

Stellar Interferometric Beam Combiners in the Context of Linear Optics Networks

C. Schmid^a

^aEuropean Southern Observatory (ESO), Karl-Schwarzschild-Strasse 2 , D-85748 Garching, Germany;

ABSTRACT

The performance of ground based optical stellar interferometers suffers from various atmospheric as well as instrumental disturbances. The mathematical modeling of any kind of disturbances is essential in order to properly design an interferometric instrument and to get realistic estimates of its envisaged performance. One of the instrumental disturbances that has recently gained particular interest is changes to the polarization state of the incoming stellar light. Conventional methods of modeling polarization of light, such as, e.g., the Mueller calculus, are usually separate from or outside the context of interference.

In this contribution I introduce to stellar interferometry a mathematical formalism that is commonly used in the field of optical Quantum Information Science. I show that this formalism allows an elegant and easy-to-use integral treatment of polarization *and* interference for any beam combination scheme based on linear optical elements. In order to illustrate the formalism I apply it exemplarily to the modeling of some existing beam combiner concepts.

Keywords: Astrometric and interferometric instruments, Interferometry

1. INTRODUCTION

The performance of ground based optical stellar interferometers suffers from various atmospheric as well as instrumental disturbances. One of the instrumental disturbances that has recently gained particular interest in the context of ESO's Phase Referenced Imaging and Micro Arcsecond Astrometry (PRIMA) facility¹ is changes to the polarization state of the incoming stellar light.² Once more it turned out that the mathematical modeling of polarization effects in stellar interferometry is essential in order to properly design an interferometric instrument and to get realistic estimates of its envisaged performance. Well established, conventional methods of modeling polarization of light, such as, e.g., the Mueller calculus, are usually separate from or outside the context of interference. Work has been done to treat polarization and interference effects at the same time³ in the framework of Optical Interferometric Polarimetry (OIP).⁴

Usually, in astronomy interference is understood in terms of classical physics dealing with beams of light (see e.g., Ref. 5). Therefore – to the best of my knowledge – any attempt for a combined modeling of polarization and interference in astronomy such as OIP relies on concepts of classical physics. However, in view of quantum physics stellar interferometry can be understood on a single photon level as well. A quantum mechanical description of interference does not only allow a different view on the underlying physics, but can furthermore provide a practical mathematical formalism to treat certain problems as well as an astrophysical exploitation of other quantum properties of light.

In this contribution I want to introduce to stellar interferometry the mathematical formalism that is commonly used in the field of optical Quantum Information Science (QIS). This branch of physics is a quantum theory of information that encodes information in single photons and uses quantum effects to enhance conventional information processing. For the encoding of information various experimental implementations exist, and many of them use either very well defined spatial modes or polarization states of photons (for extensive reviews on the topic see e.g. Refs. 6–8).

Further author information: (Send correspondence to C.S.)
C.S.: E-mail: cschmid@eso.org, Telephone: +49 (0)89 3200 6467

Given the fact that more and more beam combination schemes in stellar interferometry rely on spatially single mode wave guides (either in optical fibers⁹ or integrated optics chips¹⁰), and are affected by polarization effects (either in a positive¹¹ or negative way¹²), it seems natural to make use of the formalism developed in optical QIS to model interference and polarization in stellar beam combiners.

In the following, I will first start to introduce the formalism, and then exemplarily apply it to the modeling of some existing beam combiner concepts.

2. FORMALISM

In this section I introduce some fundamental mathematical concepts of quantum mechanics. Readers who are familiar with the mathematics of quantum mechanics and, in particular, the Dirac notation might just quickly browse through the notation subsection 2.1 and proceed directly to the application section 3.

2.1 Notation

Spatial Modes: Although the formalism can be in principle applied to interferometers with an arbitrary number of telescopes,¹³ I will consider here only the case of two telescope beam combination. Consequently I distinguish between two spatial modes a and b corresponding to the two arms of the interferometer. Further, I use the well-known Dirac notation of quantum mechanics: The symbol $|a\rangle$ (or $|b\rangle$) shall be understood as vectors of a two-dimensional Hilbert space, \mathbb{H}_2 , describing the state of a single photon propagating in the spatial mode a (or b , respectively).

As all treated optical elements are restricted to be linear it is convenient to employ a matrix formalism with a representation chosen as:

$$|a\rangle \hat{=} \begin{pmatrix} 1 \\ 0 \end{pmatrix}_{\text{spa}}, \quad |b\rangle \hat{=} \begin{pmatrix} 0 \\ 1 \end{pmatrix}_{\text{spa}} \quad (1)$$

The subscript "spa" refers to "spatial" in order to distinguish between spatial and polarization modes.

Polarization Modes: For the representation of any possible polarization, I choose as basis vectors the photon states of horizontal, h , and vertical, v , polarization. Dirac notation is also used for the polarization in combination with a proper matrix representation:

$$|h\rangle \hat{=} \begin{pmatrix} 1 \\ 0 \end{pmatrix}_{\text{pol}}, \quad |v\rangle \hat{=} \begin{pmatrix} 0 \\ 1 \end{pmatrix}_{\text{pol}}. \quad (2)$$

Analogous to before, the subscript "pol" refers to "polarization" for distinction.

Note that the polarization part of the formalism with this matrix representation is mathematically equivalent to what is known as the Jones formalism in classical polarization optics.

Combined spatial and polarization modes: As this contribution aims at the combined description of spatial interference and polarization effects, a set of basis vectors for the full state space, $\mathbb{H}_2 \otimes \mathbb{H}_2$, is needed. These basis states for the combined states of both modes can be written in terms of tensor products:

$$|a\rangle \otimes |h\rangle = |ah\rangle, \quad (3a)$$

$$|a\rangle \otimes |v\rangle = |av\rangle, \quad (3b)$$

$$|b\rangle \otimes |h\rangle = |bh\rangle, \quad (3c)$$

$$|b\rangle \otimes |v\rangle = |bv\rangle, \quad (3d)$$

where the operator " \otimes " denotes the tensor product.

