

# Minimizing the Effects of Polarization Crosstalk on the Imaging Fidelity of an Optical Interferometer

D. BUSCHER, F. BARON, AND C. HANIFF

Cavendish Laboratory, Department of Physics, University of Cambridge, JJ Thomson Avenue, Cambridge CB30HE, United Kingdom

*Received 2008 June 11; accepted 2008 December 18; published 2009 January 22*

**ABSTRACT.** Many astronomical objects are expected to be strongly polarized on the angular scales accessible with optical and infrared interferometry. Passage of stellar light through the optical trains of a long-baseline interferometer can induce cross talk between the polarized and unpolarized components of the light. As a result, the calibrated interferometric visibilities may depend on a difficult-to-separate mixture of the angular structure and the spatially varying polarization structure of the object being studied, and this will compromise the scientific usefulness of the interferometric data. We investigate the problem of designing a polarization-fidelity interferometer: one that can make accurate maps of the total intensity of an object, even when the object has a significant spatially varying polarized component. We demonstrate that taking polarization issues into account when designing the interferometric train is mandatory even when the interferometer has symmetric arms, and we identify that the key metric for such an interferometer is the diattenuation of the optical train. We evaluate the performance penalties incurred in an interferometer where polarization issues have not been adequately addressed.

## 1. INTRODUCTION

Optical interferometric arrays (in which we include arrays operating at visible and/or near-infrared wavelengths) are becoming a routine tool for astronomers studying processes occurring on angular scales tens or hundreds of times smaller than accessible with conventional telescopes. All existing and planned arrays have the same basic optical design: the radiation from a stellar object is collected at several locations by different telescopes, and the light is transported to a central location where the beams from different telescopes are combined to form interference fringes. Because of the transverse nature of electromagnetic waves, the parameters of the observed fringes depend on the polarization states of the interfering beams. The origin of the polarization in these beams is twofold: the polarization inherent to the source—indicating there is a physical phenomenon at work within the source polarizing the light emerging—and the polarization changes induced by the train of optical components in the interferometer that collects, transports, and combines the starlight. (We assume that there is no significant polarization of the signal during propagation to the Earth.) This paper examines the interaction between these two sources of polarization in the context of making accurate interferometric images.

Many astronomical objects (or “sources”) show polarization structure. At optical wavelengths, this structure most often 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, but we ignore

this case here. If the radiation source and the scattering medium have spherical symmetry, then the observed polarization pattern will have circular symmetry. If the polarized radiation is integrated over a large area, then the resulting total flux will be only weakly polarized.

On monolithic telescopes, conventional instruments such as IRPOL2 on the 3.4-m United Kingdom Infrared Telescope (UKIRT; Davis et al. 2005) have a resolution at best of  $0.1''$ . They are unable to resolve the stellar disk or most of the scattering envelopes around most stars. When observed with such telescopes, the typical intrinsic polarization shows levels of less than a few percent (Fosalba et al. 2002). Few polarimeters are used at very high angular resolution. The IR camera and spectrograph (IRCS) on the 8.2-m Subaru Telescope at Mauna Kea (Terada et al. 2004) is used with adaptive optics, but its resolution is still a few tens of milliarcseconds, much larger than most stars. On the other hand, interferometers make measurements on milliarcsecond scales where the polarization properties of the observed objects can be quite different. As the scattering environment can more often be resolved, much greater polarizations can be observed. For example, Ireland et al. (2005) measured the polarization from scattering in the envelopes of Mira-like variable stars at levels of order 15%.

The optical train of an interferometer will in general affect the state of polarization of the light passing through it. In fact in most existing interferometers, the dominant polarization effects arise from the reflections in the optical train. A quasi-monochromatic electromagnetic wave is conventionally modeled 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 632 nm incident at an angle of incidence of  $45^\circ$  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  $160^\circ$  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.

In some interferometers, single-mode optical fibers are used for beam transport and/or spatial filtering (Perrin et al. 2006). These are typically strongly birefringent devices that can therefore also affect the state of polarization of the starlight strongly.

