Reaching micro-arcsecond astrometry with long baseline optical interferometry

Application to the GRAVITY instrument

S. Lacour\textsuperscript{1}, F. Eisenhauer\textsuperscript{2}, S. Gillessen\textsuperscript{2}, O. Pfun\textsuperscript{2}, J. Woillez\textsuperscript{3}, H. Bonnet\textsuperscript{3}, G. Perrin\textsuperscript{1}, B. Lazareff\textsuperscript{4}, S. Rabien\textsuperscript{2}, V. Lapeyrère\textsuperscript{1}, Y. Clénet\textsuperscript{1}, P. Kervella\textsuperscript{1}, and Y. Kok\textsuperscript{2}

\textsuperscript{1} LESIA/Observatoire de Paris, CNRS, UPMC, Université Paris Diderot, 5 place Jules Janssen, 92195 Meudon, France
\textsuperscript{2} Max-Planck Institut für extraterrestrische Physik (MPE), Giessenbachstrasse 1, 85748 Garching, Germany
\textsuperscript{3} European Southern Observatory (ESO), Karl-Schwarzschild-Str. 2, 85748 Garching, Germany
\textsuperscript{4} Institut de Planétologie et d’Astrophysique de Grenoble, 38400 St Martin d’Hères, France

Received 3 April 2014 / Accepted 23 May 2014

\textbf{ABSTRACT}

\textbf{Context.} A basic principle of long baseline interferometry is that an optical path difference (OPD) directly translates into an astrometric measurement. In the simplest case, the OPD is equal to the scalar product between the vector that links the two telescopes and the normalized vector pointing toward the star. However, in some circumstances, too simple an interpretation of this scalar product leads to seemingly conflicting results, called here “the baseline paradox”.

\textbf{Aims.} For micro-arcsecond accuracy astrometry, we have to model the metrology measurement in full. It involves a complex system subject to many optical effects: from pure baseline errors to static, quasi-static, and high-order optical aberrations. The goal of this paper is to present the strategy used by the “General Relativity Analysis via VLT InTerferometrY” instrument (GRAVITY) to minimize the biases introduced by these defects.

\textbf{Methods.} It is possible to give an analytical formula for how the baselines and tip-tilt errors affect the astrometric measurement. This formula depends on the limit points of three type of baselines: the wide-angle baseline, the narrow-angle baseline, and the imaging baseline. We also numerically include non-common path, higher order aberrations, whose amplitudes were measured during technical time at the Very Large Telescope Interferometer (VLTI). We end by simulating the influence of high-order, common-path aberrations due to atmospheric residuals calculated from a Monte-Carlo simulation tool for adaptive optics (AO) systems.

\textbf{Results.} The result of this work is an error budget of the biases caused by the multiple optical imperfections, including optical dispersion. We show that the beam stabilization through both focal and pupil tracking is crucial to the GRAVITY system. Assuming the instrument pupil is stabilized at a 4 cm level on M1 and a field tracking below 0.2 μ/D, we show that GRAVITY will be able to reach its objective of 10 μas accuracy.

\textbf{Key words.} astrometry – instrumentation: interferometers – instrumentation: high angular resolution

\section{1. Introduction}

A particular interest of astronomical long-baseline optical interferometers is their ability to perform high accuracy angular astrometry (Shao & Colavita 1992). The basic idea is that astrometry can leverage on the high-resolution of the interferometric lab up to the telescopes. The idea behind this is that an homodyne interferometer converts an angular position on the sky into an optical path difference (OPD). In the first-order approximation, this optical path difference is equal to the difference of the star in the sky (s) projected onto the baseline vector of the interferometer (B):

\[ \text{OPD} = s \cdot B. \]  
(1)

Within this simple equation is hidden a lot of complexity. Above all, how is the baseline vector defined? The intuitive answer is to define the baseline by the vector that links the geographical position of the two telescopes. But unfortunately, it only makes sense in the case of two perfectly identical telescopes. This immediately implies some questions: What happens when the two telescopes are not identical? What happens if, for one reason or another, the optical alignment of the two telescopes is not identical?

When dealing with precision astrometry, this kind of problem is pivotal. The main scientific objective of the GRAVITY instrument is to probe the event horizon of the Galactic center black hole (Eisenhauer et al. 2011). It will have to reach unprecedented astrometric accuracy between two one-arcsec-separated objects on the order of 10 μas (ten parts per million). According to Eq. (1), it means that the baseline of the Very Large Telescope Interferometer (VLTI) must be know to a submillimeter level.

To complicate the fact that the baseline must be known to an extreme precision, the literature defines several baselines (Colavita et al. 1999; Hrynevych et al. 2004; Sahlmann et al. 2013). From the simple definition of the baseline by Lawson (2000) (the distance between two apertures), emerged the idea of a wide angle baseline (WAB), a narrow angle baseline (NAB,
Colavita 2009) and an IMaging Baseline (IMB, Woillez & Lacour 2013). The idea behind these definitions was that each baseline corresponds to a given use of the interferometer.

But all these baselines affect the astrometric measurement to various degrees. Woillez & Lacour (2013) used the so-called “limit points” to define the baselines vectors. Instead of defining the vector by its direction and length, the baseline vector is defined by the point from where it emerges to the point where it goes: the vector is also defined by its geographical position. In this paper, we establish how the position of the various limit points affect the OPD measurement. Errors due to optical aberrations are intrinsically linked to the position of the limit points, and we will see that they are an important part of astrometric limitations. Throughout this paper, we build an error budget, focusing on the biases that occur due to baseline errors and optical aberrations and showing how GRAVITY, through pupil control, plans to keep these bias below 10 μas. This paper does not intend to cover random process errors like photon noise, detector noise, or background noise.

This paper starts with a presentation of “The baseline paradox of optical interferometry” in Sect. 2. It shows in a pedagogical way how too simple an assumption on the concept of baseline leads to conflicting observations. In Sect. 3, we use a three-dimensional model of the interferometer to calculate the impact of the different baselines’ limit points on the OPD measurement. However, this analytical modeling is not enough to account for aberrations of any order higher than the tip-tilt, so in Sect. 4 we simplify the problem to two dimensions to investigate how these aberrations modify the OPD. The rest of the sections apply those results to the GRAVITY instrument. Section 5 describes how the instrument works and the technical choices made for the astrometric measurement. Section 6 puts concrete numbers on the errors and the way they affect astrometry. The end-product of this section is the error budget presented in Table 1. Section 7 concludes.

2. Necessity for multiple baselines

2.1. An apparent paradox

Consider a two-telescope interferometer as depicted in Fig. 1. In front of each telescope there is an opaque mask with a circular aperture. The aperture is not centered with respect to the optical axis of the telescope. The two masks are rotating, moving the respective holes in different directions.

In the first case (left panel), the interferometer looks at a single star. This star is exactly in the direction of observation of each telescope. For an unresolved object, the incoming wavefront arrives perpendicularly to the pupil of the telescope. As such, the masks cut out part of the incoming wavefront. But what matters is the cross-product between the aperture and the phase of the wavefront over the pupil: the wavefront being flat over the pupil, the result of the cross-product does not depend on the position of the hole. In fact, the phase averaged over the aperture is the same as it would have been if obtained without any mask. As a result, the position of the white-light fringe in the beam combiner is not affected by the rotation of the two masks, even though the distance between the two apertures changes. The optical path difference between the white-light fringe and the zero of the interferometer depends on the distance between the two telescopes ($B_2$):

$$\text{OPD} = s \cdot B_2.$$  \hspace{1cm} (2)

We assume here that no atmospheric perturbations are present.

![Fig. 1. The baseline paradox. Consider a thought experiment where an interferometer is equipped with two masks that rotate in the pupil of both telescopes. By doing so, the distance between the two apertures changes with time while the telescopes stay fixed. In the left panel, the observation is done on a single target, acquired on the axis of the telescopes (perpendicular to the plane of the aperture). The OPD measurement is done between the white-light fringe and the zero reference of the interferometer. The OPD does not change with the rotation of the masks. In the right panel, the interferometer is recording two fringe packets obtained by observing two stars in the same field. The difference in OPD is measured between the two white-light fringes. \(\delta\text{OPD}\) does change with the rotation of the masks.]

The zero of the interferometer is the point at the equal optical path distance of the pupils of the two telescopes.