In the chosen matrix representation this leads to:

$$|ah\rangle \hat{=} \begin{pmatrix} 1 \\ 0 \end{pmatrix}_{\text{spa}} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix}_{\text{pol}} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad (4a)$$

$$|av\rangle \hat{=} \begin{pmatrix} 1 \\ 0 \end{pmatrix}_{\text{spa}} \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix}_{\text{pol}} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad (4b)$$

$$|bh\rangle \hat{=} \begin{pmatrix} 0 \\ 1 \end{pmatrix}_{\text{spa}} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix}_{\text{pol}} = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad (4c)$$

$$|bv\rangle \hat{=} \begin{pmatrix} 0 \\ 1 \end{pmatrix}_{\text{spa}} \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix}_{\text{pol}} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}. \quad (4d)$$

2.2 Pure vs. Mixed States

The four vectors of Eqns. 4 form an orthonormal basis system for all possible input states of the interferometer. From a physical point of view it seems reasonable to assume that, at the entrance of the interferometer, the polarization and spatial modes are independent from each other. Therefore, an arbitrary pure input state can be considered as a product state (not as an entangled state between spatial and polarization mode) and could thus be written as

$$|\text{in}\rangle = (t_{\text{spa}}|a\rangle + r_{\text{spa}}e^{i\varphi_{\text{spa}}}|b\rangle) \otimes (t_{\text{pol}}|h\rangle + r_{\text{pol}}e^{i\varphi_{\text{pol}}}|v\rangle) \quad (5)$$

with $t_{\text{spa}}, r_{\text{spa}}, \varphi_{\text{spa}}, t_{\text{pol}}, r_{\text{pol}}, \varphi_{\text{pol}} \in \mathbb{R}$ and $t_{\text{spa/pol}}^2 + r_{\text{spa/pol}}^2 = 1$, $\varphi_{\text{spa/pol}} \in [0, 2\pi]$.

Pure states are used for the description of closed systems in which all information about the state is available.

In contrast, imagine that due to some interaction with the environment, which leads to decoherence or depolarization, only partial information about the state is available. This applies indeed to the case of stellar interferometry as starlight might be often un- or partially-polarized and might have finite spatial coherence. In this case one has to deal with open systems in which states cannot be properly described by vectors but are best represented as statistical mixtures. To this end, the formalism offers the very useful concept of so-called "mixed states" as a kind of statistical treatment. A mixed state is represented as a density operator, ϱ , which is a Hermitian, positive-semidefinite operator of trace one. In the finite dimensional cases presented here it can be written as

$$\varrho = \sum_i a_i |\alpha_i\rangle\langle\alpha_i|, \quad (6)$$

with $a_i \geq 0$, $\sum_i a_i = 1$ being the probability to find the photon in the pure state $|\alpha_i\rangle$. The symbol $\langle\alpha_i|$ denotes the dual vector to $|\alpha_i\rangle$ such that the inner product $\langle\alpha_i|\alpha_i\rangle = 1$ and $|\alpha_i\rangle\langle\alpha_i|$ can be perceived as a projection operator onto the state $|\alpha_i\rangle$.

For example, a fully unpolarized state ϱ_{pol} is given as

$$\varrho_{\text{pol}} = \frac{1}{2}|h\rangle\langle h| + \frac{1}{2}|v\rangle\langle v| = \frac{1}{2}\mathbb{1} \quad (7)$$

which is not a vector but an operator or matrix, respectively.

Readers familiar with classical methods of polarization description will realize that the density operator is the equivalent of the coherence matrix in the Stokes formalism.

2.3 Propagation Through the Optics

Within the formalism, the transformation applied by optics of the interferometer (or the beam combiner) to the input state is described as a sequence of transformations M_i . The output state after propagation through all the optical elements is then given as

$$|\text{out}\rangle = M |\text{in}\rangle \quad (8)$$

where, $M = \prod_i M_i$ is the product of all transformations, M_i , that describe the different individual optical elements acting either on the spatial mode, the polarization mode, or both.

In the case of mixed states the output state of the interferometer ϱ_{out} for a given input state ϱ_{in} is analogously obtained as

$$\varrho_{\text{out}} = M \varrho_{\text{in}} M^\dagger, \quad (9)$$

where "†" means transposition and complex conjugation.

In Sec. 2.5 I will introduce the explicit matrix form for transformations of some typical optical elements.

2.4 Observables and detection probabilities

Once the output state is obtained it can be used to predict the measurement results for particular observables. An observable is represented by an Hermitian operator A . Its expectation value, i.e. its predicted mean measurement outcome, in the pure state $|\psi\rangle$ or in the mixed state ϱ is given as

$$\langle A \rangle_\psi \equiv \langle \psi | A | \psi \rangle \quad \text{or} \quad \langle A \rangle_\varrho \equiv \text{tr}(A\varrho), \quad (10)$$

respectively.

In the context of this work, measurements of observables are often projective measurements as the photon is detected behind a single-mode waveguide or a polarizer. In this case the expectation value denotes simply the probability to detect the photon (e.g. behind the polarizer or the wave-guide). In turn, these probabilities can be used to estimate the astronomically relevant quantities such as the complex visibility*.

In the two dimensional case (two spatial modes or polarization), it is convenient to define and express an arbitrary, general observable (for polarization or spatial mode), $\sigma_{\theta,\varphi}$, in terms of the well-known Pauli operators, $\sigma_x, \sigma_y, \sigma_z$ as

$$\sigma_{\theta,\varphi} = \vec{r} \cdot \vec{\sigma}, \quad (11)$$

with $\vec{r} = (\cos(\varphi_{s/p}) \sin(\theta_{s/p}), \sin(\varphi_{s/p}) \sin(\theta_{s/p}), \cos(\theta_{s/p}))$, $\theta_{s/p} \in [0, \pi]$, $\varphi_{s/p} \in [0, 2\pi[$ and $\vec{\sigma} = \begin{pmatrix} \sigma_x \\ \sigma_y \\ \sigma_z \end{pmatrix}$.

The projector $P_{\theta,\varphi}$ on each of the eigenvectors of $\sigma_{\theta,\varphi}$ with positive or negative eigenvalue, respectively, is then expressed as

$$P_{\theta,\varphi} = \frac{1}{2} (\mathbf{1} + \sigma_{\theta,\varphi}). \quad (12)$$

To illustrate this with an example, a projective measurement of vertical polarization, e.g. by a polarizer, corresponds to $P_{\pi,0} = \frac{1}{2} (\mathbf{1} - \sigma_z)$.

This representation of two dimensional quantum observables in terms of Pauli operators has also an analogy to classical polarization optics. It is similar to the representation of the Stokes parameters on the Poincaré sphere. However, attention has to be paid as the chosen spherical parametrization, $\theta_{s/p}$ and $\varphi_{s/p}$, is not identical to the one of the Poincaré sphere, usually referred to as "orientation" and "ellipticity" angles.

To the best of my knowledge, the difference of parametrization is purely historical. The one chosen in quantum information science stems from the Bloch sphere which was originally used to describe spin-half quantum systems.

Mathematically, both, the Bloch and the Poincaré sphere representation, rely on the homomorphy between the symmetry groups $SU(2)$ and $O(3)$.

*Astronomers usually talk about fluxes. However, one has to keep in mind that the flux on the detector is nothing but the product of the detection probability behind the beam combiner times the incoming number of photons. Consequently, the detection probabilities to find a photon in the particular outputs of the beam combiner are ultimately the "observables" of interest.

2.5 Useful operators

In this section I introduce the matrix notation of some useful operators that will be used in Sec. 3. As already mentioned before, for the polarization mode these matrices are equivalent to the ones known from the Jones formalism.