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 variation will cause the polarization effects of each reflection to vary 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 Grand Interféromètre à 2 Télescopes (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 elsewhere (Rousselet-Perraut et al. 1996), 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 can be observed (Traub 1988). Among interferometers with symmetric arms are the Very Large Telescope Interferometer (VLTI; Beckers 1990a,

1990b), the Center for High Angular Resolution Astronomy Array (CHARA Array; Sturmman et al. 2003) and the Magdalena Ridge Observatory Interferometer (MROI; Buscher et al. 2006a, 2006b).

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. Because 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 will 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 ambiguity 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 difficulty 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 (Elias 2001, 2004), has been recently reported with the study of dust scattering in Mira-like variables (Ireland et al. 2005).

### 1.1. Nomenclature

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

For a single-pixel detector, the Stokes vector  $S$  is defined from intensity measurements 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  $45^\circ$ , and  $I_3$  through an ideal polarizer that 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. In all current interferometers, the combination of the telescope optics and the other interferometer optics apply a time-varying transformation to the incoming polarization so that, for example, linearly polarized light coming from a star could be received at the beam combiner as circularly polarized light at one point in time and at some time later as elliptically polarized light. This means that the design of interferometric optics to measure light from the object in a fixed polarization state or to perform the full observation of all Stokes images over a range of angles on the sky requires complex and potentially lossy interferometer optics (Tinbergen 2003).

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 cross talk 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 seeing (Sanchez Almeida and Martinez Pillet 1992) is negligible.

A beam entering one unit telescope, traveling 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 modeled 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  $J_-$  that 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 polarization 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, our model for the beam combiner incorporates a method of introducing a variable but polarization-independent phase difference between the beams. This phase difference  $\theta$  is made to vary either spatially (for example, using the geometrical variation in path length across a focal plane) or temporally (for example, using a piezoelectrically actuated mirror), and the resultant fringe intensity  $I$  is measured as a function of  $\theta$ .

We consider here only single-baseline beam combiner, 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  $\mathbf{E}_+(1)$  and  $\mathbf{E}_-(1)$  at telescope 1, and  $\mathbf{E}_+(2)$  and  $\mathbf{E}_-(2)$  at telescope 2, 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 the equation:

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

where the angle brackets indicate averaging over periods much longer than the coherence time of the radiation. We have made use of the fact that the two characteristic polarization 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 that would be seen by an ideal interferometer (i.e., one that 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:

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

and

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

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:

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

and

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

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

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

and the component oscillating sinusoidally with  $\theta$  has a complex amplitude given by

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

It is conventional in optical interferometry to measure the complex visibility of the fringes, defined as the normalized 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:

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

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:

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

The measured complex visibility can then be written as

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

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  $\tilde{S}_+(0) \approx \tilde{S}_-(0)$ , a condition that holds for most cases of astronomical interest.

Equations (14) and (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 equation (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 equation (14), the term  $[\tilde{S}_+(u, v) - \tilde{S}_-(u, v)]$  is a Fourier component of the polarization difference image  $[S_+(x, y) - S_-(x, y)]$ , while  $[\tilde{S}_+(u, v) + \tilde{S}_-(u, v)]$  is the Fourier component of the total flux  $S_0(x, y)$  at the same frequency. 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  $S_+(x, y) = S_-(x, y) = S_0(x, y)/2$ , so equation (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 § 3.2), 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% away from the stellar core (Ridgway & Keady 1988). As the diattenuation should probably stay below 20% for any well designed interferometer, the denominator remains over 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 because such properties have not been measured extensively, as most polarimetric results have so far been obtained with noninterferometric 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 milli-arcsecond scales. Several examples of sources will be discussed further in § 3.3.