In the second case (right panel of Fig. 1), we are not interested in the exact OPD of the fringes, but in the differential OPD between two fringe patterns. The interferometer is simultaneously observing two celestial sources of the same field. Equation (1) becomes a differential measurement:

$$\delta\text{OPD} = (s - p) \cdot B_2,$$  \hspace{1cm} (3)

where \((s - p)\) is the difference in astrometric position between the two sources, and \(\delta\text{OPD}\) the difference in position between the two white-light fringes. In this case, the delay in optical length between the fringe patterns matches the distance between the two apertures, projected in the direction of the source. This time, the relevant baseline is the one that is modulated \((B_2)\), and the differential position of the fringes will move when the masks rotate.

The paradox is therefore that, on two identical interferometers, two behaviors for the fringes positions can coexist. In one case (single star on-axis), the fringe pattern is fixed in absolute position, while for the other case (two stars), one fringe pattern is moving with respect to the other. In other words, how is it possible that two different baselines \(B_1\) and \(B_2\) can simultaneously exist and apply to the same interferometer?

2.2. Solution

To face this paradox (Woillez & Lacour 2013) proposed several definition of baselines, which would apply to different kinds of interferometers (or different uses of the same interferometer, as shown above). The idea is that different baselines are used depending on the way the OPD is calculated. The wide angle baseline (WAB) is defined by the two limit points that are the pivot points of the telescopes (the center of rotation of the telescopes). The imaging baseline (IMB) is defined by the two limit points that are the barycentric mean of the effective pupils of each aperture. It must be used when observing at the same time multiple
objects inside a same field of view. Last, the narrow angle baseline (NAB) is defined by the limit points that are the image of the metrology end points in primary space. In the example of the previous section, the $B_1$ baseline is the WAB and $B_2$ the IMB.

Some baselines are easier to work with than others. The baseline must be stable to give an accurate astrometric measurement. An ideal baseline would have two properties:

1. it should be fixed with respect to the sidereal reference frame or, at least, fixed within a coordinate frame that can be accurately converted into the sidereal frame (e.g., the terrestrial reference frame);
2. it should be fixed with respect to the interferometric plane where the white-light fringe is observed.

The word “fix” has the meaning that it does not change with respect to time, telescope pointing, or stellar object position. Alternatively, instead of being fixed the baseline “fix” can be replaced by “predictable”, although it will complicate data reduction. Requirement (1) is to be able to project the baseline in the reference frame of the astronomical object. Requirement (2) is to produce an OPD measurement independent of the position of the objects in the interferometric field. By definition the pivot points fulfill Req. (1), while the pupil barycenter fulfills Req. (2) (Woillez & Lacour 2013). The NAB does not necessarily fulfill any of those two properties.

### 2.3. Practical examples

Depending on the type of interferometers, some baseline definitions are more useful than others. As an example, the WAB is zero for any kind of single telescope interferometer (like SAM on NACO; Tuthill et al. 2010), as well as an interferometer like the LBT (Hill et al. 2006): $B_{WAB} = 0$. This is because the whole structure rotates around a single pivot point: each entrance aperture has the same pivot point. For these interferometers, the IMB, which is the distance between two apertures (LBT), holes (SAM), or mirrors (Michelson stellar interferometer), is the adequate baseline for interpreting the data.

Long-baseline interferometers with a single-beam combiner (for example, AMBER, PIONIER; Petrov et al. 2007; Le Bouquin et al. 2011) also require using the imaging baseline. However, the imaging baseline is hard to monitor because the effective pupil depends on the whole optical train, pupil vignetting, telescope pointing, etc. In other words, it does not follow Req. (1). Often and for convenience, data reduction softwares thus assume that the IMB is identical to the WAB. However, users of those instruments need to be aware that it introduces an error into their interferometric measurement.

Finally, in dual feed interferometers like PTI (Colavita et al. 1999), PRIMA (Delplancke 2008), GRAVITY (Eisenhauer et al. 2011), or Astra (Woillez et al. 2010), the metrological link monitors a third baseline often called the NAB. We show in the following that the NAB should closely match the IMB in order to link the OPD metrology measurement with the angular separation. Moreover, dual feed interferometers have two IMB, one for each combiner. This complicates the equations by adding quasi-static error terms in the OPD measurement.

### 3. Baseline limit-points and OPD measurement

#### 3.1. The wide angle baseline

Figure 2 shows the simplest representation possible of a purely theoretical astrometric interferometer. The telescopes are represented by two geographically-localized points $L_1$ and $L_2$. The telescopes point by rotating around themselves, i.e., $L_1$ and $L_2$ are also the pivot points. All mirrors have been suppressed in the representation, and the optical trajectories are represented by virtual rays. Because of this, the beam combiner have two virtual images, one in each arm of the interferometer (in reality, thanks to folding mirrors, the fringes are – of course – physically located at the same position).

The interferometer metrology probes two optical length measurements, one for each telescope. The astrometric measurement uses the difference between the two:

$$\text{OPD} = \text{OPD}_{1-\text{INT}} - \text{OPD}_{1-\text{INT}}.$$  

(4)

The two values $\text{OPD}_{1-\text{INT}}$ and $\text{OPD}_{2-\text{INT}}$ are measured, to a constant, between the central fringe and the telescope pivot points $L_1$ or $L_2$. In all of Sect. 3, we assume that the metrology measurement is obtained with respect to the central fringe, i.e. without any additional OPD introduced at the beam combiner level.

Since by definition the central fringe is the fringe at equal optical path, we can establish, by the following equation, the relation between $\text{OPD}_{\text{INT}}$ and $\text{OPD}_{\text{EXT}}$:

$$\text{OPD}_{1-\text{EXT}} + \text{OPD}_{1-\text{INT}} = \text{OPD}_{2-\text{EXT}} + \text{OPD}_{2-\text{INT}}.$$  

(5)

where the two numbers $\text{OPD}_{1-\text{EXT}}$ and $\text{OPD}_{2-\text{EXT}}$ are the optical path lengths up to the stellar emission, outside the interferometer.

Last, the equation that links the two external optical paths is the projection in the stellar direction of the three-dimensional

\[ \text{OPD}_{1-\text{EXT}} + \text{OPD}_{1-\text{INT}} = \text{OPD}_{2-\text{EXT}} + \text{OPD}_{2-\text{INT}}. \]

\[ \text{OPD}_{1-\text{EXT}} + \text{OPD}_{1-\text{INT}} = \text{OPD}_{2-\text{EXT}} + \text{OPD}_{2-\text{INT}}. \]

\[ \text{OPD}_{1-\text{EXT}} + \text{OPD}_{1-\text{INT}} = \text{OPD}_{2-\text{EXT}} + \text{OPD}_{2-\text{INT}}. \]
vector delimited by the two pivot points. According to the baseline definitions in Woillez & Lacour (2013), this vector is the WAB:

$$\text{OPD}_{1\text{-EXT}} - \text{OPD}_{2\text{-EXT}} = s \cdot \overrightarrow{L_1L_2}. \quad (6)$$

In the whole paper, we use an arrow to represent a vector that goes from a point $X$ to a point $Y$: $\overrightarrow{XY} = Y - X$. Here,

$$\overrightarrow{L_1L_2} = L_2 - L_1 = B_{\text{WAB}}. \quad (7)$$

The advantage of the WAB is that it is theoretically fixed with respect to the ground: it does not change with the pointing direction of the telescope.

Equations (4)–(6) directly lead to the relation that links the OPD measurement to the stellar position:

$$\text{OPD} = s \cdot B_{\text{WAB}}. \quad (8)$$

### 3.2. The narrow angle baseline

The upper panel of Fig. 3 increases the complexity with respect to Fig. 2 by including four measurements instead of two: the interferometer is now observing two different stars: $p$ and $s$. The OP measurements are made between the white-light fringes and the end points of the metrology signal, defined in geographic space: $M_1$ and $M_2$. The end points can be inside the optical train. If this is the case, $M_1$ and $M_2$ are the virtual images of the metrology end points, as seen from the star (in the sidereal reference frame). Then, the OPD measurement is twice differential:

$$\delta \text{OPD} = \text{OPD}_{2S\text{-INT}} - \text{OPD}_{1S\text{-INT}} + \text{OPD}_{1P\text{-INT}} - \text{OPD}_{2P\text{-INT}}. \quad (9)$$

Similarly to Sect. 3.1, we can combine Eq. (5), which links the internal OPs to the external OPs, to Eq. (6), which links the external OPs to the baseline vector $M_1M_2$, to derive

$$\text{OPD}_{2S\text{-INT}} - \text{OPD}_{1S\text{-INT}} = s \cdot \overrightarrow{M_1M_2} \quad (10)$$

$$\text{OPD}_{2P\text{-INT}} - \text{OPD}_{1P\text{-INT}} = p \cdot \overrightarrow{M_1M_2}. \quad (11)$$

Then, we can define the narrow angle baseline:

$$\overrightarrow{M_1M_2} = \overrightarrow{L_1L_2} + B_{\text{WAB}} + \overrightarrow{L_2L_2} = B_{\text{NAB}}. \quad (12)$$

As shown in Woillez & Lacour (2013) it is the one that must be used to derive the separation between the two stellar objects:

$$\delta \text{OPD} = (s - p) \cdot B_{\text{NAB}}. \quad (13)$$

A difficulty with the narrow angle baseline is that the limit points (the end points of the metrology) are generally defined in the telescope reference frame. As a consequence, the NAB depends on the telescope pointing. Also, if the metrology end points are not physically located in the telescope frame, but inside the interferometric optical train, the narrow angle baseline is more difficult to calculate. It is then defined by the virtual images of the metrology end point, as seen from outside the telescope.