Phase shifter This operator can act on the polarization or the spatial mode and applies a relative phase-shift of $\phi \in \mathbb{R}$:

$$PS(\phi) = \begin{pmatrix} 1 & 0 \\ 0 & \exp(-i\phi) \end{pmatrix} \quad (13)$$

General linear retarder This operator acts on the polarization mode and is also known as general wave-plate with retardation δ and angle of the axis ρ along which the retardation occurs:

$$G_{LR}(\delta, \rho) = \exp(-i\frac{\delta}{2}) \cdot \begin{pmatrix} \exp(i\frac{\delta}{2}) \cos^2(\rho) + \exp(-i\frac{\delta}{2}) \sin^2(\rho) & 2i \cos(\rho) \sin(\frac{\delta}{2}) \sin(\rho) \\ 2i \cos(\rho) \sin(\frac{\delta}{2}) \sin(\rho) & \exp(-i\frac{\delta}{2}) \cos^2(\rho) + \exp(i\frac{\delta}{2}) \sin^2(\rho) \end{pmatrix} \quad (14)$$

General beam splitter This operator transforms spatial modes, see Fig. 1. It represents a beam splitter with transmission and reflection amplitudes $t_{a/b}, r_{a/b} \in \mathbb{R}$ and phase-shifts $\varphi_{a/b} \in \mathbb{R}$:

$$BS = \begin{pmatrix} t_a & e^{i\varphi_b} r_b \\ e^{i\varphi_a} r_a & t_b \end{pmatrix} \quad (15)$$

In general, the parameters can be different for the two polarization modes such that the beam splitter becomes an operator acting on both, polarization and spatial modes,

$$G_{BS} = BS_h \otimes |h\rangle\langle h| + BS_v \otimes |v\rangle\langle v|. \quad (16)$$

With $t_l, r_l \in \mathbb{R}$ ($l \in \{ah, av, bh, bv\}$) and phase-shifts $\varphi_l \in \mathbb{R}$ this leads to:

$$\begin{aligned} G_{BS} &= \begin{pmatrix} t_{ah} & e^{i\varphi_{bh}} r_{bh} \\ e^{i\varphi_{ah}} r_{ah} & t_{bh} \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} t_{av} & e^{i\varphi_{bv}} r_{bv} \\ e^{i\varphi_{av}} r_{av} & t_{bv} \end{pmatrix} \otimes \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} t_{ah} & 0 & e^{i\varphi_{bh}} r_{bh} & 0 \\ 0 & t_{av} & 0 & e^{i\varphi_{bv}} r_{bv} \\ e^{i\varphi_{ah}} r_{ah} & 0 & t_{bh} & 0 \\ 0 & e^{i\varphi_{av}} r_{av} & 0 & t_{bv} \end{pmatrix}. \end{aligned} \quad (17)$$

3. APPLICATION TO BEAM COMBINERS

3.1 A General Model

I am using a rather general model of a two telescope beam combiner which is depicted in Fig. 2. The phase resulting from the Optical Path Difference (OPD) between the two arms of the interferometer is represented by a phase shifter $PS(\phi)$ which applies a relative phase-shift of ϕ between spatial mode a and b . In order to account for polarization changes in each spatial mode, general linear retarders G_{LR} with retardations δ_1, δ_2 and orientations of their axes of, ρ_1, ρ_2 are used. The beam combination itself happens at a general beam splitter G_{BS} with phase and amplitude coefficients as defined in Eqn. 17.

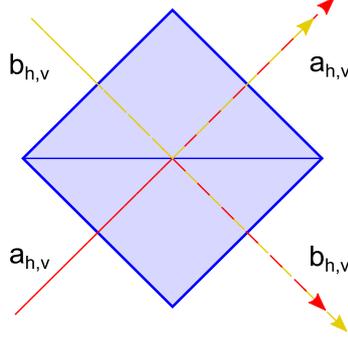


Figure 1: Beam splitter acting on spatial and polarization modes.

As input state, ϱ_{in} , I assume a state with spatiotemporal degree of coherence $\mathcal{V} \in [0, 1]^\dagger$ and degree of polarization $\mathcal{P} \in [0, 1]$:

$$\varrho_{\text{in}} = \varrho_{\text{in-s}} \otimes \varrho_{\text{in-p}}, \quad (18)$$

with

$$\begin{aligned} \varrho_{\text{in-s}} &= \mathcal{V} \left(\frac{1}{2} (|a\rangle\langle a| + |a\rangle\langle b| + |b\rangle\langle a| + |b\rangle\langle b|) \right) \\ &+ (1 - \mathcal{V}) \left(\frac{1}{2} (|a\rangle\langle a| + |b\rangle\langle b|) \right) \end{aligned} \quad (19)$$

$$\varrho_{\text{in-p}} = \mathcal{P} |p\rangle\langle p| + (1 - \mathcal{P}) \left(\frac{1}{2} (|h\rangle\langle h| + |v\rangle\langle v|) \right).$$

Obviously for $\mathcal{P} = 0$ the state is completely unpolarized, and for $\mathcal{P} = 1$ the state is fully polarized. $|p\rangle$ can be any polarization state and parameterized as

$$|p\rangle = \cos\left(\frac{\theta_{\text{pol}}}{2}\right) |h\rangle + e^{i\varphi_{\text{pol}}} \sin\left(\frac{\theta_{\text{pol}}}{2}\right) |v\rangle,$$

where I have chosen $t_{\text{pol}} = \cos\left(\frac{\theta_{\text{pol}}}{2}\right)$ and $r_{\text{pol}} = \sin\left(\frac{\theta_{\text{pol}}}{2}\right)$, compare Eqn. 5.

The output state ϱ_{out} is given as

$$\varrho_{\text{out}} = M \varrho_{\text{in}} M^\dagger, \quad (20)$$

with

$$M = G_{BS} \cdot \left(|a\rangle\langle a| \otimes G_{LR}(\delta_1, \rho_1) + |b\rangle\langle b| \otimes G_{LR}(\delta_2, \rho_2) \right) \cdot \left(PS(\phi) \otimes \mathbf{1} \right). \quad (21)$$

Starting from the output state, Eqn. 20, it is straightforward to derive the probabilities to detect a photon

[†]In this contribution I treat only the monochromatic case. However, chromatic effects might be introduced via the visibility, for example in the form of $\mathcal{V} = \mathcal{V}_0(\vec{B}, \lambda) \cdot \text{sinc}^2\left(\frac{\lambda}{2\pi w} OPD\right)$ for a rectangular band-pass, with λ being the central wavelength, w being the coherence width and \mathcal{V}_0 the visibility as function of the baseline \vec{B} .