#### 3.2. The Diattenuation Term

In general the Jones matrix  $M$  of an interferometer will be rather complex to analyze. 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 behavior. The most obvious simplification is to analyze only 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 that 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 optimized. 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 assume that the mirrors are coated with a typical commercially available protected silver coating, consisting of a layer of silver overcoated with a 170 nm layer of  $\text{SiO}_2$  and a 50 nm layer of  $\text{Al}_2\text{O}_3$ . This coating gives a level of diattenuation comparable to other silver-based or

aluminum-based coatings used in most optical and infrared interferometers. Custom coatings specifically designed to reduce diattenuation could be considered, but these are likely to be more costly and could sacrifice other desirable characteristics, e.g., high reflectivity over a broad wavelength range.

The diattenuation for this coating 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 nonnormal 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  $45^\circ$  for the diattenuation to rise above 1% in the visible.

If the optical train has not been designed to minimize the angles of incidence, then the polarization fidelity of the interferometer is likely to be low. In particular, nonnormal incidences on beam splitters inside the beam combiner could introduce large polarization effects.

If the interferometer beam relay and combination optics have been designed with polarization fidelity in mind, the angle of incidence of the starlight beam on the optics will be less than approximately  $30^\circ$ , and the diattenuation of coated surfaces is then expected to be less than 0.5%. This is the case for MROI, and we will also assume this to be true in the following discussion. Only the oblique reflections in the unit telescopes (where the angles of incidence can easily exceed  $30^\circ$ ) will then contribute significantly to polarization effects.

We consider here two typical examples of optical trains in unit telescopes that 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 (Buscher et al. 2006a, 2006b), which uses an optical train comprising three mirrors including an articulating tertiary. This first example is shown schematically in Figure 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 second example is shown schematically in Figure 2 and is presented as an example of an

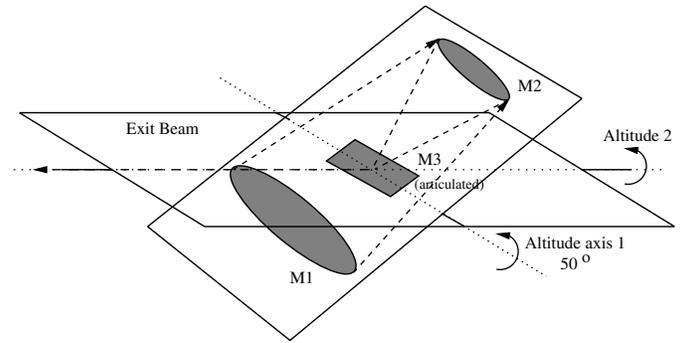


FIG. 1.—Cartoon of the three-mirror unit telescope optical train using an articulating tertiary referenced in the text. The output beam exits horizontally to the left. The geometry presenting the highest diattenuation is shown here.

optical train where polarization fidelity has not been prioritized during the design and where several mirrors receive light at angles of incidence of  $45^\circ$ . Both optical layouts deliver horizontally oriented output beams fixed in space, 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 out with Zemax to determine the expected diattenuations for both the three-mirror and the

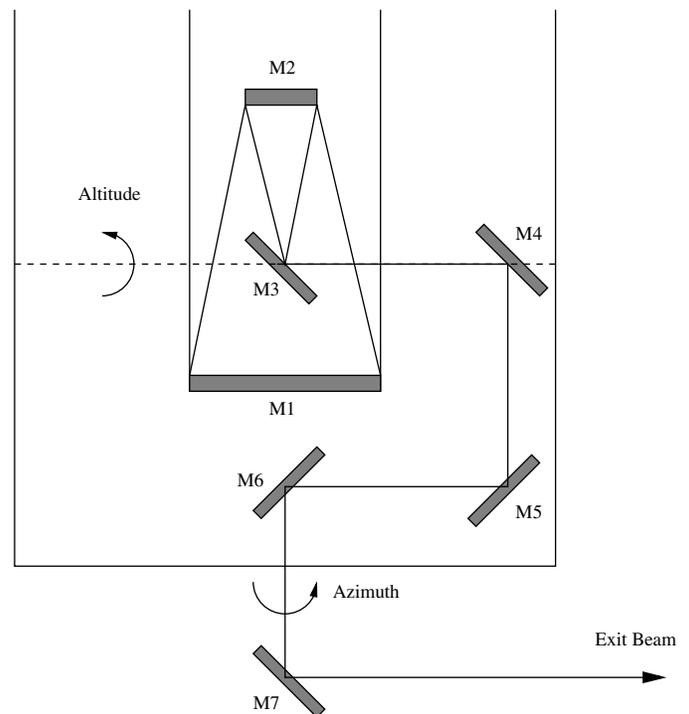


FIG. 2.—Cartoon of a seven-mirror coudé unit telescope optical train, typical of several current interferometer implementations. The collimated output beam exits horizontally to the right. The configuration shown exhibits the highest diattenuation, and corresponds to an elevation angle of  $90^\circ$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