### 3.3. The imaging baseline

It is necessary to introduce the notion of the imaging baseline once we want to consider the impact of the field position of the targets on the metrology. It can be traced back to the van-Cittert Zernike theorem, which links contrast and phase to the spatial geometry of the astrophysical object (Woillez & Lacour 2013).

For a dual beam interferometer like GRAVITY, there are two imaging baselines: one for each beam combiner. This is because the limit points of the imaging baseline are defined by the effective pupil, which is different from the telescope pupil. For a
single-mode interferometer, the effective pupil can be properly determined. It is the acceptance mode of the single-mode fibers as seen from the star (viewed through the telescope pupil). For a dual beam combiner there are four acceptance modes, thus four limit points or two IMBs. In the lower panel of Fig. 3, the limit points are noted $I_{1S}$ and $I_{2S}$ for the first beam combiner, and $I_{1P}$ and $I_{2P}$ for the second. These limit points do not have to be collocated with the telescope pupil center: this is called a pupil positioning error, which can be both lateral and longitudinal.

The propagation directions of the metrology beams are represented in Fig. 3 by the vectors $m_{1S}$, $m_{2S}$, $m_{1P}$, and $m_{2P}$. The propagation directions of the incoming light of the stellar sources are still $s$ and $p$. That means that, for example, the vector difference $(m_{2S} - s)$ characterizes the pointing error of the metrology of the first beam combiner on the second telescope toward the source denoted $s$. This is equivalent to a tip-tilt error. The imaging baseline limit points have, by definition, an important property: the optical path length between the fringes and the limit points is the same for the metrology and for the stellar light. This is a property of the effective pupil: the pupil is field-independent. A consequence, $OP_{1S-INT}$ defined between the beam combiner and the imaging baseline limit point does not depend on the $(m_{2S} - s)$ tilt error. This is also the case for the three other optical path lengths: $OP_{2S-INT}$, $OP_{1P-INT}$, and $OP_{2P-INT}$.

However, the $\delta OPD$ measurement is done up to the metrology end point. Following the conventions from the lower panel of Fig. 3, Eq. (9) must be rewritten to include the optical path length corresponding to the projected distance between the imaging and metrology end points (the four $m \cdot IM$ terms). Thus, $\delta OPD$ measured between $M_1$ and $M_2$ becomes

$$\delta OPD_{M_12} = OP_{2S-INT} + m_{2S} \cdot I_{2S} \, I_{2M_2}^2$$

$$- OP_{1S-INT} - m_{1S} \cdot I_{1S} \, I_{1M_1}^2$$

$$+ OP_{1P-INT} + m_{1P} \cdot I_{1P} \, I_{1M_1}^2$$

$$- OP_{2P-INT} - m_{2P} \cdot I_{2P} \, I_{2M_2}^2. \quad (14)$$

On the other hand, the internal and external optical path lengths must satisfy Eq. (5). In the same way that we obtained Eqs. (10) and (11), we can derive

$$OP_{2S-INT} - OP_{1S-INT} = s \cdot I_{1S} \, I_{1S}^2$$

$$OP_{2P-INT} - OP_{1P-INT} = p \cdot I_{1P} \, I_{1P}^2, \quad (15)$$

where the vectors $I_{1S} \, I_{1S}$ and $I_{1P} \, I_{1P}$ correspond to the two imaging baselines:

$$I_{1S} \, I_{1S} = I_{1S} \, I_{L1} + B_{WAB} \, I_{L2} \, I_{2S} = B_{IMB-S} \quad (17)$$

$$I_{1P} \, I_{1P} = I_{1P} \, I_{L1} + B_{WAB} \, I_{L2} \, I_{2P} = B_{IMB-P}. \quad (18)$$

Combining together Eqs. (14)–(18), we can link the $\delta OPD_{M_12}$ measurement to the differential angular position $(s - p)$ and to the wide angle baseline $B_{WAB}$:

$$\delta OPD_{M_12} = (s - p) \cdot B_{WAB} + m_{2S} \cdot (m_{2S} - s) \cdot I_{1S} \, I_{1M_1}^2$$

$$+ m_{1P} \cdot (m_{1P} - p) \cdot I_{1P} \, I_{1M_1}^2$$

$$+ s \cdot (I_{1S} \, I_{1L1} + I_{1S} \, I_{L2})$$

$$- p \cdot (I_{1P} \, I_{L1} + I_{2P} \, I_{L2}). \quad (19)$$

This equation simplifies to

$$\delta OPD_{M_12} = (s - p) \cdot B_{WAB} + (m_{2S} - s) \cdot I_{1S} \, I_{1M_1}^2$$

$$+ (m_{1P} - p) \cdot I_{1P} \, I_{1M_1}^2 \quad (20)$$

or to

$$\delta OPD_{M_12} = (s - p) \cdot B_{WAB} + (m_{2S} - s) \cdot I_{1S} \, I_{1M_1}^2$$

$$+ (m_{1P} - p) \cdot I_{1P} \, I_{1M_1}^2. \quad (21)$$

Hence, $\delta OPD_{M_12}$ results from the sum of two components. The first component $(s - p) \cdot B_{WAB}$ is the main contributor to the full measurement. It reflects the assertion stated in Woillez & Lacour (2013) that the NAB baseline is the main baseline of a dual feed astrometric interferometer. However, several second-order terms are present because the imaging baseline limit points are not necessarily the same as the metrology end points $(\vec{IM} \neq \vec{0})$. Since both $\vec{IM}$ and $(m - s)$ are small, the amplitude of the scalar product will be almost negligible compared to the total optical path difference. However, if no care is taken, it adds to the $\delta OPD_{M_12}$ in such a way that it will bias the estimation of the angular separation $(s - p)$.

It is possible to combine the second-order terms to show the influence of the differential length between the imaging baselines $B_{IMB}$ and the narrow angle baseline $B_{NAB}$:

$$\delta OPD_{M_12} = (s - p) \cdot B_{NAB}$$

$$+ (m_{2S} + m_{1S}) / 2 - s \cdot (B_{NAB} = B_{IMB-S})$$

$$+ (m_{2P} + m_{1P}) / 2 - p \cdot (B_{NAB} = B_{IMB-P})$$

$$+ (m_{2S} - m_{1S}) \cdot I_{1S} \, I_{1M_1}^2$$

$$+ (m_{2P} - m_{1P}) \cdot I_{1P} \, I_{1M_1}^2$$

$$+ (m_{2S} - m_{1S}) \cdot I_{1S} \, I_{1M_1}^2$$

$$+ (m_{2P} - m_{1P}) \cdot I_{1P} \, I_{1M_1}^2 \cdot (I_{1S} \, I_{1S} + I_{1S} \, I_{1P}$$

$$+ (m_{2P} - m_{1P}) \cdot I_{1S} \, I_{1M_1}^2) + (m_{2S} - m_{1S}) \cdot I_{1S} \, I_{1M_1}^2$$

$$+ (m_{2P} - m_{1P}) \cdot I_{1P} \, I_{1M_1}^2 \cdot (I_{1S} \, I_{1S} + I_{1S} \, I_{1P}$$

$$+ (m_{2P} - m_{1P}) \cdot I_{1S} \, I_{1M_1}^2). \quad (22)$$

According to this formula, $\delta OPD_{M_12}$ can be equal to $\delta OPD$ as spelled out in Eq. (13) if all the three following conditions are fulfilled:

1. The baselines match each other: $B_{NAB} = B_{IMB-S} = B_{IMB-P}$
2. The limit points of the two imaging baselines are superposed: $I_{1P} \, I_{1S} = I_{1S} \, I_{2P}$ (equal to 0 assuming 1)
3. The metrology beams have the same propagation angle on each telescope: $m_{1S} - m_{1P} = m_{2S} - m_{2P}$.

