Optical measurements of phase steps in segmented mirrors – fundamental precision limits

L. NOETHE† and H.-M. ADORF‡

† European Southern Observatory, Garching bei München, Germany
‡ Max-Planck-Institut für Astrophysik, Garching bei München, Germany

May 22, 2006

Abstract

Phase steps are an important type of wavefront aberrations generated by large telescopes with segmented mirrors. In a closed-loop correction cycle these phase steps have to be measured with the highest possible precision using natural reference stars, that is with a small number of photons. In this paper the classical Fisher information of statistics is used for calculating the Cramér-Rao minimum variance bound, which determines the limit to the precision with which the height of the steps can be estimated in an unbiased fashion with a given number of photons and a given measuring device. Four types of such measurement devices are discussed: a Shack-Hartmann sensor with one small cylindrical lenslet covering a subaperture centred over a border, a modified Mach-Zehnder interferometer, a Foucault test, and a curvature sensor. The Cramér-Rao bound is calculated for all sensors under ideal conditions, that is narrowband measurements including photon shot noise, but without other forms of noise or disturbances. This limit is compared with the ultimate quantum statistical limit for the estimate of such a step, which is independent of any measuring device. For one device, the Shack-Hartmann sensor, the effects on the Cramér-Rao bound of broadband measurements, finite sampling, and disturbances such as atmospheric seeing and detector readout noise are also investigated. The methods presented here can be used to compare the precision limits of various devices for measuring differential segment phases, and for optimising the devices. Under ideal conditions the Shack-Hartmann and the Foucault devices nearly attain the ultimate quantum statistical limit, whereas the Mach-Zehnder and the curvature devices each require approximately twenty times as many photons in order to reach the same precision.

1 Introduction

In future extremely large optical telescopes at least one of the mirrors will probably be segmented. During the operation of such a telescope the wavefront errors introduced by segmentation have to be reduced to the order of a few tens of nanometers. If the segments are regarded as rigid bodies, the wavefront errors can only be generated by tip-tilt errors of individual segments or differential axial displacements. This paper assumes that the segments are already perfectly aligned in tip and tilt. A correction of remaining piston errors – from now on called differential segment phases – requires measurements of the heights of the wavefront steps at the borders of adjacent segments with a precision of a few nanometers, and the suppression of the relative segment phases by appropriate rigid body movements of the segments.

Several methods are used or have been proposed for the measurement of differential segment phases in segmented mirrors. One of these is based on the Shack-Hartmann technique and applied in a slightly modified fashion at the Keck telescope [1][2][3]. In an image of the segmented mirror a circular lens is placed over a subaperture centred on an intersegment border. The phase step across the border is estimated from the form of the diffraction pattern, for narrowband and broadband measurements, and in the presence of atmospheric disturbances. Other methods proposed in the literature use spatial filtering in a modified Mach-Zehnder interferometer [4], measurements in defocussed images and pupils [5] [6] [7] [8], phase filtering techniques [9], or a knife

edge technique in the pyramid wavefront sensor [10][11][14][15]. The performance of four types of sensors will be evaluated and compared in the Active Phasing Experiment [12][13].

All the sensors discussed in this paper generate diffraction patterns in the plane of a detector. These patterns can be regarded as a likelihood function $p(x \mid \theta)$ describing the "likelihood" to register a photon on the detector as a function of the spatial variable x on the detector, given a certain value of a parameter, in this paper the relative segment phase θ. There are various possibilities to define, based on $p(x \mid \theta)$ and its derivatives with respect to θ, an information content in the likelihood function. One of them is the classical Fisher information [16]:

$$I_{\mathrm{F}} = \int_{-\infty}^{+\infty} \mathrm{d}x \; p(x \mid \theta) \left(\frac{\partial \ln p(x \mid \theta)}{\partial \theta} \right)^2 . \tag{1}$$

Estimating the differential segment phases is particularly difficult at low light levels, e.g. when only the light of faint stars is available for the measurements. The question arises: what precision can ultimately be obtained with a given number of detected photons? In statistics, precision is characterized by the conditional variance of an estimate of the parameter. For a parameter of interest, in this paper the relative segment phase θ, the minimum variance of an unbiased estimate of the parameter, the so-called Cramér-Rao minimum variance bound σ^2_{CRB} [16], is given by the inverse of the classical Fisher information $I_{\mathrm{F}}(\theta)$:

$$\sigma^2_{\mathrm{CRB}}(\theta) = 1/I_{\mathrm{F}}(\theta) \quad . \tag{2}$$

In the following the shorter expression Cramér-Rao bound will be used for the square root σ_{CRB} of the Cramér-Rao minimum variance bound.

Two of the assumptions in the derivation of equation (2) are, unless obvious, checked for all cases investigated in this paper. First, the integral of the likelihood function over the interval for which it is defined must not depend on θ, at least not for a particular value of θ for which the Fisher information is calculated. Second, the likelihood function must be regular with respect to θ. Under the first assumption the likelihood function can be normalised:

$$\int_{-\infty}^{+\infty} \mathrm{d}x \, p(x \mid \theta) = 1. \tag{3}$$

Assuming a 100% quantum efficiency of the detector, this is equivalent to the assumption that one photon passes through the aperture.

If $q(x \mid \theta)$ denotes a real-valued likelihood amplitude, defined by

$$q(x \mid \theta) = \sqrt{p(x \mid \theta)}, \tag{4}$$

the integrand in equation (1), which we will call the *Fisher information density* $i_{\mathrm{F}}(x \mid \theta)$, can be written as

$$i_{\mathrm{F}}(x \mid \theta) = 4 \left(\frac{\partial q(x \mid \theta)}{\partial \theta} \right)^2 . \tag{5}$$

In principle there is an ambiguity related to the definition of the Fisher information density. This is due the identity

$$\int_{-\infty}^{+\infty} \mathrm{d}x \left(\frac{\partial q(x \mid \theta)}{\partial \theta} \right)^2 = - \int_{-\infty}^{+\infty} \mathrm{d}x \, q(x \mid \theta) \, \frac{\partial^2 q(x \mid \theta)}{\partial \theta^2} \tag{6}$$

which can be obtained by differentiating equation (3) twice with respect to θ. The Fisher information defined as

$$I_{\mathrm{F}} = 4 \int_{-\infty}^{+\infty} \mathrm{d}x \left[A \left(\frac{\partial q(x \mid \theta)}{\partial \theta} \right)^2 + B \, q(x \mid \theta) \, \frac{\partial^2 q(x \mid \theta)}{\partial \theta^2} \right] \tag{7}$$

would be independent of the choice of A and B provided that $A + B = 1$. However, the term following the coefficient B can be negative, possibly giving rise to a negative Fisher information density. Since this seems

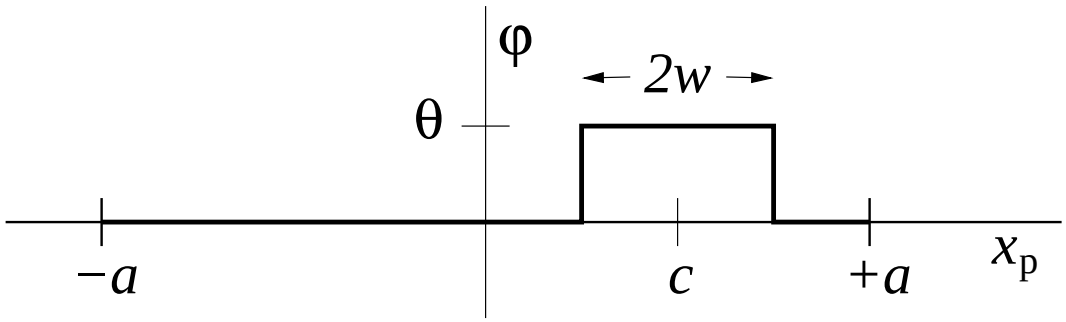

Figure 1: Wavefront phase φ with a relative segment phase θ generated by a segment of width $2w$ and a centre at c within a subaperture of width $2a$.

unreasonable, B is set to zero and the Fisher information density is defined as in equation (5). The Fisher information will therefore be computed as

$$I_{\mathrm{F}} = \int_{-\infty}^{+\infty} i_{\mathrm{F}}(x \,|\, \theta) = 4 \int_{-\infty}^{+\infty} \mathrm{d}x \, \left(\frac{\partial q(x \,|\, \theta)}{\partial \theta} \right)^2 . \tag{8}$$

According to equation (2) the computation of the Fisher information offers a possibility of calculating analytically the potential bounds to the precision of various measurement methods. These bounds can then be compared with the ultimate limit which is independent of any measuring device. The latter is calculated in section 3 from quantum statistical principles.

