J. J., “The IRC +10216 circumstellar envelope. II - Spatial measurements of the dust,” *Astrophysical Journal* **326**, 843–858 (Mar. 1988).