This is not a situation that we assume in the following GRAVITY error budget.
4. High-order optical aberrations

The notion of imaging baseline is intrinsically related to the presence of a pointing error, or tip tilt (\(m-s \neq 0\)). These tilt-tilt errors are equivalent to low-order optical aberrations and are considered in the previous section. In this section, we include the effect of the higher order Zernike polynomials.

To be exact, a full diffraction analysis would have to be performed over the whole metrology train. That would mean including a Fresnel analysis over the optics down to the fiber for the stellar light and up to the metrology sensor for the metrology beam. However, we can considerably simplify the problem by neglecting the effect of the \(z\) direction inside the optical train. It is equivalent to considering that the atmosphere, the metrology sensor, the pupil of the telescopes, the pupil of the instruments, and the optical aberrations are conjugated inside a single plane. In that plane, only the two lateral directions \((x, y)\) matter. Then, the effect of the high-order aberrations can be seen as a problem of a two-dimensional coupling of the light onto the modes of the beam combiners.

4.1. Aberrations affecting the stellar light

In the case of a single-mode interferometer, the modes of the beam combiners can be identified as the modes of acceptance of the single-mode fibers. The phase delays due to the optical aberrations are therefore the argument of the coupling coefficient of the abnormal light onto the mode of each fiber: \(\delta \Phi = \arg(\rho)\), where \(\delta \Phi\) is the phase error caused by the aberration, and \(\rho\) the coupling coefficient of the light into the fiber (Robieux 1959; Shaklan & Roddier 1988):

\[
\rho = \left( \frac{\int |E(x, y)|^2 dxdy}{\int |E_p(x, y)|^2 dxdy} \right)
\]

(23)

where \(E(x, y)\) is the incoming aberrated wavefront, and \(E_p(x, y)\) is the acceptance mode for the electric field of the single-mode fiber. The cross-correlation is easier to understand in the focal plane, i.e., at the entrance of the fiber. However, according to the Parseval-Plancherel theorem, it can also be calculated in the pupil plane, which turns out to be more convenient for modeling the aberrations.

To calculate the impact on the OPD measurement, we have to consider the modes of all four fibers \((E_{p1}, E_{p2}, E_{p1}, \text{and } E_{p2})\) and of the four distinct incoming wavefront \((E_{s1}, E_{s2}, E_{s1}, \text{and } E_{s2})\). Then, the phase error on the OPD measurement will be

\[
\delta \Phi_{\text{stellar}} = \arg \left( \int_{\text{pupil}} E_{s1}(x, y)E_{p1}(x, y)\, dxdy \right)
\]

\[ - \arg \left( \int_{\text{pupil}} E_{s2}(x, y)E_{p2}(x, y)\, dxdy \right) \]

\[ + \arg \left( \int_{\text{pupil}} E_{s1}(x, y)E_{p2}(x, y)\, dxdy \right) \]

\[ - \arg \left( \int_{\text{pupil}} E_{s2}(x, y)E_{p1}(x, y)\, dxdy \right) \]

(24)

where the integration is done over the electric fields present inside the pupil.

4.2. Aberrations affecting the metrology light

While the phases measured inside the beam combiners are affected by the full aberration of the stellar light, the metrology value is modified solely by the optical aberrations inside the optical train. These aberrations can be quantified according to the Zernike polynomials \(Z^n(x, y)\) (\(n\) according to numbering by Noll 1976) of variance unity. Each one of the four metrology beams can be affected by these aberrations. However, only a differential aberration will affect the differential metrology measurement. This is a beam-walk effect: the two stellar beams have different footprints on the optics, which can result in a differential optical aberration. Assuming that the aberrations are a sum of Zernike of amplitude \(k_{s1, p1}, k_{s2, p1}, k_{s1, p2}, \text{and } k_{s2, p2}\) (one for each one of the four metrology beams), we can derive a phase error to the metrology:

\[
\delta \Phi_{\text{met}} = \int_{\text{receivers tel1}} \sum_{n} \left( k_{s1, p1} - k_{s1, p2} \right) Z^n(x, y)\, dxdy
\]

\[ - \int_{\text{receivers tel2}} \sum_{n} \left( k_{s2, p1} - k_{s2, p2} \right) Z^n(x, y)\, dxdy. \]

(25)

The full metrology measurement is a phase measurement averaged over all the metrology receivers. This is a general case. Depending on the type of metrology system and position/number of receivers, this integral can be a simple summation of several measurements at given \((x, y)\) positions.

5. The GRAVITY instrument

5.1. GRAVITY metrology principle

As mentioned before, GRAVITY is a four-telescope interferometer for the VLTI. It is operating in the \(K\) band and is able to couple the four 8.2 m Unit Telescopes (UTs) or the four 1.8 m auxiliary telescopes (ATs). It is assisted with a visible and infrared adaptive optics (AO) system for the UTs. GRAVITY is equipped with two beam combiners in integrated optics (Jocou et al. 2012), one for a reference source and the other for the science target. In this dual mode, the goal of GRAVITY is to provide an accurate differential phase measurement to achieve astrometry with ten micro-arcsecond accuracy in a two-arcsecond field (4-arcsec field with the ATs).

The reference source is used to cancel atmospheric perturbations over the six baselines thanks to a dedicated fringe tracking system. To do so, the GRAVITY instrument will have a fast feedback control loop using steering mirrors on piezo actuators inside the cryostat (with a cutoff bandwidth close to 30 Hz) and a slow feedback (for offloading) to the main VLTI delay lines. GRAVITY will have to fringe track on a \(K = 10\) unresolved object with 300 nm OPD residuals, setting the basic specification of the system.

The basic functions impending the OPD measurement are presented in Fig. 4. The metrology is injected into both beam combiners (Gillesen et al. 2012). On one beam combiner, the metrology is phase-modulated with respect to the other with a lithium niobate phase modulator. Going through the optics in the reverse direction of the stellar light, the metrology beams encounter in the cryostat (Haug et al. 2012): the fibered differential delay lines, the fibered polarization rotators, the pupil actuator mirror, and the fast piezo actuators. Before exiting the cryostat, part of the metrology light is recombined on internal metrology diodes to provide constant and reliable feedback for the differential fiber delay lines (without risk of beam interruption).
rest of the metrology light is back-propagated through the VLTI optical train, across all the optics including M1, up to receivers situated on each four spider arms of each of the four telescopes. For each telescope a phase value is extracted from the mean of the phase value of the four telescope diodes: \( \Phi_Mi \), with \( i \) the telescope number.

The stellar light is following a similar path in reverse. Once in the cryostat, after the pupil actuator mirror and before injection into the fluoride fibers, the astrophysical sources are separated with a roof mirror acting as a field separator (Pfuhl et al. 2012). The light is guided inside the single-mode fibers up to the pupil actuator mirror, and imaged on infrared cameras. On one hand, the light of the bright reference source is sent to a fast eAPD detector (the SAPHIRA detector, Finger et al. 2012) from which a phase measurement is read for each five wavelength channels of the fringe tracker: \( \Phi_P(\lambda_i) \). This phase is kept close to zero by the means of the OPD error equal to \( 0 \), as the OPD length is proportional to the index of refraction of the medium through which it propagates. It writes as

\[
\Delta \text{OPD error} = 2 \pi \left( \Phi_P(\lambda_i) - \Phi_Mi \right) \frac{\Delta \lambda}{\lambda}.
\]

5.2. Dispersion in GRAVITY

In Eq. (26) is missing an important term caused by dispersion. Dispersion can have an important impact on the contrast of the white-light fringes (Coudé du Foresto et al. 1995). But it also have an important effect on the astrometry. It comes from the fact that the apparent optical path length depends on the wavelength, as the OP length is proportional to the index of refraction of the medium through which it propagates. It writes as

\[
\text{OP}(\lambda) = n(\lambda) \cdot L,
\]

where \( \text{OP}(\lambda) \) is the apparent optical path length and \( L \) the physical length. For an addition of several media along the optical path length measured by the metrology, dispersion introduces an OPD error equal to

\[
D(\lambda_m, \lambda_s) = \sum_{\text{med}} \Delta_\text{med} \cdot \left( n^\text{med}_s - n^\text{med}_m \right).
\]

where \( \Delta_\text{med} \) is the difference of physical length of each optical medium, and \( n^\text{med}_s \) is the refractive index of the corresponding medium at a wavelength \( \lambda_s \).