Section 4 discusses the Cramér-Rao bounds for the Shack-Hartmann, the Mach-Zehnder, the Foucault and the curvature sensor under ideal conditions, that is for monochromatic light with photon shot noise, but without any effects introducing additional noise or light losses or aberrations. The effects of broadband measurements, finite sampling, and external disturbances such as seeing and detector readout noise can also be included in the calculations. The methods are outlined in detail only for the Shack-Hartmann sensor in section 6, but could readily be applied to the other sensors as well.

In general, the limit given by the Cramér-Rao bound can only be attained, if an efficient estimator is available for the analysis of the data. The issue, whether such an efficient estimator exists, will, however, not be discussed in this paper. Furthermore, all calculations will be carried out for the one-dimensional case, which is sufficient for all devices discussed below.

2 Definition of the wavefront error

Figure 1 shows the wavefront after the reflection by the mirror. The one-dimensional spatial variable is denoted by x_{p} and the pupil defined by the interval $-a$ to $+a$. An out-of-phase segment, centred at c, generates a constant wavefront error or phase difference $\varphi(x_{\mathrm{p}}) = \theta$ over its full width $2w$ compared with an also constant phase over the rest of the pupil. There are two wavefront discontinuities, that is phase steps of height θ, at the segment borders. Apart from an irrelevant average phase, the complex amplitude describing the wavefront is given by

$$U_{\mathrm{s}}(x_{\mathrm{p}} \,|\, \theta) = \frac{1}{\sqrt{2a}} \, e^{ik\varphi(x_{\mathrm{p}})} \, W(x_{\mathrm{p}}, -a, +a) \quad , \tag{9}$$

with the phase function

$$\varphi(x_{\mathrm{p}}) = \left\{ \begin{array}{lll} 0 & \text{for} & -a \le x_{\mathrm{p}} \le c - w \\ \theta & \text{for} & c - w < x_{\mathrm{p}} < c + w \\ 0 & \text{for} & c + w \le x_{\mathrm{p}} \le a \end{array} \right\} \tag{10}$$

and the boxcar function

$$W(x_{\mathrm{p}}, -a, +a) = \left\{ \begin{array}{lll} 0 & \text{for} & x_{\mathrm{p}} < -a \\ 1 & \text{for} & -a \le x_{\mathrm{p}} \le +a \\ 0 & \text{for} & +a < x_{\mathrm{p}} \end{array} \right\} \quad . \tag{11}$$

Here $k = 2\pi/\lambda$ is the wavenumber and λ is the wavelength of the light. The restriction of the definition to the interval $[-a, +a]$ signifies that only the photons passing through the aperture of width $2a$ can be used for estimating the relative segment phase θ, and the factor in front of the exponential in equation (9) normalizes the wave function such that the number of photons within this subaperture is equal to 1, that is

$$\int_{-a}^{+a} dx_{\mathrm{p}}\, U_{\mathrm{s}}^{*}(x_{\mathrm{p}} \,|\, \theta)\, U_{\mathrm{s}}(x_{\mathrm{p}} \,|\, \theta) = 1 \quad , \tag{12}$$

where U_{s}^{*} denotes the complex conjugate of U_{s}. In addition to the general *two-border* configuration shown in figure 1, two special configurations will be discussed in more detail. In the first configuration the segment is located at the edge of the aperture $[-a, +a]$, that is if $c = a - w$. This configuration will subsequently be called the *single-border* configuration. In the second special configuration, which is a special case of the single-border configuration, the border is located at the centre of the aperture $[-a, +a]$, that is at $c = w = a/2$. This configuration, which is of particular interest for the Shack-Hartmann type sensor, will be called the *centred single-border* configuration. For the general configuration the parameter θ will be referred to as the relative segment phase, and for the single-border configurations as the phase step.

3 Quantum statistical precision limit for the estimate of a relative segment phase

The ultimate precision limit for the measurement of differential segment phases can, if at all, be obtained with a measuring device which is capable of transferring to the data the maximum relevant information contained in the wavefront immediately after the reflection by the mirror. In addition, one needs an appropriate, that is efficient data analysis procedure.

As stated in the introduction, the limit on the data side can be obtained from the Fisher information which is derived from a classical, real-valued likelihood function. However, immediately after the reflection of an incoming monochromatic plane wavefront by the mirror configuration shown in figure 1, the full information is contained in a complex quantum statistical wave function $\psi(x \,|\, \theta)$, which is equivalent to the complex amplitude function (9) familiar from optical theory. Therefore, the calculation of the information in the wavefront requires the quantum statistical equivalent of the expression (8) for the Fisher information. For a pure state ψ and its complex conjugate ψ^{*} this is given [17] by

$$I_{\mathrm{QM}}(\theta) = 4 \left(\int_{-a}^{+a} \mathrm{d}x_{\mathrm{p}} \frac{\partial}{\partial\theta}\psi^{*}(x_{\mathrm{p}} \,|\, \theta) \frac{\partial}{\partial\theta}\psi(x_{\mathrm{p}} \,|\, \theta) \;-\; \left| \int_{-a}^{+a} \mathrm{d}x_{\mathrm{p}}\, \psi^{*}(x_{\mathrm{p}} \,|\, \theta) \frac{\partial}{\partial\theta}\psi(x_{\mathrm{p}} \,|\, \theta) \right|^{2} \right) \quad . \tag{13}$$

The first term on the right hand side is similar to the definition of the classical Fisher information density in equation (8). The second term, in effect, suppresses a dependence of I_{QM} on an undetectable average phase across the full aperture $[-a, +a]$.

By replacing the complex wave function ψ in equation (13) by the complex amplitude U_{s} of equation (9) one obtains the quantum statistical equivalent of the Fisher information about the relative segment phase contained in the complex amplitude immediately after the reflection by the mirrors:

$$I_{\mathrm{QM}}(\theta) = 4\,k^{2}\,\frac{w}{a}\left(1 - \frac{w}{a}\right) \quad . \tag{14}$$

The quadratic dependence of the quantum statistical information about the segment width, which is independent of the segment position, is shown in figure 13 by the solid line. The maximum

$$I_{\mathrm{QM,max}}(\theta) = k^{2} \tag{15}$$

is obtained, if the segment covers half of the aperture, that is for $w = a/2$.

For a single photon, the lower bound for the root-mean-square (RMS) of the uncertainty in the estimate of the parameter θ, that is the quantum statistical equivalent of the Cramér-Rao bound, is given by

$$\sigma_{\mathrm{QM,CRB,1}}(\theta) = \frac{1}{2k}\frac{a}{\sqrt{w(a-w)}} \quad . \tag{16}$$

The minimum value is obtained, if the segment covers half of the pupil:

$$\sigma_{\mathrm{QM,CRB},1,\mathrm{min}}(\theta) = \frac{1}{k} = \frac{\lambda}{2\pi} \quad . \tag{17}$$

It depends only on the wavelength of the light, and is of the order of 100 nm for visible wavelengths. For N_{phot} registered photons the information $I_{\mathrm{QM}}(\theta)$ increases by a factor of N_{phot}, the quantum statistical Cramér-Rao bound decreases by $1/\sqrt{N_{\mathrm{phot}}}$, and the lower bound for the RMS of the uncertainty in the estimate of the parameter θ is correspondingly given by

$$\sigma_{\mathrm{QM,CRB,min}}(\theta) = \frac{1}{\sqrt{N_{\mathrm{phot}}}} \frac{\lambda}{2\pi} \quad . \tag{18}$$

It is important that this fundamental limit does not refer to any particular measurement device or a particular data analysis procedure. It is based on the maximum information that can potentially be extracted from the wavefront, without necessarily implying that such an efficient measurement process exists.