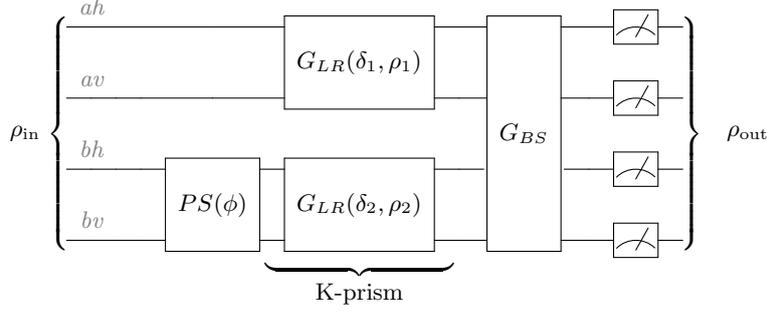


Figure 2: Circuit representation of the beam combiner model.

with a particular polarization in the spatial modes a or b . To this end I define the following projection operators:

$$A = |a\rangle\langle a| \otimes \frac{1}{2} (\mathbb{1} + \sigma_{\theta_A, \varphi_A}), \quad (22a)$$

$$B = |a\rangle\langle a| \otimes \frac{1}{2} (\mathbb{1} + \sigma_{\theta_B, \varphi_B}), \quad (22b)$$

$$C = |b\rangle\langle b| \otimes \frac{1}{2} (\mathbb{1} + \sigma_{\theta_C, \varphi_C}), \quad (22c)$$

$$D = |b\rangle\langle b| \otimes \frac{1}{2} (\mathbb{1} + \sigma_{\theta_D, \varphi_D}). \quad (22d)$$

The analytic expressions for the detection probabilities $\langle A \rangle_{\rho_{\text{out}}}$, $\langle B \rangle_{\rho_{\text{out}}}$, $\langle C \rangle_{\rho_{\text{out}}}$, $\langle D \rangle_{\rho_{\text{out}}}$ for the most general case are very long and not easy to grasp. Therefore they are not explicitly shown in this document. Instead, some special cases that allow to get insight in some interesting aspects are studied in the following subsections.

3.2 The PRIMA FSU

The PRIMA dual-field facility¹ of ESO's Very Large Telescope Interferometer (VLTI) uses two twin beam combiners called Fringe Sensor Units (FSUs). Their optical layout and working principle is explained in detail in Ref. 14. Here, will just shortly sketch some of their key elements which are relevant for the modeling within the context of this work.

3.2.1 The Ideal Case

The light of the two arms of the interferometer is combined at a non-polarizing 50:50 beam splitter.[‡] This means in the ideal case the parameters for the G_{BS} are

$$t_{ah} = t_{av} = t_{bh} = t_{bv} = \sqrt{1/2}; \quad (23a)$$

$$r_{ah} = \sqrt{1 - t_{ah}^2}; \quad (23b)$$

$$r_{av} = \sqrt{1 - t_{av}^2}; \quad (23c)$$

$$r_{bh} = \sqrt{1 - t_{bh}^2}; \quad (23d)$$

$$r_{bv} = \sqrt{1 - t_{bv}^2}; \quad (23e)$$

$$\varphi_{ah} = \varphi_{av} = \varphi_{bh} = \varphi_{bv} = \frac{\pi}{2}. \quad (23f)$$

The targets observed for narrow angle astrometry are point-like and unpolarized which means

$$\mathcal{V} = 1; \quad (24a)$$

$$\mathcal{P} = 0. \quad (24b)$$

[‡]The actual splitting ratio is not perfectly 50:50, but I will start from the ideal values.

In order to retrieve a phase measurement from the interference signal the fringes are sampled at four points with a known separation in phase space. The phase modulation is spatial and uses a kind of "polarization multiplexing" in order to obtain the quadratures. To this end, a K-prism that applies a $\frac{\pi}{2}$ phase shift between horizontal and vertical polarization is inserted in spatial mode b prior to the beam combination. This can be modeled by setting the parameters of the general linear retarder

$$\delta_2 = \frac{\pi}{2}; \quad (25a)$$

$$\rho_2 = 0. \quad (25b)$$

$$(25c)$$

For the moment I do not assume a differential polarization change between arm a and b due to some effects within the VLTI optical train. Therefore $\delta_1 = 0$ and $\rho_1 = 0$.

Finally the phase quadratures are obtained by analyzing the polarization in the h/v -basis after beam combination in each spatial mode:

$$A = |a\rangle\langle a| \otimes \frac{1}{2}(\mathbb{1} + \sigma_z) = |ah\rangle\langle ah|, \quad (26a)$$

$$B = |a\rangle\langle a| \otimes \frac{1}{2}(\mathbb{1} - \sigma_z) = |av\rangle\langle av|, \quad (26b)$$

$$C = |b\rangle\langle b| \otimes \frac{1}{2}(\mathbb{1} + \sigma_z) = |bh\rangle\langle bh|, \quad (26c)$$

$$D = |b\rangle\langle b| \otimes \frac{1}{2}(\mathbb{1} - \sigma_z) = |bv\rangle\langle bv|, \quad (26d)$$

means $\theta_A = \theta_C = 0$, $\theta_B = \theta_D = \pi$, $\varphi_A = \varphi_B = \varphi_C = \varphi_D = 0$.

It is straight forward to verify that indeed for these ideal parameters the detection probabilities are as expected and given as

$$\langle A \rangle_{\rho_{\text{out}}} = \frac{1}{4}(1 + \sin(\phi)), \quad (27a)$$

$$\langle B \rangle_{\rho_{\text{out}}} = \frac{1}{4}(1 + \cos(\phi)), \quad (27b)$$

$$\langle C \rangle_{\rho_{\text{out}}} = \frac{1}{4}(1 - \sin(\phi)), \quad (27c)$$

$$\langle D \rangle_{\rho_{\text{out}}} = \frac{1}{4}(1 - \cos(\phi)). \quad (27d)$$

This result in itself is not very exciting but provides a good consistency check of the model. Starting from there one can assume imperfect values for the different parameters and study the consequences.

3.2.2 Imperfections

K-prism and non-polarizing beam splitter: It seems particularly interesting to assume imperfections of the K-prism and/or the non-polarizing beam splitter. To this end, I recalculate the detection probabilities with

the corresponding parameters undetermined and variable. This leads to

$$\begin{aligned}
\langle A \rangle_{\varrho_{\text{out}}} &= \frac{1}{4} \left(r_{ah}^2 + t_{ah}^2 \right. \\
&+ r_{ah} t_{ah} \left(2 \sin \left(\frac{\delta_2}{2} \right) \cos(2\rho_2) \sin \left(\frac{\delta_2}{2} - \varphi_{bh} + \phi \right) \right. \\
&\left. \left. + \cos(\delta_2 - \varphi_{bh} + \phi) + \cos(\phi - \varphi_{bh}) \right) \right) \quad (28a)
\end{aligned}$$

$$\begin{aligned}
\langle B \rangle_{\varrho_{\text{out}}} &= \frac{1}{4} \left(r_{bv}^2 + t_{av}^2 \right. \\
&+ r_{bv} t_{av} \left(-2 \sin \left(\frac{\delta_2}{2} \right) \cos(2\rho_2) \sin \left(\frac{\delta_2}{2} - \varphi_{bv} + \phi \right) \right. \\
&\left. \left. + \cos(\delta_2 - \varphi_{bv} + \phi) + \cos(\phi - \varphi_{bv}) \right) \right) \quad (28b)
\end{aligned}$$