In GRAVITY, the wavelength of the metrology (\( \lambda_m = 1.908 \, \mu m \)) is different from the wavelength of the stellar light (1.95 \( \mu m < \lambda_s < 2.45 \, \mu m \)). On the path of the metrology, there are three different media: air, vacuum (inside the cryostat), and fluoride glass (fibers). The dispersion of vacuum is zero, so the dispersion term can be written as

\[
D(\lambda_m, \lambda_s) = \Delta\text{air} \cdot \left( n^\text{air}_s - n^\text{air}_m \right) + \Delta\text{fiber} \cdot \left( n^\text{fiber}_s - n^\text{fiber}_m \right).
\]
the phase measurements to the angular separation of the targets on sky:

\[ (s - p) \cdot B_{\text{NAB}} = (\Phi_{\text{S}} - \Phi_{\text{p}}) \frac{\lambda_1}{2\pi} + (\Phi_{\text{M}_s} - \Phi_{\text{M}_p}) \frac{\lambda_m}{2\pi} - D(\lambda_m, \lambda_s) + \text{error terms (2nd order).} \]  

(30)

It is not foreseen to calculate the value of the second-order error terms. The goal is to keep them low enough in the error budget. They come from baseline errors and optical aberrations, as spelled out in Eqs. (21), (24), and (25):

error terms \[= (m_{1S} - s) \cdot \overrightarrow{I}_S M'_1 - (m_{2S} - s) \cdot \overrightarrow{I}_S M'_2 + (m_{1p} - p) \cdot \overrightarrow{I}_p M'_1 - (m_{2p} - p) \cdot \overrightarrow{I}_p M'_2 + \delta \Phi_{\text{stellar}} \frac{\lambda_1}{2\pi} + \delta \Phi_{\text{met}} \frac{\lambda_m}{2\pi}. \]  

(31)

5.3. Geographical position of the limit-points of GRAVITY

The idea that drove the design of the GRAVITY instrument is therefore to have monitoring and steering possibilities for both the pupil and the field plane. This is paramount for keeping the error terms in Eq. (31) under control. Focal and pupil plane positions are simultaneously supervised during observation by a dedicated camera situated inside the cryostat (Amorim et al. 2012).

The feedback system of this camera – called the acquisition camera – is presented in yellow in Fig. 4. The first terms to minimize are the field terms: \( m_{1S} - s \) and \( m_{2S} - s \) for each telescope \( i \). To do so, the targets are imaged in the \( H \) band on the acquisition camera, and the sources are maintained on the fibers thanks to the tri-axis piezo mirror situated in the pupil plane of the instrument. The second terms to minimize are the mismatch between the imaging baseline limit points and the metrology endpoints: i.e., \( \overrightarrow{I}_S M'_1 \) and \( \overrightarrow{I}_p M'_1 \). To do so, four light sources (emitting at a wavelength of 1.2 \( \mu \)m) are placed on the spider arms of the telescopes. They are the center of rotation of the telescope, the telescopes are, in theory, the easiest to locate in geographical coordinates. They are the center of rotation of the telescope, which can be determined by external measurements. However, problems can come from bending of the structure that may hinder the notion of a fixed pivot point.

On the other hand, the problem of the NAB and IMB limit points are that they are defined in telescope frame coordinates. Because of the telescope optical train (notably M2), the telescope frame coordinates are not necessarily colinear to the celestial frame of the object: the telescope reference frame can be several arcsec away from the direction of the target. The terrestrial coordinates are defined as (\( u, v, w \)). The telescope coordinates are defined as (\( x, y, z \)), with (\( 0, 0, 0 \)) the coordinate of the pivot point, (\( x, y \)) the transverse coordinates, and (\( z \)) the longitudinal coordinates.

On the UTs, the metrology measurements will be obtained thanks to four photodiodes positioned on the spider arms, symmetrically around the pupil. The (\( x, y, z \)) coordinates are (1.945, 1.692), (−1.945, 1.692), (1.945, −1.692), and (−1.945, −1.692) in meters. The positions were chosen to cancel any small defocus in the non-common path (Sect. 5.4). The metrology measurement \( \Phi_{\text{NAB}} \) will be the mean of the four phase measurements. The resulting limit point of the NAB will thus be the mean position of the four metrology sensors. It gives a NAB limit point close to the center of the beam (\( (x_{\text{NAB}}, y_{\text{NAB}}) \approx 0 \)). In longitudinal coordinates, the spider arms are at \( z \approx 10 \text{ m} \) above the pivot point.

The uncertainty of the position of the imaging baseline limit points is larger due to the relay optics between the cryostat and the telescope pupil. As a result, the effective pupils can move with respect to the telescope pupil. However, on GRAVITY, the internal pupil tracking system will stabilize the effective pupils close to the telescope pupil. The imaging limit points are expected to be at (\( x, y \pm 4 \text{ cm} \) from the central optical axis. The tracking in the longitudinal direction should also be able to perform with an expected accuracy of \( z \approx 10 \text{ km} \) between the metrology end points and the imaging baseline limit points.

5.4. GRAVITY vulnerability to high-order aberrations

The two incoming stellar fields \( E_{\text{ff}}(x, y) \) and \( E_{\text{p}}(x, y) \) are defined over the eight meters of the telescope pupil except for the central 1.116 m in diameter due to secondary mirror obscuration (0.1395 of the pupil radius). Their phases are affected by both the atmospheric turbulence and the optical aberrations in the optical train.

The electromagnetic fields accepted by the fibers are also defined in the pupil: \( E_{\text{fib}}(x, y) \) and \( E_{\text{p}}(x, y) \). These fields have a near-Gaussian profile whose amplitude can be approximated by

\[ E_{\text{fib}}(x, y) = E_{\text{p}}(x, y) = \exp \left( -\frac{x^2 + y^2}{2\sigma_{\text{p}}^2} \right). \]

(32)

where \( \sigma_{\text{p}} \) was chosen to maximize the injected flux at optimal correction (for a perfectly flat incoming wavefront), and \( \sigma_{\text{p}} = 0.644573 \) in units of pupil radius. The phases of the \( E_{\text{fib}} \) and \( E_{\text{p}} \) electric fields are only important relative to the phase of the incoming stellar fields (\( E_{\text{s}} \) and \( E_{\text{p}} \), respectively), so we suppose the phases of the acceptance fields equal to zero. The phase errors resulting from these fields are therefore

\[ \delta \Phi_{\text{stellar}} = \arg \left( \int \int_{0.1395 < r < 1} E_{\text{fib}}(x, y) \exp \left( -\frac{x^2 + y^2}{2\sigma_{\text{p}}^2} \right) \text{d}x\text{d}y \right) \]

\[ - \arg \left( \int \int_{0.1395 < r < 1} E_{\text{p}}(x, y) \exp \left( -\frac{x^2 + y^2}{2\sigma_{\text{p}}^2} \right) \text{d}x\text{d}y \right) \]

\[ - \arg \left( \int \int_{0.1395 < r < 1} E_{\text{p}}(x, y) \exp \left( -\frac{x^2 + y^2}{2\sigma_{\text{p}}^2} \right) \text{d}x\text{d}y \right) \]

\[ + \arg \left( \int \int_{0.1395 < r < 1} E_{\text{fib}}(x, y) \exp \left( -\frac{x^2 + y^2}{2\sigma_{\text{p}}^2} \right) \text{d}x\text{d}y \right) \]

where \( \rho = \sqrt{x^2 + y^2}. \)

\[ z = 10 \text{ km} \text{ in the 8 m beam is equivalent to } z = 1 \text{ m} \text{ in the 80 mm beam in the VLTI laboratory.} \]
GRAVITY will have four metrology receivers on each telescope. According to Eq. (25), the effect of internal aberrations is therefore

\[
\delta \Phi_{\text{met}} = \frac{1}{4} \sum_{d=1}^{4} \sum_{n} \left[ (k_{n}^x - k_{p1}^x) - (k_{n}^y + k_{p2}^y) \right] Z_{n}(x_{d}, y_{d}) \tag{33}
\]

where \((x_{d}, y_{d})\) are the positions of the four metrology sensors, and \(k_{n}^x, k_{p1}^x, k_{n}^y, k_{p2}^y\) the amplitude of the aberrations affecting each one of the four metrology beams.