4 Cramér-Rao bounds for the measurement of piston steps with different sensors under ideal conditions

The Cramér-Rao bound for measuring the relative segment phase or phase step θ is calculated for each of the four sensors mentioned above, under ideal conditions, that is under the following assumptions: light from a quasi-monochromatic source, a sufficiently fine sampling, and, apart from the photon shot noise, no additional noise introduced by atmospheric effects or the readout characteristics of the detector.

The phase corrections in a segmented mirror are performed in two steps. First, after the installation of new segments, initial large differential segment phases of the order of possibly several wavelengths have to be reduced to fractions of a wavelength. Usually, this only needs to be performed once and can therefore be done with bright stars. Since these supply a large number of detectable photons, high precision can be obtained during this process. Second, during observations, one has to be content with small number of photons, but has to correct only comparatively small differential segment phases. It is therefore sufficient to calculate the precision limits or Cramér-Rao bounds only for the case of very small differential segment phases or phase steps, that is for $\theta \to 0$.

4.1 Shack-Hartmann method

Usually, the Shack-Hartmann method used for phasing measurements consists of placing a circular lens across the border between two segments. The information on the phase step is contained in the diffraction pattern of the lens. In a one-dimensional setting, a cylindrical lens is used with its axis parallel to the border, and its centre ideally exactly placed on the border. The normal configurations for the Shack-Hartmann sensor are therefore the single-border and the centred single-border configuration, as defined in section 2. However, to be able to compare the Shack-Hartmann sensor with the other phasing wavefront sensors, the computation of the likelihood function in the focal plane and the Fisher information, presented in appendix A1, are also carried out for the general configuration shown in figure 1, that is for a single lens covering the full aperture with one out-of-phase segment in an arbitrary location. The likelihood function $p(\xi\,|\,\theta)$ and the Fisher information density $i_{\mathrm{F}}(\xi\,|\,\theta)$ are both functions of the phase step θ and a dimensionless position variable ξ. The latter is related to the true position variable x in the focal or detector plane by $\xi = x/f$, where f is the focal length of the lens. The general as well as the special configurations will be discussed in this section.

Section 4.1.1 discusses the likelihood function, the Fisher information density, and the Fisher information. A variation of the parameter θ causes a shift of the location of the maximum of the diffraction pattern as well as a change of its form. From both characteristics partial information about θ can be obtained. Section 4.1.2 shows that three quarters of the information are contained in the shift of the location of the maximum, and section 4.1.3 that one quarter of the full information is contained in the change of the form.

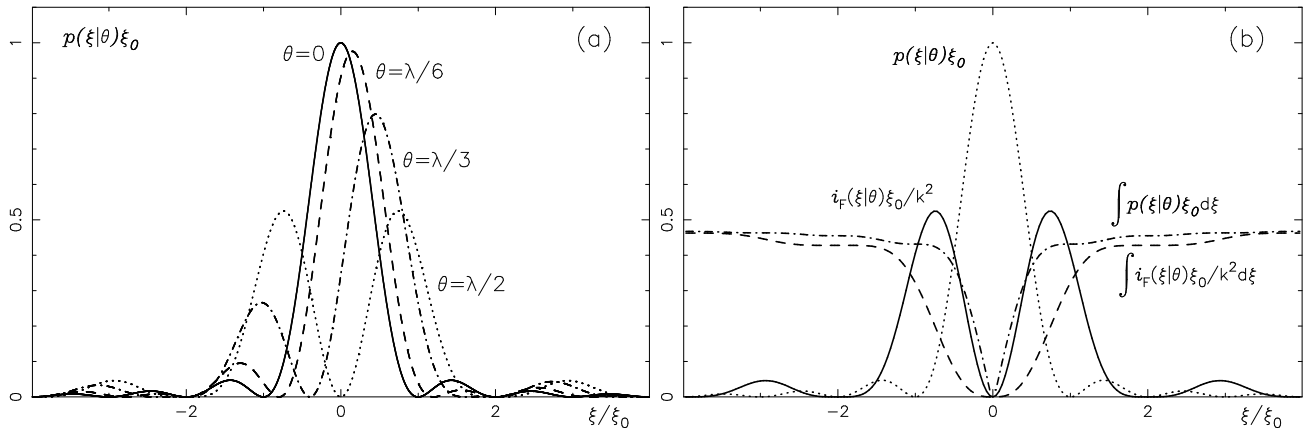

Figure 2: **Shack-Hartmann sensor:** (a) Likelihood functions for four differential segment phases θ as a function of a the normalised coordinate in the focal plane of the Shack-Hartmann lens. (b) Likelihood function and corresponding functions related to the Fisher information for $\theta = 0$: normalised likelihood function $p(\xi \,|\, \theta)\xi_0$ (dotted), normalised Fisher information density $i_{\mathrm{F}}(\xi \,|\, \theta)\xi_0/k^2$ (solid), integrated normalised likelihood function (dashed-dotted), and integrated normalised Fisher information density (dashed).

4.1.1 Likelihood function, Fisher information density and Fisher information

For the centred single-border configuration figure 2a shows the likelihood functions $p(\xi \,|\, \theta)$, for four different phase steps θ ranging from zero to $\lambda/2$. The location of the first zero in the diffraction pattern for $\theta = 0$ will be denoted by $\xi_0 = \lambda/(2a)$. For the centred single-border configuration and for arbitrary θ the Fisher information density and the likelihood function are related by

$$i_{\mathrm{F}}(\xi \,|\, \theta) = k^2 \, p(\xi \,|\, \theta - \lambda/2) \quad . \tag{19}$$

Figure 2b shows that for the case $\theta = 0$ the first maxima of the Fisher information density on both sides of the origin are lying just inside the first minima of the likelihood function, represented by the dotted curve. The dashed line in figure 2b shows the Fisher information density integrated from $\xi = 0$ to the ξ value on the abszissa and divided by k^2, and the dashed-dotted line the corresponding integrated likelihood function. The likelihood function is more concentrated around the origin than the corresponding Fisher information density. For the case $\theta = \lambda/2$ the likelihood function has a minimum at $\xi = 0$. Consequently, the Fisher information density has its maximum at $\xi = 0$ and is therefore more concentrated around the origin than the corresponding likelihood function.

The Fisher information is obtained by integrating the Fisher information density from $\xi = -\infty$ to $\xi = +\infty$. For the general configuration in the limiting case $\theta \to 0$ the integration of the Fisher information density given in equation (A-5) can be done analytically and yields

$$I_{\mathrm{F,SH}} = k^2 \, \frac{2}{a} \min(w, |c|) \quad . \tag{20}$$

The maximum of $\min(w, |c|)$ equals $a/2$, confirming that the Fisher information in equation (20) is always bounded by the quantum statistical information given in equation (15). For the special case of the single-border configuration this can be seen in figure 13a, which demonstrates the linear dependence of the Fisher information for the Shack-Hartmann sensor on the segment width $2w$.

For the general configuration and for a given $w = 0.1a$, figure 13b shows the Fisher information for the Shack-Hartmann sensor as a function of the location of the segment centre. Surprisingly the Fisher information is zero for a segment in the centre of the aperture.

The Fisher information for the Shack-Hartmann sensor attains its maximum for the centred single-border configuration, that is for $w = |c| = a/2$:

$$I_{\mathrm{F,SH,max}} = k^2 \quad , \tag{21}$$

and is identical to the quantum statistical information. Therefore, if all the relevant information about the relative segment phase in the diffraction pattern could be extracted with an efficient estimator, one could reach the highest possible precision, that is the fundamental bound (15) for the estimate of the phase step θ.