$$\begin{aligned}
\langle C \rangle_{\varrho_{\text{out}}} &= \frac{1}{4} \left(r_{ah}^2 + t_{bh}^2 \right. \\
&+ 2 r_{ah} t_{bh} \left(\sin^2(\rho_2) \cos(\delta_2 + \varphi_{ah} + \phi) \right. \\
&\left. \left. + \cos^2(\rho_2) \cos(\varphi_{ah} + \phi) \right) \right) \quad (28c)
\end{aligned}$$

$$\begin{aligned}
\langle D \rangle_{\varrho_{\text{out}}} &= \frac{1}{4} \left(r_{av}^2 + t_{bv}^2 \right. \\
&+ 2 r_{av} t_{bv} \left(\cos(\varphi_{av} + \phi) (\cos(\delta_2) \cos^2(\rho_2) + \sin^2(\rho_2)) \right. \\
&\left. \left. - \sin(\delta_2) \cos^2(\rho_2) \sin(\varphi_{av} + \phi) \right) \right) \quad (28d)
\end{aligned}$$

This more general form of the expressions allows, for instance, to study the influence of different deviations, such as a misalignment of the K-prism coordinate system with respect to the one of the polarization analysis, or a non-perfect retardation of the K-prism, see Fig. 3. As can be seen, an imperfection in the alignment, $\rho_2 \neq 0$, as well as in the retardation, $\delta_2 \neq \frac{\pi}{2}$, introduces a deviation from the ideal phase difference between detector quadrants A-B and C-D. While the dependence of δ_2 is perfectly linear with slope 1, see Fig. 3(b), the dependence of ρ_2 is quadratically periodic as $\arctan(\cot(\rho_2^2)) - \arctan(\tan(\rho_2^2))$, see Fig. 3(a).

Another interesting aspect to look at is an imperfection of the beam combiner. To this end I consider the overall detection probability $\langle A \rangle_{\varrho_{\text{out}}} + \langle B \rangle_{\varrho_{\text{out}}} + \langle C \rangle_{\varrho_{\text{out}}} + \langle D \rangle_{\varrho_{\text{out}}}$. It is straightforward to show that for a perfect lossless beam splitter the only physical solution for the phase acquired upon reflection is $\varphi_l = \frac{\pi}{2}$; otherwise the overall detection probability could become greater than one.

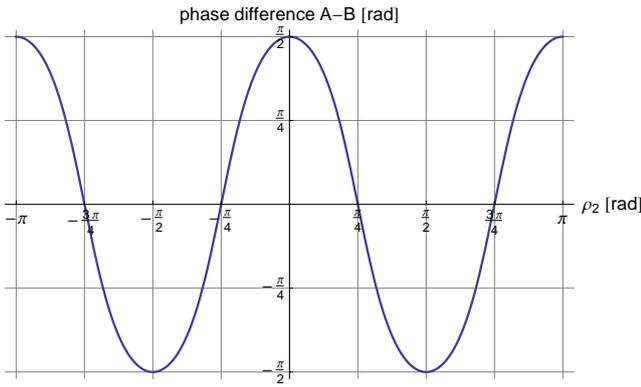
The reflection and transmission amplitudes for the PRIMA FSU have been measured to be

$$t_{av} = t_{bv} = r_{ah} = r_{bh} = \sqrt{0.49}; \quad (29a)$$

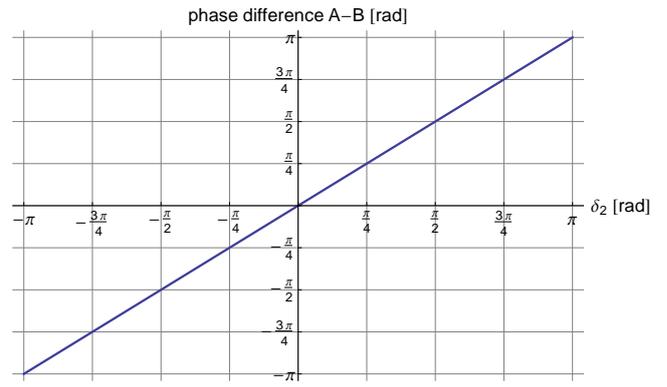
$$t_{ah} = t_{bh} = \sqrt{0.42}; \quad (29b)$$

$$r_{av} = r_{bv} = \sqrt{0.44}. \quad (29c)$$

This allows for the beam splitter phases φ_l to deviate from their ideal value by $\Delta\varphi$. Fig. 4(a) displays the overall detection probability versus OPD, ϕ , for an arbitrarily assumed value of $\Delta\varphi = 2.5^\circ$. As can be seen, the overall detection probability varies with OPD; something that is indeed observed for real data. This must, of course, *not* be seen as a violation of energy conservation as long as the detection probability is still below one for all values

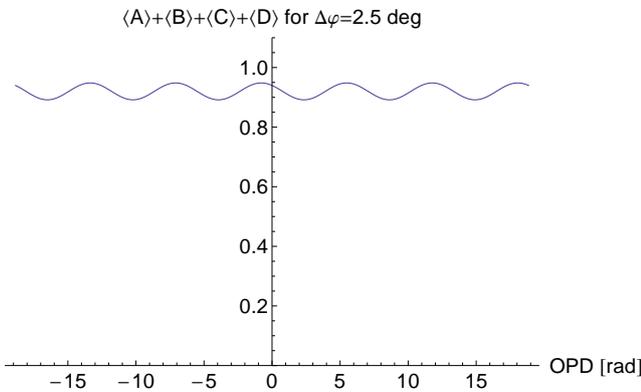


(a) Phase shift between quadratures A and B versus K-prism angle ρ_2 in degrees. The dependence is quadratically periodic as $\arctan(\cot(\rho_2^2)) - \arctan(\tan(\rho_2^2))$.

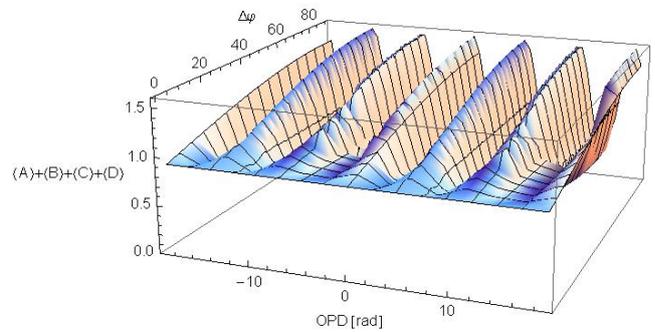


(b) Phase shift between quadratures A and B versus K-prism retardation δ_2 in degrees. The dependence is perfectly linear with a slope of 1.