The four \((x_{d}, y_{d})\) sensors' coordinates can be chosen so that optical aberrations affecting both stellar and metrology beams can be fully compensated: \(\delta \Phi_{\text{stellar}} + \delta \Phi_{\text{met}} = 0\). Because of the limited number of metrology receivers, the number of Zernike polynomials whose effect can be nullified is limited. To cancel a polynomial \(n\), the \((x_{d}, y_{d})\) coordinates must satisfy the relation

\[
\frac{1}{4} \sum_{d=1}^{4} Z_{n}(x_{d}, y_{d}) = \int_{0.1395 < \rho < 1} \int_{0.1395 < \rho < 1} Z_{n}(x, y) \exp \left( -\frac{x^2 + y^2}{2\sigma^2} \right) \mathrm{d}x \mathrm{d}y,
\]

in the approximation of low-amplitude phase errors.

The choice made for GRAVITY was to position the metrology sensors to ensure that piston, tip, tilt, and defocus are canceled by the metrology. For example, the fourth polynomial, \(Z_{4}(x, y) = 2 \sqrt{3}(x^2 + y^2) - \sqrt{3}\) (defocus, Noll 1976), puts a constraint on the distance between the center of the pupil to the position of the sensors:

\[
\rho_{d}^2 = \frac{\int_{0.1395 < \rho < 1} 2\pi \rho^3 \exp \left( -\frac{\rho^2}{2\sigma^2} \right) \mathrm{d}\rho}{\int_{0.1395 < \rho < 1} 2\pi \rho^2 \exp \left( -\frac{\rho^2}{2\sigma^2} \right) \mathrm{d}\rho}
\]

(35)

gives \(\rho_{d} = 0.644573\) in unit of pupil radius. However, depending on the position of the metrology sensor, the influence of the other Zernike polynomials in the optical train can be amplified. The idea for GRAVITY was to cancel the aberrations that can be caused by a simple displacement of the single-mode fibers in the focal plane. The influence of the higher order polynomials is investigated numerically in Sect. 6.4.1.

5.5. Polarization in GRAVITY

Much of the preceding discussion has revolved around the spatial properties of the GRAVITY instrument associated with the VLTI optical train, their differential effects on science and reference stars, and the metrology light. The discussion would not be complete without also addressing the polarization states of light. The combined optical train of the VLTI and the GRAVITY instrument is birefringent. This causes phase errors in two ways. First, the differential phase between the metrology beams for the science and reference stars is impacted (in any telescope) if they do not enter the optical train with identical polarization orientation. For the expected magnitude of birefringence, this phase error can be made negligible with an easily achievable alignment.

\(\rho_{d} = \sigma_{p}\), a consequence of injection optimisation.
accuracy. A second, more subtle effect, results from the combination of a) differential birefringence between optical trains and b) different polarization states of the science and reference starlights. In fact, the light from gas orbiting the galactic center black hole (a prime science target for GRAVITY) is known to be partially linearly polarized.

A detailed account is beyond the scope of the present paper, but will be given in a forthcoming publication (Lazareff et al., in prep.). We plan to operate with the assumption of an unpolarized reference star, and a partially linearly polarized science target. Under that condition, fringes detected in a linear polarization (at output) perpendicular to either eigenvector are immune from the phase error. Then, the operation can be summarized as follows: i) define (in our particular context) eigenvectors of the optical train as linear polarizations (at input) that result in linear polarization at output. ii) Rotate the polarization controllers of GRAVITY to realize the orthogonality condition. Because the birefringence of each mirror train is a function of the pointing parameters, this alignment condition will be determined in situ within a short time interval from the actual observation or according to polarization modeling acquired during dedicated campaigns.

6. Astrometric error budget

The following sections derive GRAVITY’s astrometric error budget from the expression of the astrometric measurement established from Eqs. (30) and (31). The goal is to estimate the bias that may affect the measurement of \( s - p \). Section 6.1 considers the imprecisions on the \( B_{WAB} \) baseline vector. Section 6.2 focuses on the combined effect of the tip-tilt error (the \( m_{s} - s \) and \( m_{p} - p \) terms) with a pupil-matching error (\( IM_{\text{ff}} \)). Section 6.3 quantifies the errors brought by the dispersion in the interferometer. (The \( D(\lambda_{s}, \lambda_{p}) \) value.) Last, in Sect. 6.4, the errors caused by the high-order aberrations are computed: \( \Delta \Phi_{\text{stellar}} \) and \( \Delta \Phi_{\text{met}} \).

6.1. NAB determination in celestial coordinates

That the limits points of the metrology are not at the same position as the pivot points (\( L_{\text{ff}} \neq 0 \) in Eq. (21)) mostly raises a problem of reference systems. The WAB baseline is defined in terrestrial coordinates and fixed in this frame of reference. On the other hand, the narrow angle baseline is defined in a different frame system: the telescope coordinates. Finally, the stellar target and the terrestrial coordinates depend on i) the knowledge of the target coordinates and ii) the local stellar time (LST). We assume that time accuracy will be the limiting factor. Not because the LST is not precisely known, but simply because integration times on the science camera are several seconds. Putting a precise time stamp on the data will therefore be difficult, especially since it would have to be weighted by injection coefficients during the exposure. We estimate our uncertainties to \( 1/15 \, \text{s} \), meaning \( 1'' \) on sky, knowing that the earth rotate by 15 arcsec per second. The error caused by an arcsec on the projection of a 100 m baseline is approximately 0.5 mm (decreasing with the proximity to the zenith).

With respect to a 100 m baseline, a 0.5 mm error in baseline will correspond to an error in the astrometric accuracy of 5 ppm, hence 5 \( \mu \)s for two stars separated by 1''. The NAB indetermination factors are summarized in Table 1.

6.2. IMB to NAB mismatch

The IMB has an effect on the astrometric measurement only if i) the metrology beam is not exactly pointing toward the stellar source (\( m_{s} - s \neq 0 \)) and if ii) the limits points of the IMB and the NAB do not agree (\( IM_{\text{ff}} \neq 0 \)). The effect on the OPD measurements is an additional term for each telescope: \( (m_{S} - s) \cdot I_{S} M_{S} - (m_{P} - p) \cdot I_{P} M_{P} \), with \( i \) the telescope number. Although it is technically possible to model its influence, owing to the complexity of this term, it is better to try to nullify it. This is achieved in GRAVITY thanks to the pupil-tracking camera mentioned in Sect. 5.3. The residuals are taken into account in the error budget.

The pupil-tracking camera will track the pupil both on the longitudinal and lateral directions. But owing to its design, the longitudinal accuracy is considerably lower than for the lateral. With the pupil tracking activated, the pupil vector \( IM_{\text{ff}} \) is expected to be within 10 km in the longitudinal and 4 cm in the lateral direction (\( z = 0 \pm 10 \, \text{km} \) and \( |x, y| = 0 \pm 4 \, \text{cm} \)). In the GRAVITY design (as shown in Fig. 5) the metrology beam is injected inside the IO beam combiner. As a result, by optimizing the injection of the stellar flux inside the fibers, GRAVITY will also make sure that the metrology beam has a propagation direction nearly parallel to the stellar light (but with opposite direction), so for optimum injection, \( m_{s} - s = 0 \). Practically speaking, the GRAVITY acquisition camera should stabilize the injection during observation enough to be able to reach \( |m - s| \leq 0.2\, \text{L/D} \), or \( \approx 10 \, \text{mas} \) at \( \lambda = 2.2 \, \mu \text{m} \).

We can split the OPD error induced by the IMB to NAB mismatch into two terms: longitudinal and lateral. The lateral contribution is the 4 cm of the pupil displacement projected toward the normalized (\( m_{s} - s \)) vector: \( \sin(10 \, \text{mas}) \times 4 \, \text{cm} = 2 \, \text{nm} \). The longitudinal contribution will be \( |1 - \cos(10 \, \text{mas})| \times 10 \, \text{km} = 0.01 \, \text{nm} \). Related to the total OPD expected between two target separated by 1'' observed with a baseline of 100 m, it means that the IMB...
to NAB mismatching will result for GRAVITY in an additional astrometric error of 4 μas. To account for each of the four terms, this value is multiplied by \(\sqrt{4}\) in Table 1.