The formulae given above have been derived for one lens covering the full aperture $[-a, +a]$. However, in a phasing wavefront sensor, a single lenslet will only cover a subaperture across a border. For these lenslets the configuration will ideally be the centred single-border configuration. It can readily be verified that, due to the linear dependence of the Fisher information on the segment width, the information I_{F} obtained with a single lens across the full aperture is identical to the information obtained with lenslets across the borders and centred on the borders, provided that the lenslets cover the full segment width and the segment centre is not too close to the centre of the aperture. Since the other sensors work over the full aperture, the Fisher information for the Shack-Hartmann method can therefore be compared with the other methods by using the results for a single lens covering the full aperture.

4.1.2 Information contained in the shift of the location of the maximum of the likelihood function

Figure 2b, applying to the centred single-border configuration, shows that in the most interesting case of $\theta \ll \lambda$ nearly 50% of the Fisher information is contained in the interval between the first zeroes of the likelihood function on both sides of the origin. The maximum of the likelihood function is well defined, even in case of broadband measurements or in the presence of atmospheric disturbances as shown in sections 6.1 and 6.2. Furthermore, the location of the maximum depends on the phase step θ, and is not too sensitive to small misalignments between the centre of the lens and the border between segments. For small phase steps θ and small deviations from the centred single-border configuration, described by the difference $c - a/2$, that is for $k\theta \ll 1$ and $k(c - a/2) \ll 1$, one obtains for the location $\xi_{\max}(\theta)$ of the maximum of the likelihood function

$$\xi_{\max}(\theta) = -\frac{3}{4}\,\frac{\theta}{a}\,(1 - 4\frac{(c - a/2)^2}{a^2}) \quad . \tag{22}$$

This equation shows that the location of the maximum is only weakly affected by the misalignment $c - a/2$ provided that $c - a/2 \ll a$. A similar expression for $\xi_{\max}(\theta)$ without the effect of the decentring has been given in [2] for circular lenslets.

Therefore, it is of interest to estimate the phase step θ only from the core of the likelihood function which is defined here as the interval within which the likelihood function exceeds 20% of its maximum value. Figure 2b shows that for the centred single-border configuration the integral of the Fisher information density over this interval is approximately $0.3 I_{\mathrm{F,SH,max}}$, that is approximately 30% of the total information can be extracted from the core.

The central core can quite accurately be fitted by a Gaussian

$$p_{\mathrm{gauss}}(\xi \mid \theta) = \frac{0.6ka}{\sqrt{\pi}}\, e^{-[0.6ka(\xi - 3\theta/(4a))]^2} \quad , \tag{23}$$

and the Fisher information contained therein is given by

$$I_{\mathrm{F,gauss}} = 0.4\, k^2 \approx 0.4\, I_{\mathrm{F,SH,max}} \quad . \tag{24}$$

This is only marginally larger than the information content $0.3 I_{\mathrm{F,SH,max}}$ in the core. The reason for the small difference is probably that the integration of the Gaussian probability function extends from $-\infty$ to $+\infty$, whereas the one of the diffraction pattern only over the central core.

The shift $\Delta\xi_{\mathrm{mean}}$ of the mean position, that is the centre of gravity of the likelihood function, is not identical to the shift of the central maximum of the diffraction pattern. $\Delta\xi_{\mathrm{mean}}$ should be smaller than $\Delta\xi_{\max}$ since the shift of $\Delta\xi_{\max}$ to one side is accompanied by an increase of the first peak on the opposite side. For the centred single lens configuration one obtains from equation (A-4) for the centre of gravity

$$\Delta\xi_{\mathrm{mean}} = -\frac{1}{2ka}\,\sin(k\theta) \quad . \tag{25}$$

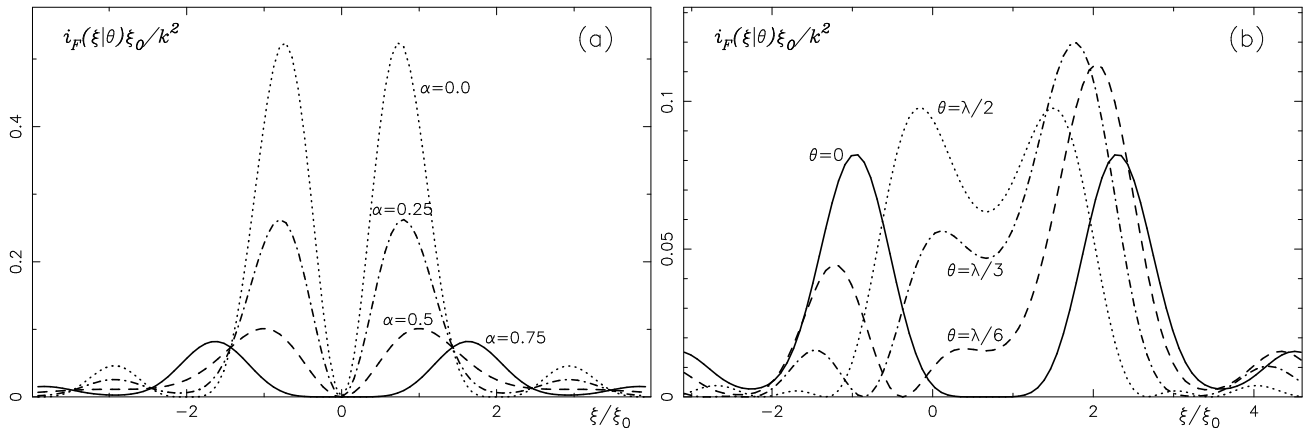

Figure 3: **Shack-Hartmann sensor:** (a) Normalised Fisher information density $i_{\mathrm{F}}(\xi \mid \theta)\xi_0/k^2$ compensated by various shifts of $\alpha\theta/a$ for the limit $\theta \to 0$. The values for α are given in the figure. (b) Normalised Fisher information density in the form of the likelihood function for four phase steps θ.

For $k\theta \ll 1$ this becomes

$$\Delta\xi_{\mathrm{mean}} = -\frac{\theta}{2a} \quad . \tag{26}$$

A comparison with equation (22) shows that this is indeed smaller than the shift $\Delta\xi_{\mathrm{max}}$ of the maximum by a factor of 1.5. It is equal to half of the average tilt of the wavefront $\varphi(x_{\mathrm{p}})$ over the aperture in the pupil.

4.1.3 Information contained in the form of the likelihood function

In some applications a reference pattern generated by the Shack-Hartmann lenslet array may not be available. Furthermore, an adaptive optics system may partially correct a phase step, in the simplest case by introducing a constant tilt over a subaperture covered by a lenslet. Such a tilt, the amount of which may not be known, will lead to a shift of the diffraction pattern. Under such circumstances no information can be obtained from the position of the diffraction pattern, and the parameter θ can only be estimated from the changes in the form of the likelihood function.

The Fisher information contained in the form, from now on called the "form-only Fisher information", can be calculated as follows. The likelihood function for a given phase step θ can be shifted such that its maximum is closer to the origin. The Fisher information can then be calculated as a function of this shift, with the information contained in the form being defined as the minimum of the Fisher information as a function of the shift. For the centred single-border configuration a likelihood function p_{s} shifted by $\alpha\theta/a$ is described by

$$p_{\mathrm{s}}(\xi \mid \theta) = \frac{1}{\xi_0} \frac{1}{(\pi\xi/\xi_0 - \alpha 2\pi\theta/\lambda)^2} \frac{1}{(\xi - \alpha\theta/a)^2} \left[\sin(\pi\xi/\xi_0 - \alpha 2\pi\theta/\lambda + 2\pi\theta/\lambda) - \sin(2\pi\theta/\lambda)\right]^2 \quad . \tag{27}$$

Figure 3a shows for a few values of α the Fisher information densities for $\theta = 0$. For arbitrary values of θ the minimum Fisher information is analytically obtained for $\alpha = 3/4$. According to equation (22), for $\theta \ll \lambda$ such a shift is equivalent to moving the maximum of the likelihood function back to $\xi = 0$. Figure 3b shows the form-only Fisher information densities for four phase steps ranging from $\theta = 0$ to $\theta = \lambda/2$. A comparison with figure 2b shows that for $\theta = 0$ the peaks of the Fisher information densities are much lower than in case of the full information and are located further away from the origin.