Figure 3: Influence of non-perfect K-prism. These figures display the phase difference between quadratures A and B versus retardation and rotation angle of the K-prism.



(a) Overall detection probability versus OPD for beam splitter phase deviation $\Delta\varphi = 2.5^\circ$.



(b) Overall detection probability versus OPD and beam splitter phase deviation $\Delta\varphi$.

Figure 4: Effects of a non-perfect beam combiner on the overall detection probability.

of OPD[§]. However, to ensure an overall detection probability less or equal to one (and energy to be conserved) the value of phase deviation $\Delta\varphi$ is constrained to a certain range for a given set of transmission and reflection amplitudes[¶] as in Eqns. 29. This can be directly seen in Fig. 4(b).

What cannot be seen from Fig. 4, but easily derived from Eqns. 28 is that the non-ideal phase-shifts of a non-ideal beam splitter change also the phase differences between the quadratures A and C (or B and D) from their ideal values of π .

Differential birefringence within the VLTI optical train: The model can also be used to show that any retardation, $\delta_1 \neq 0$, between horizontal and vertical polarization components which is differential between spatial modes a and b (apart from the K-prism), and which might be introduced by the VLTI optical train, shifts the phases between quadratures A and B and A and D linearly. If this retardation occurs not along h and v but along any other direction, $\rho_1 \neq 0$, it reduces in addition the visibility of the interference.

Application to photometry: The PRIMA FSUs have no photometric channels. Therefore they require careful calibration of the photometry in order to correctly estimate interferometric phases and visibilities.¹⁵ Due to the aforementioned polarization multiplexing the photometric measurement is particularly susceptible to partially polarized input states. This can be also verified by the model. The photometry is simply obtained by setting the transmission and reflection coefficients for spatial modes a or b to zero. This corresponds to placing a shutter in one of the input beams.

So, in the following I use

$$PHOT_a : \quad t_{bh} = t_{bv} = r_{bh} = r_{bv} = 0$$

and

$$PHOT_b : \quad t_{ah} = t_{av} = r_{ah} = r_{av} = 0.$$

This yields for $PHOT_a$:

$$\langle A \rangle_{\ell_{\text{out}}} = \langle C \rangle_{\ell_{\text{out}}} = \frac{1}{8} (1 + \mathcal{P} \cos(\theta_{\text{pol}})) \quad (30a)$$

$$\langle B \rangle_{\ell_{\text{out}}} = \langle D \rangle_{\ell_{\text{out}}} = \frac{1}{8} (1 - \mathcal{P} \cos(\theta_{\text{pol}})) \quad (30b)$$

and for $PHOT_b$ (this is the input channel that contains the K-prism):

$$\begin{aligned} \langle A \rangle_{\ell_{\text{out}}} = \langle C \rangle_{\ell_{\text{out}}} = & \frac{1}{8} \left(\mathcal{P} \sin(\theta_{\text{pol}}) \left(\sin^2 \left(\frac{\delta_2}{2} \right) \cos(\varphi_{\text{pol}}) \sin(4\rho_2) - 2 \sin(\delta_2) \sin(\varphi_{\text{pol}}) \sin(\rho_2) \cos(\rho_2) \right) \right. \\ & \left. + \mathcal{P} \cos(\theta_{\text{pol}}) \left(\sin^2 \left(\frac{\delta_2}{2} \right) \cos(4\rho_2) + \cos^2 \left(\frac{\delta_2}{2} \right) \right) + 1 \right) \end{aligned} \quad (31a)$$

$$\begin{aligned} \langle B \rangle_{\ell_{\text{out}}} = \langle D \rangle_{\ell_{\text{out}}} = & \frac{1}{8} \left(\mathcal{P} \sin(\theta_{\text{pol}}) \left(\sin(\delta_2) \sin(\varphi_{\text{pol}}) \sin(2\rho_2) - \sin^2 \left(\frac{\delta_2}{2} \right) \cos(\varphi_{\text{pol}}) \sin(4\rho_2) \right) \right. \\ & \left. - \mathcal{P} \cos(\theta_{\text{pol}}) \left(\sin^2 \left(\frac{\delta_2}{2} \right) \cos(4\rho_2) + \cos^2 \left(\frac{\delta_2}{2} \right) \right) + 1 \right) \end{aligned} \quad (31b)$$

As can be seen, for an unpolarized input state, $\mathcal{P} = 0$, the photometry is constant and identical at all outputs A, B, C, D . However, for partially polarized input state, $\mathcal{P} \neq 0$, the situation is different:

[§]Note that in the data only those photons are taken into account that are registered in the detector which are not *all* the photons collected by the telescopes.

[¶]The analytic form for the allowed range of $\Delta\varphi$ can be derived for a given set of transmission and reflection amplitudes. However, this is out of the scope of this work.

For an ideal K-prism, $\delta_2 = \frac{\pi}{2}$, $\rho_2 = 0$, both photometric measurements, $PHOT_a$ and $PHOT_b$, are the same, but vary depending on the polarization state, θ_{pol} and φ_{pol} . For partially polarized input state and a non-ideal K-prism both photometric measurements, $PHOT_a$ and $PHOT_b$, are different and vary depending on the polarization state, θ_{pol} and φ_{pol} . Similar considerations apply to differential polarization changes in the VLTI optical train, $\delta_1 \neq 0$, $\rho_1 \neq 0$.

3.3 Polarization Based FSU-like Scheme

Conventional beam combination concepts might suffer from false fringe signals caused by incoherent light, if the photometry is not properly taken into account and the combination is not perfectly balanced (e.g., no perfect 50:50 beam splitter). This is because incoherent light might leak unequally in both output channels of the beam combiner and create spurious fringe signals.

Within the scope of a feasibility study for a second generation fringe tracker for the VLTI, a topology for beam combination has been proposed that deals with this difficulty in an inventive and simple way.¹¹

The scheme suggested in Ref. 11 relies on the combination of two orthogonal polarizations in one spatial mode. In the following I will show that the model depicted in Fig. 2 can be used to simulate such a method of beam combination. I will further show that the polarization based scheme is mathematically equivalent to the PRIMA FSU case treated in the previous section.