### 6.3. Dispersion

According to Eq. (29) the dispersion term writes as

\[
D(\lambda_m, \lambda_s) = \Delta_{\text{air}} \cdot (n_{\text{air}}^\lambda - n_{\text{air}}^{\lambda_s}) + \Delta_{\text{fiber}} \cdot (n_{\text{fiber}}^\lambda - n_{\text{fiber}}^{\lambda_s}),
\]

(36)

where \(\Delta_{\text{air}}\) and \(\Delta_{\text{fiber}}\) are the differential delays in air and glass, respectively. Ideally, the differential path is entirely compensated for in fibers: \(\Delta_{\text{air}} \approx 0\) and \(\Delta_{\text{fiber}} = \delta_{\text{OPD}}M\). This is not exactly the case owing to the non-common path in air or vacuum. The main errors come from uncertainties in the physical length of fiber (\(\Delta_{\text{fiber}}\)) and of air (\(\Delta_{\text{air}}\)). But uncertainties in the knowledge of the refractive indices (\(n_{\text{air}}^\lambda - n_{\text{air}}^{\lambda_s}\)) can also be a problem, since they depend on environmental conditions.

Following the Owens (1967) formula, we computed the dispersion of air for standard atmospheric conditions. It gives

\[
\Delta_{\text{air}} = 2.2 \mu m, \quad \Delta_{\text{fiber}} = 10 \mu m.
\]

(37)

hence a dispersion effect of 80 pm for \(\Delta_{\text{air}} = 1 \mu m\). This is a conservative upper limit, since the two stellar beams share the same optics up to the cryostat. (The star separator is inside the cryostat.) In the data reduction software, air dispersion will therefore be neglected.

The fluoride fibers are, however, a lot more dispersive (Coudé du Foresto et al. 1995). Thus, their effect cannot be neglected. Because the fibers are inside the cryostat, we assume that the dispersion in the fibers will not change and that it can be precisely modeled. First laboratory results show dispersion coefficients on the order of

\[
\Delta_{\text{fiber}} = 2 \times 10^{-4}.
\]

(38)

This is in accordance with Coudé du Foresto et al. (1995). For a physical length \(\Delta_{\text{fiber}} = 1 \mu m\), the correction term to the OPD is therefore on the order of 200 nm. With a knowledge of the fiber stretching of the differential delay lines of \(\sigma(\Delta_{\text{fiber}}) = 10 \mu m\), we can model the \(D(\lambda_m, \lambda_s)\) to a precision of 2 nm.

### 6.4. Optical aberrations

These aberrations can be split into two distinct categories:

- The common-path aberrations that affect both stellar signals. This is the case of atmospheric perturbations, partially corrected by the AO on the UTs.
- The non-common path aberrations that affect only one of the stars. This is typically a consequence of a beam-walk effect, when the two stellar beams do not have the same imprint on the optical surfaces.

We assume that the metrology will monitor the non-common path optical aberrations, but not the common-path aberrations. We neglect isoplanatism as being beyond the scope of this paper.

#### 6.4.1. Non-common path aberrations – beam-walk defects

The goal of the metrology is to monitor and account for the non-common optical path length between the two beam combiners. By construction, the metrology laser does perfectly probe any piston induced in the optical train. Higher orders are not necessarily measured as well. To simulate the effect of non-common path aberrations, we consider the influence of Zernike polynomial (\(n\)) individually, affecting only one beam of the metrology (\(k^1 = k^0 = k^p = k^q = 0\)). Then,

\[
\delta \Phi_{\text{stellar}} = \frac{1}{4} \sum_{d=1}^{\infty} k^d \cdot Z^d(x, y),
\]

(39)

The same aberration also affects the stellar light, but with a negative sign,

\[
E_s(x, y) = \exp(-ik^dZ^d(x, y)),
\]

(40)

hence

\[
\delta \Phi_{\text{stellar}} = \arg \left( \int \exp(-ik^dZ^d(x, y)) \exp \left( -\frac{x^2 + y^2}{2\sigma^2} \right) \, dx \, dy \right).
\]

(41)

For PRIMA, analytical models predict a precision of 10 μas for one hour of observations between two targets separated by 10 arcsec (Sahlmann et al. 2013).
Table 2. Error caused by non-common path aberrations.

<table>
<thead>
<tr>
<th>Zernike#</th>
<th>rms (b)</th>
<th>(\delta \Phi_{\text{stellar}} \frac{\lambda}{2 \pi})</th>
<th>(\delta \Phi_{\text{met}} \frac{\lambda}{2 \pi})</th>
<th>(\delta \Phi_{\text{stellar}} \frac{\lambda}{2 \pi} + \delta \Phi_{\text{met}} \frac{\lambda}{2 \pi})</th>
</tr>
</thead>
<tbody>
<tr>
<td>4(f)</td>
<td>8 nm</td>
<td>−2.3 nm</td>
<td>−2.3 nm</td>
<td>0.0 nm</td>
</tr>
<tr>
<td>6(f)</td>
<td>8 nm</td>
<td>0.0 nm</td>
<td>1.1 nm</td>
<td>−1.1 nm</td>
</tr>
<tr>
<td>11(f)</td>
<td>2 nm</td>
<td>−0.1 nm</td>
<td>−2.7 nm</td>
<td>2.6 nm</td>
</tr>
<tr>
<td>12</td>
<td>2 nm</td>
<td>0.0 nm</td>
<td>−0.6 nm</td>
<td>0.6 nm</td>
</tr>
<tr>
<td>14</td>
<td>2 nm</td>
<td>0.0 nm</td>
<td>−1.4 nm</td>
<td>1.4 nm</td>
</tr>
<tr>
<td>22(f)</td>
<td>1 nm</td>
<td>0.1 nm</td>
<td>0.7 nm</td>
<td>−0.7 nm</td>
</tr>
<tr>
<td>24</td>
<td>1 nm</td>
<td>0.0 nm</td>
<td>0.1 nm</td>
<td>−0.1 nm</td>
</tr>
<tr>
<td>26</td>
<td>1 nm</td>
<td>0.0 nm</td>
<td>1.6 nm</td>
<td>−1.6 nm</td>
</tr>
<tr>
<td>28</td>
<td>1 nm</td>
<td>0.0 nm</td>
<td>−0.1 nm</td>
<td>0.1 nm</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td></td>
<td>3.7 nm</td>
</tr>
</tbody>
</table>

Notes. (a) Zernike numbering as defined by Noll (1976); (b) amplitude of the aberration due to the Zernike in nm; (c) defocus; (d) astigmatism; (e) 3rd-order spherical; (f) 5th-order spherical.

giving an OPD error according to Eq. (31) of \(\delta \Phi_{\text{stellar}} \frac{\lambda}{2 \pi} + \delta \Phi_{\text{met}} \frac{\lambda}{2 \pi}\) in nanometers.

The amplitude of the non-common path aberrations were measured during two test campaigns at Paranal in August and October 2010. They were used to validate the metrology principle. Two coherent beams were overlapped at an angle corresponding to the angular separation of two stars in the field of view of GRAVITY. The two planar waves interfere to create a characteristic fringe pattern inside the VLTI lab. This fringe pattern was projected into the VLTI, including beam expanders and delay lines, onto the pupil of UT4. Aberrations were measured on the M2 mirror from the scattered light, observed with a commercial IR camera mounted to the telescope structure. We found that deviations occurring at the edge of the pupil are smaller than 8 nm/2 nm for a time base of 30 min (for quadratic and forth order respectively, negligible above).

Table 2 summarizes the errors on the OPD measurement for each Zernike polynomial up to the seventh order (Zernike \# ≤36). Owing to the point symmetry of the fundamental mode of the fibers, as well as of the metrology sensor, most of the polynomials do not bias the measurement. Only the ones affecting the phase are shown in Table 2. The variance considered for each one of these Zernike is 8 nm for the quadratic order, 2 nm for the fourth order, and 1 nm for the sixth order. The quadratic sum of all these terms is shown in the last row, which must be multiplied by \(\sqrt{2}\) to account for errors induced by the two arms of the interferometer.