For all values of θ the minimum Fisher information in the centred single-border configuration is given by

$$I_{\mathrm{F,SH,form}} = k^2/4 \quad . \tag{28}$$

Therefore, if an estimate of the phase step θ can only be based on the form of the likelihood function, one needs four times the number of photons to reach the same precision as is obtainable from the full information, which includes the information from the position of the pattern.

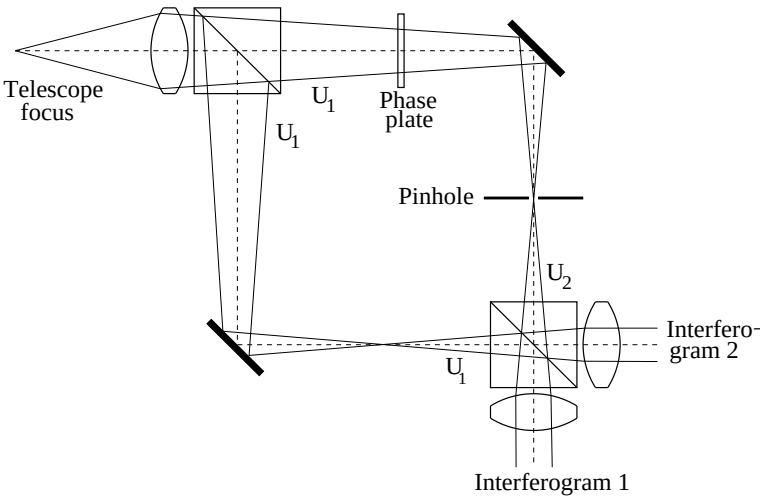

Figure 4: **Mach-Zehnder sensor:** Principle of a Mach-Zehnder interferometer when used for the detection of differential segment phases.

4.2 Mach-Zehnder interferometer

The principle of the Mach-Zehnder interferometer as used for the detection of differential segment phases [4] is shown in figure 4. The incoming light is focussed and split into two beams. One of them, which serves as the reference beam, contains a spatial filter in the focal plane and a phase plate introducing a phase shift ϕ. The light from the two beams is then recombined behind a second beam splitter, collimated and registered on two detectors in planes which are conjugated to the entrance pupil. The spatial filter is either a pinhole with a sharp edge or with a Gaussian transmission function with 100% transmission at the centre. If ξ is an dimensionless coordinate in the focal plane, the transmission functions of the filters are given by

$$t(\xi) \quad = \quad \left\{ \begin{array}{ll} \left\{ \begin{array}{lll} 1 & \text{for} & |\xi| \le \zeta \\ 0 & \text{for} & |\xi| > \zeta \end{array} \right\} & \text{sharp} - \text{edge} \\ \exp(-\xi^2/(2\zeta^2)) & \text{Gaussian} \end{array} \right\} \quad , \tag{29}$$

where ζ, also in dimensionless coordinates, is either half of the diameter of the sharp pinhole or the RMS of the Gaussian transmission function. The computation of the likelihood function and of the Fisher information density is outlined in appendix A2. Figure 5a shows the likelihood functions for sharp-edge pinholes with four different half-diameters ζ and $\theta = \lambda/10$, expressed as fractions or multiples of $\xi_0 = \lambda/(2a)$. Figure 5b presents the corresponding Fisher information densities. Similar curves for a Gaussian pinhole are shown in figures 6a and 6b. Obviously, the Fisher information is maximised for some intermediate value of ζ. On the one hand, for very small pinhole widths the form of the likelihood function closely resembles the form of the phase in the wavefront after the reflection by the mirrors. However, since only a small fraction of the light entering the second beam interferes with the light from the reference beam, the Fisher information is small. In the limit $\zeta \to 0$ one has $I_{\mathrm{F,MZ}} \propto \zeta^2$. On the other hand, for large pinholes the form in the interference pattern and the Fisher information density are concentrated at the locations of the borders between the segments. In the limit of very large pinhole sizes, the two beams are, apart from a constant phase shift, identical, and the information content goes to zero as $1/\zeta$ for $\zeta \to \infty$. The maximum Fisher information must therefore be obtained for an intermediate size of the pinhole. For a segment half-width of $w = 0.2a$ this maximum is obtained for $\zeta \approx 0.8\xi_0$ in case of a sharp pinhole and for $\zeta \approx 0.4\xi_0$ in case of a Gaussian pinhole, indicated by the dashed-dotted lines in the figures 5 and 6, respectively.

For the centred single-border configuration the maxima can be obtained from figure 7a which shows the normalised sum I_{F}/k^2 of the Fisher information in the two beams as a function of the ratio of ζ to ξ_0. In this case the maximum of the Fisher information is obtained for $\zeta \approx 0.465\xi_0$ in case of a sharp-edge pinhole and $\zeta \approx 0.235\xi_0$ in case of Gaussian pinhole. Figure 7b shows that these sizes of the pinholes maximising the Fisher information are, for both types of pinholes, comparable to the size of the core of the image in the focal

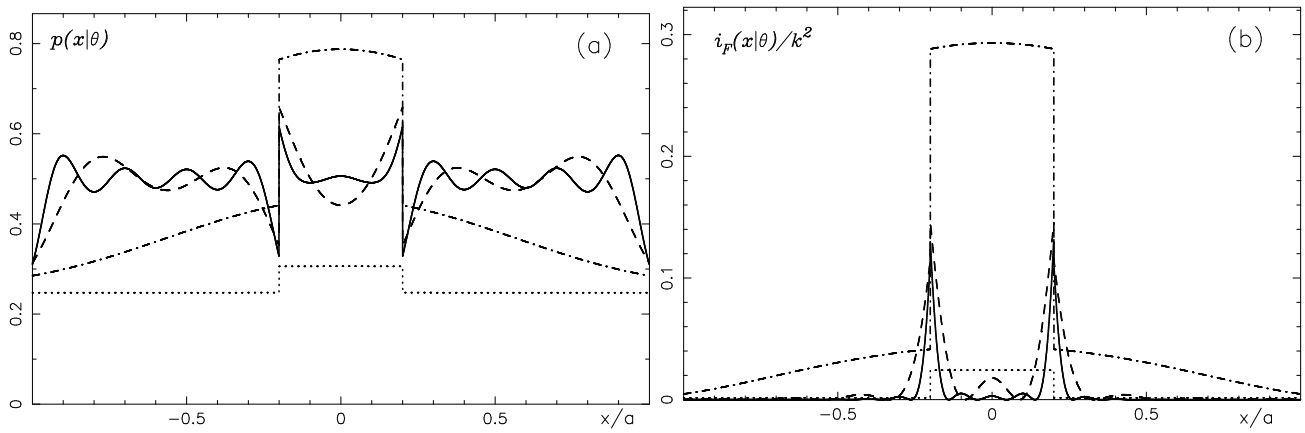

Figure 5: **Mach-Zehnder sensor:** (a) Likelihood functions generated by a Mach-Zehnder interferometer with sharp edge pinholes with radii of 0.1 (dotted), 0.8 (dashed-dotted), 4 (dashed) and 10 times $\lambda/(2a)$ (solid), for a relative segment phase $\theta = \lambda/10$. (b) Corresponding normalised Fisher information densities i_{F}/k^2 for $\theta = 0$.