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 three-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  $0^\circ$  or  $90^\circ$ , and the azimuth angle must be a multiple of  $90^\circ$ . Among those remaining possibilities, we selected the worst-case diattenuation figures as described below.

For the three-mirror train, the diattenuation is maximized when the normal to the tertiary forms the maximum angle with regards to the exit beam. In the case of the MROI, the telescope is designed to allow the observation of all sources at astronomical declinations of greater than  $-10^\circ$  for up to 3 hr after transit at the meridian. This constraint, when combined with the location of MROI (latitude 34N) and the telescope mount orientation chosen (the fixed outer axis lies approximately  $104^\circ$  East of North), results in a maximum angle of incidence on the tertiary mirror of  $65^\circ$ . At this angle of incidence and for wavelengths longer than about 800 nm, the maximum diattenuation remains under 1.5%. However, for observations in the visible around 600 nm, the diattenuation rises to 3%. We will use this worst-case diattenuation for illustrative purposes.

For the seven-mirror coudé design, we can discard the configuration at  $0^\circ$  elevation as not being representative of real observations and select an elevation angle of  $90^\circ$  and an azimuth angle of  $0^\circ$  (the incidence angle remaining  $45^\circ$ ). For sources observed at 600 nm each of the five mirrors of the coudé train increases the diattenuation by approximately 1.5% each, and so the total resulting diattenuation is about 8%.

We note that in both designs the effect of the other mirrors in the interferometer optical chain may be either to lessen or to increase the total diattenuation. It may be possible to design the interferometer so that the diattenuation of the beam relay optics cancels that of the unit telescopes for one particular orientation of the telescopes. However, for other orientations of the telescope, such a design may yield worse diattenuations than the design for the beam relay optics we have assumed here, namely that the beam relay diattenuation is minimal.

### 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 hot spot” model, consisting of a uniform unpolarized stellar disk and an unresolved spot that emits 10% of the flux of the star. We assume the hot spot is linearly polarized in a direction that happens to match the  $J_+$  characteristic polarization of the interferometer. The relative and absolute visibility errors given by equations (14) and (16) are shown as a function of the baseline length in Figure 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, so  $S_-(u, v) \approx 0$ . 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 7% for the three-mirror one. On the baselines resolving the disk, the average relative errors in the measured visibility are about 3% (three-mirror) and 8% (coudé). The fringe visibility there is of order 10%, so the absolute visibility errors are of order 0.3% and