6.4.2. Common-path aberrations – atmospheric perturbations

The atmosphere perturbations affect the phase of the electromagnetic field coming from the stars. If we neglect the influence of anisoplanatism, we can consider a similar electric field coming from both stars \(E_p(x, y) = E_s(x, y) = E(x, y)\). The atmospheric perturbations are therefore common-path aberrations. They are not probed by the metrology (\(\delta \Phi_{\text{met}} = 0\)) and affect the optical path difference as

\[
\delta \Phi_{\text{stellar}} = \arg \left( \iint_{\text{pupil}} E(x, y) E_p^*(x, y) \, dx \, dy \right) - \arg \left( \iint_{\text{pupil}} E(x, y) E_p^*(x, y) \, dx \, dy \right) \quad (42)
\]

Table 3. Error caused by atmospheric perturbations.

<table>
<thead>
<tr>
<th>(\Delta \rho)</th>
<th>UTs (\rho)</th>
<th>UTs(^a) (\rho)</th>
<th>ATs (\rho)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0/1/D</td>
<td>0 nm</td>
<td>0 nm</td>
<td>0 nm</td>
</tr>
<tr>
<td>0.14/1/D</td>
<td>4 nm</td>
<td>96 nm</td>
<td>53 nm</td>
</tr>
<tr>
<td>0.24/1/D</td>
<td>9 nm</td>
<td>173 nm</td>
<td>123 nm</td>
</tr>
<tr>
<td>0.34/1/D</td>
<td>15 nm</td>
<td>240 nm</td>
<td>190 nm</td>
</tr>
<tr>
<td>0.44/1/D</td>
<td>22 nm &gt; 300 nm(^b)</td>
<td>243 nm</td>
<td></td>
</tr>
<tr>
<td>0.54/1/D</td>
<td>32 nm &gt; 300 nm(^b)</td>
<td>&gt; 300 nm(^a)</td>
<td></td>
</tr>
</tbody>
</table>

Notes. (a) Without AO; (b) lower limit due to \(2\pi\) wrapping.

where \(E_p^*(x, y)\) and \(E_s^*(x, y)\) are the accepted mode of the fibers.

The effect of atmospheric turbulence is computed here for one telescope alone, since the atmosphere on the two telescopes can be considered as uncorrelated.

In the absence of any non-common path aberrations \((E_p^*(x, y) = E_s^*(x, y))\), common path aberrations do not change the measured optical path length: \(\delta \Phi_{\text{stellar}} = 0\). It is therefore necessary to combine them with some level of field and/or pupil offset. Considering a field offset, i.e., a tip-tilt error, we used the following equations to simulate the acceptance fields:

\[
E_p(x, y) = \exp \left( 2\pi i \frac{\Delta \rho \cdot x}{\lambda s} - \frac{x^2 + y^2}{2\sigma_p^2} \right) \quad (43)
\]

\[
E_p(x, y) = \exp \left( -\frac{x^2 + y^2}{2\sigma_p^2} \right) \quad (44)
\]

where \(\Delta \rho\) is the tip-tilt error. Such a tip-tilt error will necessarily be present because the alignment of the fiber cannot be perfectly on the star. For \(E(x, y)\), we used the YAO AO simulator\(^7\) to generate arrays of phase masks for three different conditions: i) observation with the UTs with AO locked on a \(Kmg = 7\) off-axis reference star; ii) observation with the UTs without AO; and iii) observation with the ATs\(^8\). The seeing conditions are average \((r_0 = 25 \text{ cm at } \lambda = 0.5 \text{ m})\). The results are presented in Table 3 for tip-tilt errors from 0 to 0.51/D.

The OPD errors average with time. However, the coherence time of the common path aberrations depends highly on the seeing conditions, wind speed, etc. A rough, conservative solution is to consider a coherence time of 100 ms, hence a reduction by a factor 10 for a 10 s observation. In the case of the 0.21/D stated in Sect. 6.2, it means a 0.9 nm OPD error on the UTs and 12.3 nm error on the ATs. A factor \(\sqrt{2}\) is needed to account for the atmospheric perturbations on each telescope. This conservative value is reported in Table 1.

It is important to note that the fiber positioning (responsible for the tip-tilt error) is crucial to an operation without AO. On the ATs, if the alignment is below 0.51/D, the phase measurement will wrap over \(2\pi\), preventing averaging of the error, and thus any realizable astrometric

7 http://frigaut.github.io/yao/index.html
8 As yet, the ATs are not equipped with an AO system.

7. Discussion and conclusion

7.1. Fifteen years of experimentation

The predominantly theoretical exercise of this paper is supported by 15 years of experimentation carried out on various interferometers. The first astrometric observations with a long baseline interferometer were achieved around 1999 on the Palomar...
Testbed Interferometer (Colavita et al. 1999), based on the theoretical work of Shao & Colavita (1992). Using small siderostats on 100-m-long baselines, the instrument reached 160 μas precision measurements on bright pairs (Shao et al. 1999). PTI was a technology demonstrator in preparation for the four 1.8 m telescopes planned to be added to the Keck Interferometer. The Keck Outrigger astrometry project was the first to tackle the issue of implementing astrometric baselines on large telescopes. The concept foresees transfer of the metrology end point to primary space by a dedicated baseline monitor (Hrynevych et al. 2004). However, the project was cancelled in 2006 before seeing first light. Started in 2008, and even though it never managed an on-sky astrometric demonstration, the Astra project of astrometric extension of the two ten-meter Keck Interferometer was the occasion to study the implementation of metrologies further for astrometric interferometer on large telescopes.

The design of PRIMA (Delplancke 2008), the VLTI equivalent to Astra, and the Keck Outriggers started in 2000. In 2011, during its first on-sky observations, PRIMA achieved an astrometric precision of 30 μas, but the astrometric accuracy was limited to about 3 mas, mainly because the metrology end points were located at the telescopes Coude foci, far from primary space (Sahlmann et al. 2013). A second on-sky demonstration run in 2013, with the metrology extended to the telescope secondary mirror, resulted in an astrometric accuracy of about 100 μas.

The GRAVITY metrology (Gillessen et al. 2012) follows a slightly different concept than the previous experiments. Instead of measuring the optical path difference in double pass between two telescopes for each star, as done for PRIMA and Astra, the GRAVITY metrology measures the optical path difference between the two beams for each telescope in single pass. And other than in PRIMA and Astra, the GRAVITY metrology end points are mounted above the primary mirror on the telescope spider, therefore covering the full optical path and providing a physical realization of the narrow-angle baseline endpoints. A second laser is sent back from the telescope to actively track the pupil on the metrology endpoints (Amorim et al. 2012).

7.2. On the GRAVITY astrometric error budget

A sophisticated error budget was a central design driver for GRAVITY. The idea is to cancel as many error terms as possible thanks to technical choices. For example, the metrology is sent up to the spider arms, above M1. It is located in geographical coordinates to minimize error when transposing the physical narrow-angle baseline to the sidereal reference frame (Sect. 6.1). The metrology is also sent by the same single-mode fibers that accept the star light. The difference in propagation direction between the metrology and the starlight is therefore minimized (Sect. 6.2).

The necessity of active control over the pupils (the imaging baseline limit points) is one of the main requirements given by the error budget. The imaging baselines introduce an error term if the instrument pupil and focal plane are not correctly stabilized to a given position (Sect. 3.3). This cannot be canceled out by technical design, so GRAVITY includes a pupil and field tracking system inside its own cryostat. This system controls the position of the instrument pupil in real time with respect to four laser beacons on the spider arms of the telescopes. It also monitors the position of the stellar objects to keep them accurately positioned (within 0.2λ/D) on the fibers.

Field tracking is also fundamental to minimize the influence of common-path aberrations. At first sight, common path aberrations (e.g., atmospheric perturbations) do not matter. This is not the case if there is an additional error on the field tracking: it is the cross-term between field error and atmospheric perturbation that matters. In Sect. 6.4.2, we showed that atmospheric perturbations have the potential to be the major limit of the astrometric precision, depending on the level of star tracking. Moreover, we show that GRAVITY’s accuracy will depend on the level of AO correction. Notably, the absence of AO on the ATs does put a stringent requirement on the affordable tip-tilt errors for the system to work (see Table 3).

The main result of the error budget is that GRAVITY should fulfill its objective of 10 μas accuracy on the UTs if i) the telescope pupils are positioned within the instrument to an accuracy of 0.5% of their diameter (4 cm on M1) and ii) the two stellar objects are tracked within 0.2λ/D.

Acknowledgements. S.L. would like to thank J.D. Monnier for fruitful discussions about the impact of common-path aberrations.

References

Lawson, P. R. 2000, Principles of Long Baseline Stellar Interferometry (Pasadena: NASA, JPL, CIT)
Robieux, J. 1959, IEEE Transactions on Antennas and Propagation, 7, 118