As two orthogonal modes of polarization shall be combined in the same spatial mode, the core element of the scheme is a polarizing beam splitter. This implies for the parameters of G_{BS}

$$t_{ah} = t_{bh} = r_{av} = r_{bv} = 1; \quad (32a)$$

$$t_{av} = t_{bv} = r_{ah} = r_{bh} = 0; \quad (32b)$$

$$\varphi_{ah} = \varphi_{av} = \varphi_{bh} = \varphi_{bv} = \frac{\pi}{2}. \quad (32c)$$

The general linear retarder in spatial mode a is used to simulate a half-wave plate at an angle of 45° that flips the polarization in this spatial mode, and that guarantees that two orthogonal polarizations of different *input* spatial modes are indeed combined into the same *output* spatial mode behind the beam splitter

$$\delta_1 = \pi; \quad (33a)$$

$$\rho_1 = \frac{\pi}{4}. \quad (33b)$$

In order to maintain four signals in phase quadrature, the general linear retarder in spatial mode b is assumed to be the same K-prism like in the FSU

$$\delta_2 = \frac{\pi}{2}; \quad (34a)$$

$$\rho_2 = 0. \quad (34b)$$

At this stage "which-path" information is contained in the polarization which prevents the observation of interference. This information can, however, be easily erased by analyzing the polarization in each spatial mode in a basis that is conjugate to the one defined by G_{BS} . One possible choice for the observables that fulfills this criterion is

$$A = |a\rangle\langle a| \otimes \frac{1}{2}(\mathbb{1} + \sigma_x), \quad (35a)$$

$$B = |a\rangle\langle a| \otimes \frac{1}{2}(\mathbb{1} - \sigma_x), \quad (35b)$$

$$C = |b\rangle\langle b| \otimes \frac{1}{2}(\mathbb{1} + \sigma_x), \quad (35c)$$

$$D = |b\rangle\langle b| \otimes \frac{1}{2}(\mathbb{1} - \sigma_x), \quad (35d)$$

means $\theta_A = \theta_B = \theta_C = \theta_D = \frac{\pi}{2}$, $\varphi_A = \varphi_C = 0$, $\varphi_B = \varphi_D = \pi$.

Actually, it is this choice of polarization basis where the strength of this topology lies: While splitting ratios of non-polarizing beam splitters are usually fixed by their dielectric coatings, the direction of polarization analysis can be easily adjusted to be conjugate, or – in other words – to be 50:50 balanced.

With these parameter settings the detection probabilities are

$$\langle A \rangle_{\rho_{\text{out}}} = \frac{1}{4} (1 + \cos(\phi)), \quad (36a)$$

$$\langle B \rangle_{\rho_{\text{out}}} = \frac{1}{4} (1 - \cos(\phi)), \quad (36b)$$

$$\langle C \rangle_{\rho_{\text{out}}} = \frac{1}{4} (1 + \sin(\phi)), \quad (36c)$$

$$\langle D \rangle_{\rho_{\text{out}}} = \frac{1}{4} (1 - \sin(\phi)), \quad (36d)$$

which is equivalent to Eqns. 27 and proves that the two schemes are mathematically the same.

From a physical point of view it is interesting to note that:

- In the original FSU scheme the interference happens in the spatial mode and is independent of the polarization mode, while
- in the modified polarization based FSU scheme the interference happens in the polarization mode and is independent of the spatial mode.

This can also be seen from the fact that the conjugate signals in phase opposition (with π phase-shift in between) are

- two different spatial modes A and C (B and D) with the same polarization, Eqns. 27, in the original FSU scheme, and
- two different polarization modes A and B (C and D) with the same spatial mode, Eqns. 36, in the modified FSU scheme.

Analogous to Sec. 3.2 one could use the model to study all possible disturbances and imperfections and make a comparison of the two schemes, which would be a diligent but routine piece of work that I will not detail any further here. Instead, I would like to comment on a part of the original work of Ref. 11.

In their publication the authors state that:

"[...] this topology [...] only combines one polarization of one beam with one polarization of the other beam. Therefore it is not able to duplicate the functionality of the two input two output beam combiner [...] in which both input beams include both polarizations and both output polarizations in each output beam fall upon the same detector."

I think this statement is too restrictive as it does not account for the full potential of the polarization-based scheme. Actually, the splitting of the beams according to polarization *prior* to their re-combination is not necessary. As can be seen from my analysis above, indeed a two input/two output scheme does exist.

Naturally, as the polarization of photons constitutes a *two-level* quantum system it cannot be used for any symmetric splitting $1 \rightarrow n$ with $n > 2$ ($n \in \mathbb{N}$). Still, I conjecture that the polarization-based beam re-combination can be employed in *any pairwise* combination scheme of multi-telescope interferometers. The only provision is that a splitting $1 \rightarrow n$ ($n > 2$) has to take place in the spatial mode, based on intensity rather than polarization, see Fig. 5 for an example.

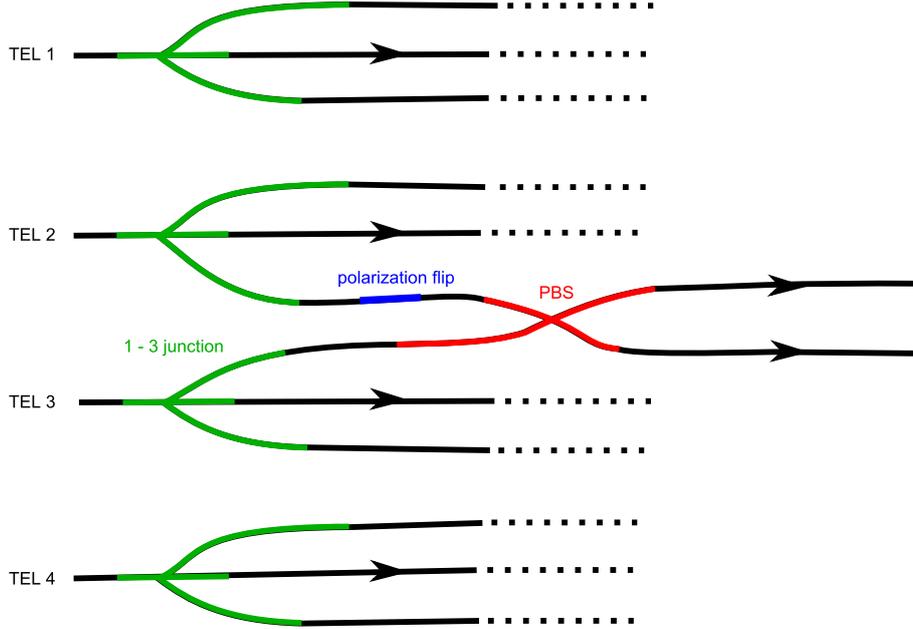


Figure 5: Example for an extended polarization-based beam combination scheme: The light of four telescopes is combined pairwise on all six baselines. To this end, the light of each telescope is first split into three spatial modes at a $1 \rightarrow 3$ junction according to intensity. Afterwards the recombination happens pairwise at a polarizing beam splitter (PBS). Prior to the re-combination the polarization is flipped in one of the arms.

In particular, this implies that – provided a realization of a polarizing beam splitter exist in integrated optics for a wavelength band relevant to stellar interferometry^{||} – any of the integrated optics pairwise combination schemes used so far for other interferometric projects (see e.g., Refs. 18 or 19) could be realized in a polarization-based beam combination variation and profit from the particular advantages.

4. CONCLUSION

In this contribution I have modeled beam combiners of stellar interferometers as linear optics networks. To this end I introduced a mathematical formalism that is commonly used in the field of QIS. The formalism, though being similar to methods of classical physics as far as it concerns polarization, offers a different view on interference and allows an integrated treatment of both polarization and spatial interference.