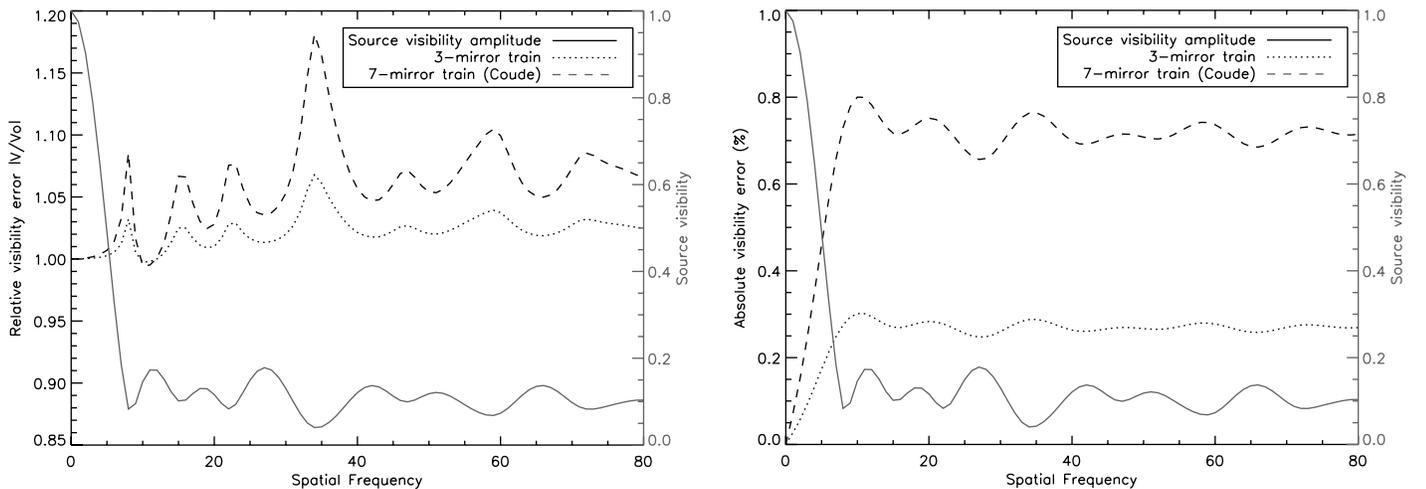


FIG. 3.—Simulations of visibility errors for a stellar disk plus hot spot model when observed with the three-mirror and seven-mirror designs. The visibility  $|V_0|$  of the source is plotted along with the fractional error  $|V/V_0|$  (left) and the absolute visibility errors (right) as a function of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

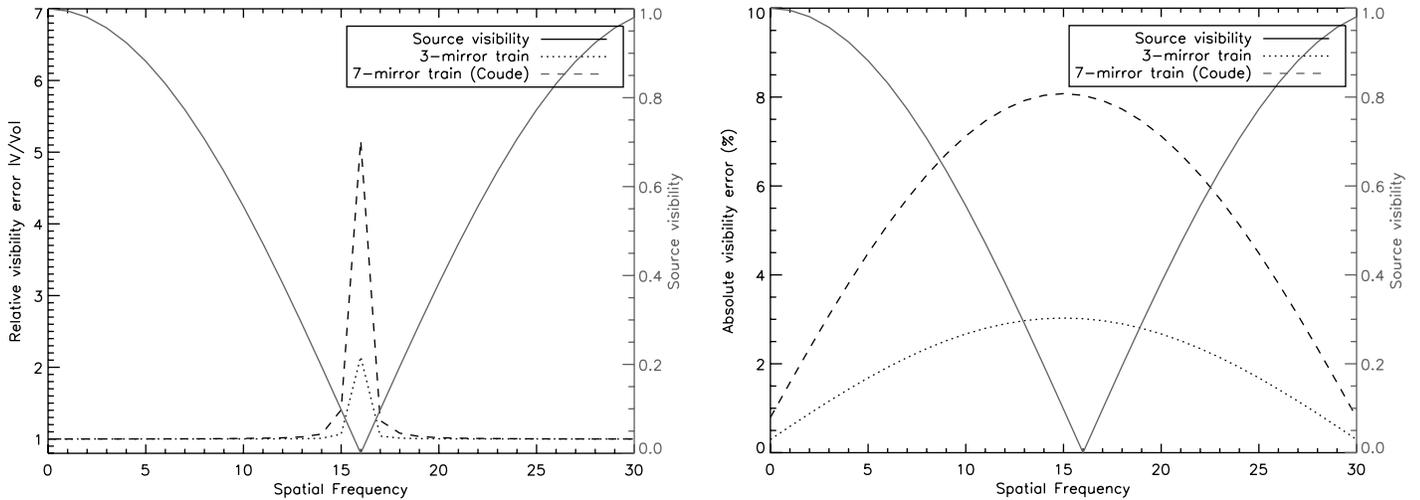


FIG. 4.—Simulations of visibility errors for an extreme binary model when observed with the three-mirror and seven-mirror designs. The visibility  $|V_0|$  of the source is plotted along with the fractional error  $|V/V_0|$  (left) and the absolute visibility errors (right) as a function of the spatial frequency. Typical noise levels of 1% on the 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.