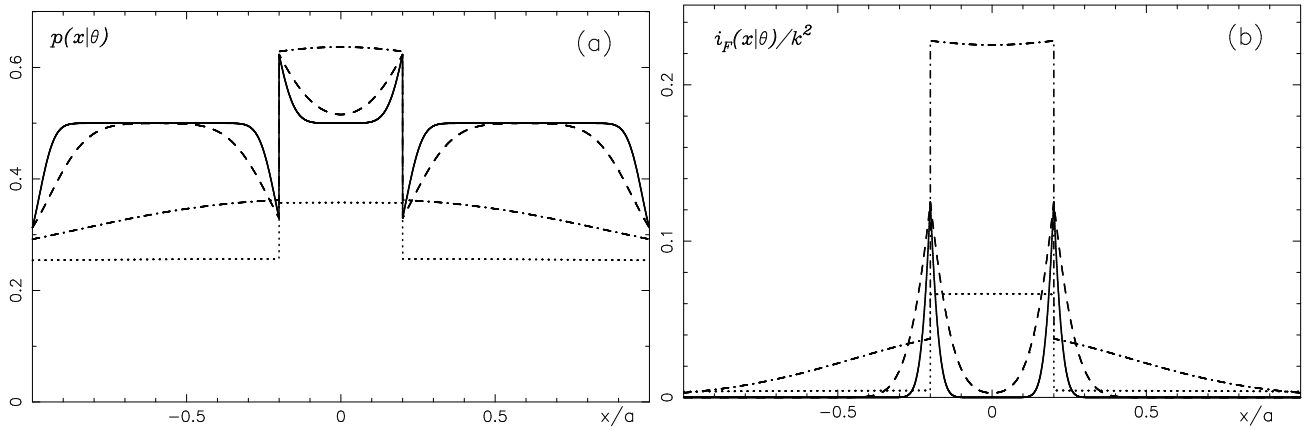

Figure 6: **Mach-Zehnder sensor:** (a) Likelihood functions generated by a Mach-Zehnder interferometer with Gaussian pinholes with RMS values of 0.1 (dotted), 0.4 (dashed-dotted), 2 (dashed) and 5 times $\lambda/(2a)$ (solid), for a relative segment phase $\theta = \lambda/10$. (b) Corresponding normalized Fisher information densities i_{F}/k^2 for $\theta = 0$.

plane. However, even under ideal conditions, the maximum Fisher information is only approximately 12% of the quantum statistical Fisher information in case of the sharp-edge pinhole, and 9% in case of the Gaussian pinhole.

For the single-border configuration and a width of the pinhole which maximizes the Fisher information the dotted line in figure 13a shows that $I_{\mathrm{F,MZ}}(\theta)$ depends quadratically on the segment width, and figure 13b shows that for the general configuration the Fisher information is slighly larger for segments closer to the centre of the aperture. However, if the Mach-Zehnder method is used in a telescope with several segments, the signals from neighbouring borders have to be well separated. Somewhat arbitrarily it will be defined that this condition is fulfilled, if the distances between the border and, in case of the sharp-edge pinhole, the first zero, or, in case of the Gaussian pinhole, the x-coordinate where the signal has dropped to 10% of the value at the border, are both approximately equal to $w/3$, as shown by the solid lines in the figures 5 and 6. This requires widths of $\zeta = \lambda/w$ for the sharp-edge pinhole and $\zeta = \lambda/(2w)$ for the Gaussian pinhole. Figure 8 shows that with these choices of the pinhole widths the Fisher information for the single-border configuration is now proportional to the segment half-width w. However, for such widths of the pinholes the Fisher information is much lower than for the widths maximising the Fisher information.

For well separated signals the Fisher information densities in the neighbourhood of a border are independent of the segment position. As long as the signals from both borders in a two-border configuration are well within

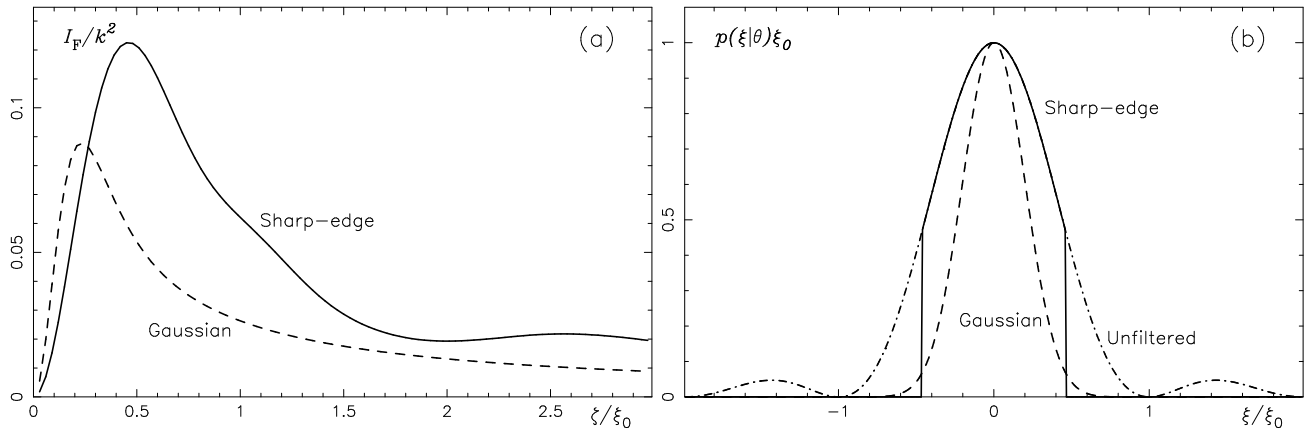

Figure 7: **Mach-Zehnder sensor:** (a) Normalised Fisher information $I_{\mathrm{F,MZ}}/k^2$ as a function of the normalised pinhole radius ζ/ξ_0 for pinholes with a sharp edge (solid) and with a Gaussian transmission function (*dashed*) for the centred single-border configuration. (b) Intensities filtered with a sharp-edge pinhole (solid), and with a Gaussian pinhole (dashed) with widths which maximise the Fisher information. The dashed-dotted curve represents the unfiltered diffraction pattern for a phase step $\theta = 0$.

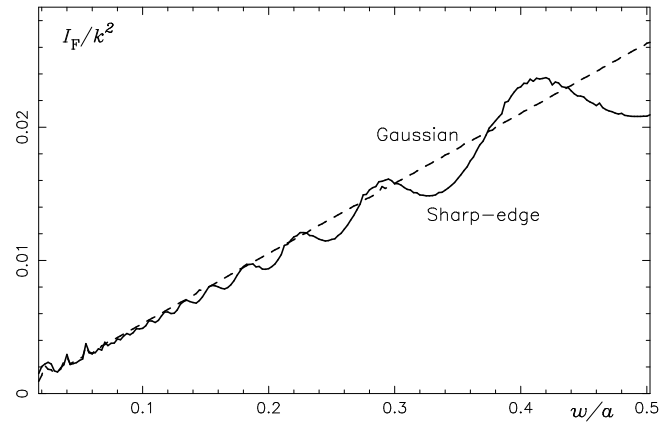

Figure 8: **Mach-Zehnder sensor:** Normalised Fisher information $I_{\mathrm{F,MZ}}/k^2$ as a function of half of the segment half-width w in a single-border configuration with pinhole sizes of λ/w for the sharp edge pinhole (solid), and $\lambda/(2w)$ for the Gaussian pinhole (dashed).

the aperture, the Fisher information is therefore also independent of the segment location.