I used the formalism to analyze two allegedly different topologies for two-telescope beam combination and explained why they are indeed equivalent from a mathematical point of view. On the other hand, as one of the schemes has particular advantages from a physical point of view, I showed how it might be extended to multi-telescope beam combination.

I am convinced that the presented formalism can be extremely helpful for the development and the design of future interferometric instruments. Although I could only present a small set of techniques within this contribution, QIS offers a big tool box with a whole lot more of methods that might be useful in the context of stellar interferometry, such as e.g. quantum state²⁰ and process tomography²¹ just to name two. In this sense, my aim was to start building a bridge between two fields of physics that are per se completely different but still share similar problems. I hope others will follow this path.

^{||}It seems that different concepts based on various physical effects do exist for polarization beam splitters in optical fibers¹⁶ as well as integrated optics chips¹⁷ at telecom wavelengths.

ACKNOWLEDGMENTS

C.S. would like to thank Bernard Lazareff, Frank Eisenhauer and Jean-Philippe Berger for fruitful discussions, and Jean-Philippe Berger for proofreading of the manuscript and providing lots of useful comments. The circuit diagrams were typeset by using the L^AT_EX package Q-circuit:
<http://physics.unm.edu/CQuIC/Qcircuit/>.

REFERENCES

1. G. van Belle, J. Sahlmann, R. Abuter, M. Accardo, L. Andolfato, S. Brillant, J. de Jong, F. Derie, F. Delplancke, T. Duc, C. Dupuy, B. Gilli, P. Gitton, P. Haguenaue, L. Jocou, A. Jost, N. Di Lieto, R. Frahm, S. Mnardi, S. Morel, J.-M. Moresmau, R. Palsa, D. Popovic, E. Pozna, F. Puech, S. Lvque, A. Ramirez, N. Schuhler, F. Somboli, S. Wehner, and T. E. Consortium, “The vlti prima facility,” *The Messenger*, pp. 6–11, Dec. 2008.
2. J. Woillez and et al., “Improving the astrometric performance of vlti-prima,” 2014.
3. J.-B. LeBouquin, *Imagerie par synthese d’ouverture optique, application aux etoiles chimiquement particulieres*. PhD thesis, Laboratoire d’AstrOphysique de Grenoble, UJF/CNRS, BP 53, F-38041 Grenoble Cedex 9, Nov. 2005.
4. M. Elias II, Nicholas, “Optical interferometric polarimetry. ii. theory,” *The Astrophysical Journal* **611**(2), p. 1175, 2004.
5. P. Lawson, *Principles of long baseline interferometry – Course notes from the 1999 Michelson Summer School, August 15-19, 1999*, no. 00-009 07/00 in JPL Publication, Jet Propulsion Lab (JPL), 2000.
6. P. Kok, W. J. Munro, K. Nemoto, T. C. Ralph, J. P. Dowling, and G. J. Milburn, “Linear optical quantum computing with photonic qubits,” *Rev. Mod. Phys.* **79**, pp. 135–174, Jan 2007.
7. J. L. O’Brien, “Optical quantum computing,” *Science* **318**(5856), pp. 1567–1570, 2007.
8. J.-W. Pan, Z.-B. Chen, C.-Y. Lu, H. Weinfurter, A. Zeilinger, and M. Żukowski, “Multiphoton entanglement and interferometry,” *Rev. Mod. Phys.* **84**, pp. 777–838, May 2012.
9. S. B. Shaklan, “Fiber optic beam combiner for multiple-telescope interferometry,” *Optical Engineering* **29**(6), pp. 684–689, 1990.
10. H. kai Hsiao, K. A. Winick, J. D. Monnier, and J.-P. Berger, “An infrared integrated optic astronomical beam combiner for stellar interferometry at 3-4 μm ,” *Opt. Express* **17**, pp. 18489–18500, Oct 2009.
11. J. A. Meisner, W. J. Jaffe, R. S. Le Poole, S. F. Pereira, A. Quirrenbach, D. Raban, and A. Vosteen, “The polarization-based collimated beam combiner and the proposed nova fringe tracker (nft) for the vlti,” 2010.
12. Lazareff, B., Le Bouquin, J.-B., and Berger, J.-P., “A novel technique to control differential birefringence in optical interferometers,” *A&A* **543**, p. A31, 2012.
13. C. Schmid, “Beam combination in stellar interferometry: A quantum point of view,” *in preparation*.
14. J. Sahlmann, S. Ménardi, R. Abuter, M. Accardo, S. Mottini, and F. Delplancke, “The prima fringe sensor unit,” *A&A* **507**, pp. 1739–1757, Dec. 2009.
15. N. Gomes, C. Schmid, J. Sahlmann, S. Menardi, R. Abuter, A. Merand, and F. Delplancke, “Estimating visibility amplitudes with the prima fringe trackers,” 2012.
16. I. Yokohama, K. Okamoto, and J. Noda, “Analysis of fiber-optic polarizing beam splitters consisting of fused-taper couplers,” *Journal of Lightwave Technology* **LT-4**(9), p. 1352, 1986.
17. Q. Wang, G. Farrell, and Y. Semenova, “Design of integrated polarization beam splitter with liquid crystal,” *IEEE Journal of Selected Topics in Quantum Electronics* **12**(6), p. 1349, 2006.
18. E. Choquet, F. Cassaing, G. Perrin, F. Eisenhauer, W. Brandner, C. Straubmeier, K. Perraut, and S. Gillessen, “Gravity: design and performance of the fringe tracker,” 2010.
19. Le Bouquin, J.-B., Berger, J.-P., Lazareff, B., Zins, G., Haguenaue, P., Jocou, L., Kern, P., Millan-Gabet, R., Traub, W., Absil, O., Augereau, J.-C., Benisty, M., Blind, N., Bonfils, X., Bourget, P., Delboulbe, A., Feautrier, P., Germain, M., Gitton, P., Gillier, D., Kiekebusch, M., Kluska, J., Knudstrup, J., Labeye, P., Lizon, J.-L., Monin, J.-L., Magnard, Y., Malbet, F., Maurel, D., Menard, F., Micallef, M., Michaud, L., Montagnier, G., Morel, S., Moulin, T., Perraut, K., Popovic, D., Rabou, P., Rochat, S., Rojas, C., Roussel, F., Roux, A., Stadler, E., Stefl, S., Tatulli, E., and Ventura, N., “Pionier: a 4-telescope visitor instrument at vlti,” *A&A* **535**, p. A67, 2011.

20. D. F. V. James, P. G. Kwiat, W. J. Munro, and A. G. White, "Measurement of qubits," *Phys. Rev. A* **64**, p. 052312, Oct 2001.
21. I. L. Chuang and M. A. Nielsen, "Prescription for experimental determination of the dynamics of a quantum black box," *Journal of Modern Optics* **44**(11-12), pp. 2455–2467, 1997.