0.8%, as shown in Figure 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 become 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 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. While unrealistic, this is a useful example for illustration purpose. The evolution of the relative visibility is presented in Figure 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 practice, 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 that are finite even when the fringe visibility is zero. In Figure 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 hot spot model, here the absolute visibility errors due to polarization vary smoothly with the baseline. Figure 4 (right) shows that the errors follow bell-shaped curves. As the assumption on which equation (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 3% for the three-mirror optical train, and 8% for the seven-mirror coudé optical train. In practice 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 error due to random noise 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 optical and infrared interferometry, a discrepancy in the polarization states of the combined beams may result both in a significant degradation of the visibility signal and in the emergence of an ambiguity between the source morphology and the source polarization structure in the scientific interpretation of the measurement.

The polarization states of the beams, initially determined by the source, are mainly affected by the oblique reflections within the optical train of each interferometer arm. While designing an interferometer with symmetric arms or trains partially solves the first problem by allowing all beams to be recombined in

the same polarization state, we have demonstrated in this paper that this alone is not sufficient to resolve the measurement ambiguity.

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  $30^\circ$  is generally sufficient to keep the diattenuation small and

hence minimize the final fringe visibility errors. However, in most modern interferometric implementations, this condition is usually violated both in the unit telescopes and in the associated beam relay trains conveying the light out. 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

- Beckers, J. M. 1990a, in Proc. SPIaE Conf., 1989 Aug. 9–11, Polarization Considerations for Optical Systems II, ed. R. A. Chipman, vol. 1166 (A91-21262 07-74; Bellingham, WA: SPIaE), 380–390
- . 1990b, Proc. SPIE, 1236, 364–371 (A91-23201 08-89)
- Buscher, D. F., et al. 2006a, Proc SPIE, 6307, 63070B doi: 10.1117/12.684422
- . 2006b, Proc. SPIE, Advances in Stellar Interferometry, 6268, 62682I doi: 10.1117/12.671915
- Davis, C. J., Chrysostomou, A., & Hough, J. 2005, in ASP Conf. Ser. 343, Astronomical Polarimetry: Current Status and Future Directions, ed. A. Adamson, C. Aspin, C. Davis, & T. Fujiyoshi (San Francisco: ASP), 79
- Elias, N. M., II 2001, ApJ, 549, 647–668 doi: 10.1086/319046
- . 2004, ApJ, 611, 1175–1199 doi: 10.1086/422212
- Fosalba, P., Lazarian, A., Prunet, S., & Tauber, J. A. 2002, ApJ, 564, 762–772 doi: 10.1086/324297
- Ireland, M. J., Tuthill, P. G., Davis, J., & Tango, W. 2005, MNRAS, 361, 337–344 doi: 10.1111/j.1365-2966.2005.09181.x
- Perrin, G., et al. 2006, Science, 311, 194
- Ridgway, S., & Keady, J. J. 1988, ApJ, 326, 843–858 doi: 10.1086/166142
- Rousselet-Perraut, K., Vakili, F., & Mourard, D. 1996, Opt. Eng., 35, 2943–2955
- Sanchez Almeida, J., & Martinez Pillet, V. 1992, A&A, 260, 543–555
- Sturmann, J., ten Brummelaar, T. A., Ridgway, S. T., Shure, M. A., Safizadeh, N., Sturmann, L., Turner, N. H., & McAlister, H. A. 2003, Proc. SPIE, 4838, 1208–1215
- Terada, H., Kobayashi, N., Tokunaga, A. T., Pyo, T.-S., Nedachi, K., Weber, M., Potter, R., & Onaka, P. M. 2004, Proc. SPIE, 5492, 1542–1550, doi: 10.1117/12.552516
- Tinbergen, J. 2003, Proc. SPIE, 4843, 122–136
- Traub, W. A. 1988, in ESO Astrophysics Symposium 29, 1029–1038