4.3 Foucault knife-edge test

Differential segment phases can also be measured by a Foucault test. The incoming beam is focussed, a knife-edge is placed in the centre of the focal plane, and finally the beam is collimated and detected in a plane conjugated to the pupil. By replacing the knife-edge with a glass prism, both halves of the beam can be used for the analysis [10][11][14]. The computation of the likelihood function and the Fisher information density is outlined in appendix A3. Figure 9a shows the likelihood functions for three locations of the segment centre with a segment half-width of $w = 0.2a$ and for $\theta/\lambda = 0.1$, and figure 9b the Fisher information densities for $\theta = 0$. The widths of the peaks in the Fisher information densities in the neighbourhood of the borders are approximately equal to the widths of the peaks in the likelihood functions. The ratio of the width of the Fisher information density in the neighbourhood of a border to the width $2w$ of the segment is independent of the other parameters, and approximately equal to 0.2. The information from adjacent borders can therefore always be well separated. A numerical integration shows that for the centred single-border configuration the Fisher information in one of the

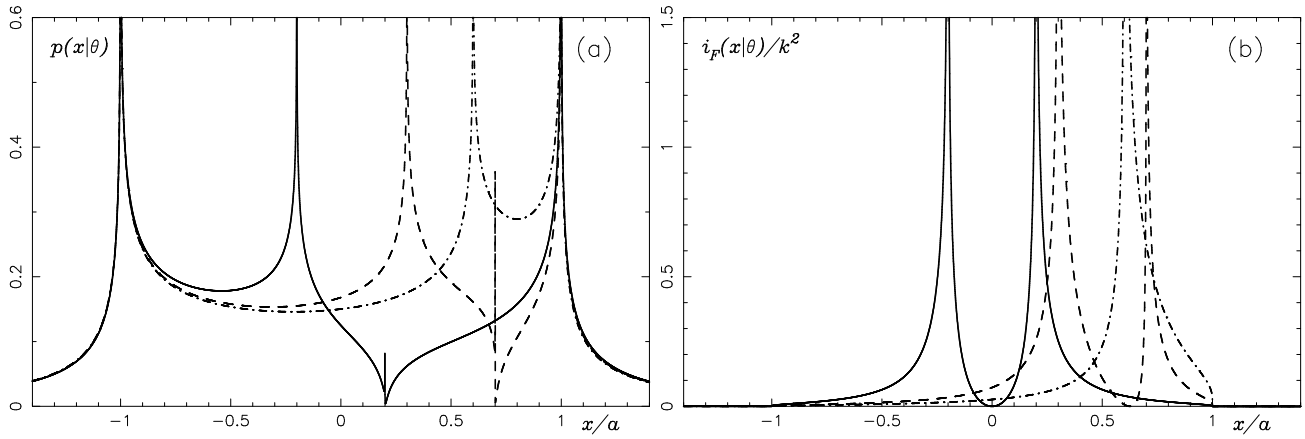

Figure 9: **Foucault sensor:** (a) Likelihood function for a relative segment phase $\theta = \lambda/10$. (b) Normalised Fisher information density for a relative segment phase $\theta = 0$ and $w = 0.2a$, and, for both plots, three different positions $c = 0$ (solid), $c = 0.5a$ (dashed) and $c = 0.8a$ (dashed-dotted) of the segment centre.

exit pupil equals $k^2/4$. This may intuitively be expected, since only half of the light enters one of the two exit pupils and only half of the image in the focal plane is used to retrieve information. Adding up the information in both channels gives $I_{\mathrm{F,FC}} = k^2/2$. The Foucault sensor with a glass prism therefore extracts potentially half of the quantum statistical information.

The dashed-dotted line in figure 13a shows the Fisher information in the single-border configuration as a function of the segment width. The dependence is quadratic, similar to case of the quantum statistical information. For all segment widths the Fisher information of the Foucault sensor equals half of the quantum statistical information. Figure 13b shows that the Fisher information for the general configuration is effectively independent of the segment location. Figure 9b shows therefore that the Fisher information density is larger in the neighbourhood of the single-border in the single-border configuration than in the neighbourhood of the two borders in the general configuration.

4.4 Curvature sensor

In a curvature sensor the detector is placed at a distance l in front or behind the focal plane, as shown in figure 10. The diffraction pattern is spread over an interval of approximately $2af/l$, and the intensities are therefore approximately proportional to l/f. It follows that the detector coordinates in the figures are normalised by a multiplication with $f/(la)$, and the likelihood function and the Fisher information density are normalised by a multiplication with l/f. Except for detector positions very close to the aperture, the diffraction pattern can be calculated with the Fresnel approximation, which is outlined in appendix A4. In the limit $l \to 0$ the expressions for the likelihood function and for the Fisher information density converge to the respective expressions for the Shack-Hartmann sensor derived in appendix A1. However, the curvature method requires the signals from the two segment borders to be clearly separated. In other words, the Fisher information densities corresponding to the two borders should not overlap. For $w = 0.2a$ such a separation in the likelihood function as well as in the Fisher information density seems to be established for $l/f \geq 0.002$, as shown in the figures 11a and 11b. Figure 12 shows in a log-log-plot, using the same parameters as in figure 11, the Fisher information as a function of the ratio l/f. The regime where $I_{\mathrm{F,CV}}$ is approximately constant is the Fraunhofer regime. Let the pure curvature regime be defined as the regime where the signals from two adjacent borders are well separated. The ratio $l/f = 0.002$, marking the lower limit for which a separation is guaranteed, lies at the beginning of the regime with a constant slope in the log-log-plot, where $I_{\mathrm{F,CV}} \propto \sqrt{f/l}$ and which extends at least up to $l/f \approx 0.1$. With a choice of $l/f > 0.002$ one is well outside the caustic, that is, the correlation of the subapertures is identical with that in the pupil.

A precise definition of the signal width which guarantees a separation of signals from adjacent borders is

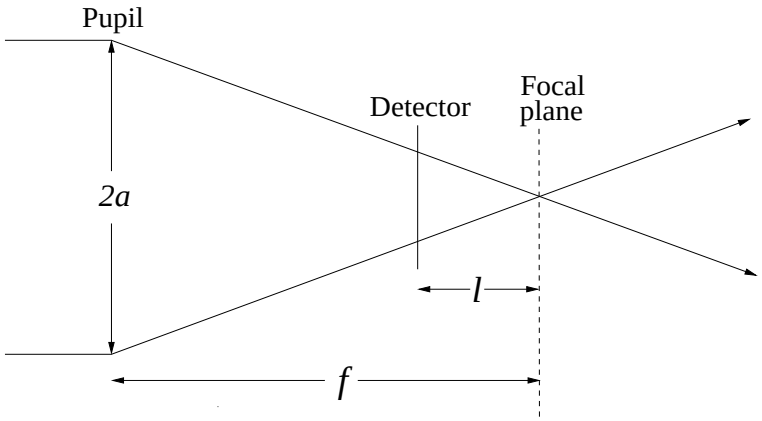

Figure 10: **Curvature sensor:** Principle of the curvature wavefront sensor with the detector plane being out-of-focus by a distance l with respect to the focal plane of the telescope.

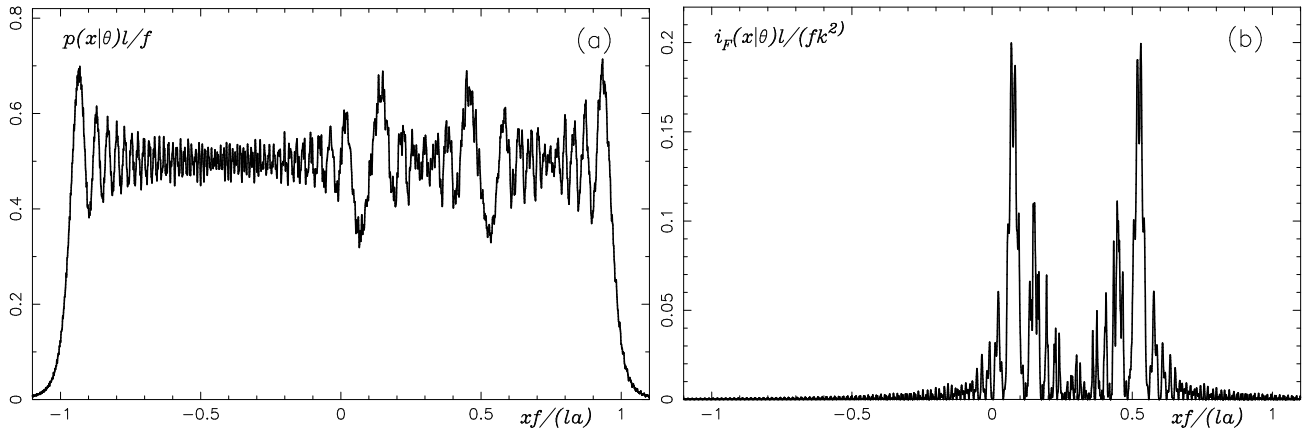

Figure 11: **Curvature sensor:** (a) Normalised likelihood function for a defocussing ratio $l/f = 0.002$ and $\theta = \lambda/10$. (b) Normalised Fisher information density for $l/f = 0.002$, both as functions of the normalised detector coordinate $xf/(la)$. For both plots : $a = 1m$, $f = 20m$, $c = 0.3a$, and $w = 0.2a$.

somewhat arbitrary. Requiring the distance of the first zero of the oscillating pattern to be less than a fraction $1/z$ of the segment half-width w leads to the condition

$$\frac{l}{f} > \left(1 + \frac{w^2}{0.8z^2\lambda f}\right)^{-1} \quad . \tag{30}$$

If, for different widths $2w$ of the segment, the parameter l/f is always chosen such that the condition (30) is fulfilled for $z = 10$, the dependence of the Fisher information on the segment width is linear, as in case of the Shack-Hartmann sensor, that is in the Fraunhofer limit. However, for the chosen non-overlap condition (30) the values for the Fisher information are approximately 30 times lower than in the Fraunhofer limit.

$I_{\mathrm{F,CV}}(\theta)$ is effectively independent of the segment location. This is similar to case of the Shack-Hartmann sensor, except that the central dip in the curve for the Shack-Hartmann sensor in figure 13 is not present in case of the curvature sensor with sufficient defocussing.

5 Comparison between the four sensors

The Cramér-Rao bounds for each of the four types of sensors have been discussed in detail in the preceeding sections. This section provides a comparison of the relative precision limits which can, in principle, be achieved

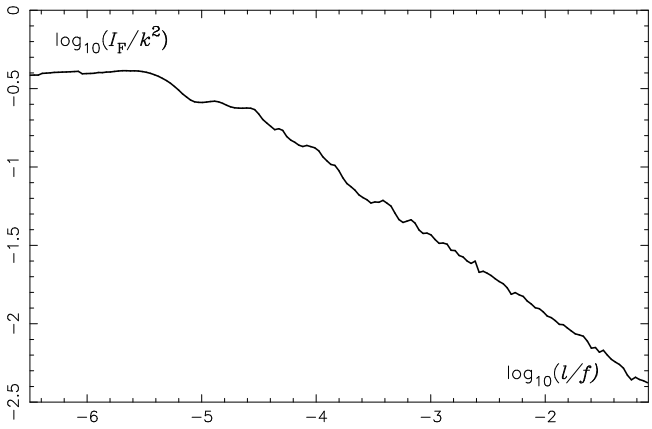

Figure 12: **Curvature sensor:** Normalised Fisher information as a function of the defocussing ratio l/f, and $a = 1m$, $f = 20m$, $c = 0.3a$, and $w = 0.2a$.

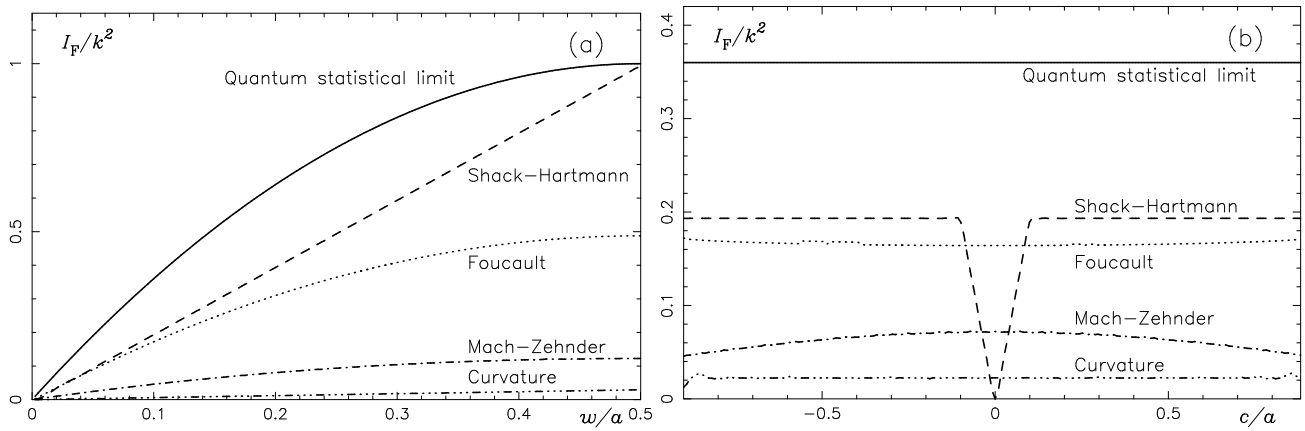

Figure 13: **All sensors:** Normalised Fisher information I_F/k^2 for the single-border configuration as functions (a) of the segment width $2w$ and (b) of the segment location for $w = 0.1a$. Also shown is the limiting quantum statistical information, which depends quadratically on the segment width. The curves for the Mach-Zehnder device are valid for a sharp-edge pinhole with a pinhole diameter $\zeta = 0.465\xi_0$ which maximises the information, and the curves for the curvature sensor for a defocussing which generates a clear separation of signals from adjacent borders.

with these sensors for the detection of differential segment phases. Figure 13a shows for the single-border configuration under ideal conditions the performance, defined by the Cramér-Rao bound, of the four sensors discussed in this paper, together with the corresponding quantum statistical limit.

Figure 13a demonstrates that the performance of the Shack-Hartmann sensor is superior to the one of the Foucault sensor, which in turn outperforms the Mach-Zehnder sensor for all segment widths. For small segment widths the Fisher information from the Shack-Hartmann and the Foucault sensors are effectively identical. However, for non-zero segment widths only the Shack-Hartmann sensor in the centred single-border configuration and a segment width of $w = a$ attains the quantum statistical limit.

For the Shack-Hartmann sensor the Fisher information depends linearly on the segment widths. The dependence is quadratic for the quantum statistical limit, the Foucault sensor, and the Mach-Zehnder sensor with pinhole diameters maximising the Fisher information.

For a Mach-Zehnder sensor with large pinhole diameters, which guarantee that the signals from adjacent borders are well separated, the Fisher information is a linear function of the segment width. The maximum value which is attained for the centred single-border configuration is approximately 0.03 if the signals from adjacent

borders are just separated, and decreases with a better separation of the signals. Also the Fisher information of the curvature sensor depends, like the Shack-Hartmann sensor as its limiting case for small amounts of defocussing, linearly on the segment width. However, the slope depends strongly on the amount of defocussing, and the curve in figure 13 has been calculated for a defocussing which guarantees that the half width of the segment is approximately ten times larger than the distance between a border and the position of the first zero of the oscillating pattern. See section 4.4 for further details.

Figure 13b shows that for all sensors, except for the Mach-Zehnder device, the Fisher information is effectively independent of the segment location within the pupil. However, for large pinhole diameters this also applies to the Mach-Zehnder sensor. The dip in the curve for the Shack-Hartmann sensor is irrelevant: when such a sensor is used in a telescope the subapertures would always be defined by smaller lenslets covering only the neighbourhood around a segment border, and the configuration would then be the single-border configuration.

6 Effects of noise and other measurement parameters

The effects of broadband illumination, atmospheric disturbances, detector readout noise, and discrete sampling of the likelihood function will only be discussed for one sensor, namely the Shack-Hartmann sensor in the centred single-border configuration. However, the methods used in this section can readily be applied to all other types of sensors and configurations.

6.1 Effects of broadband measurements

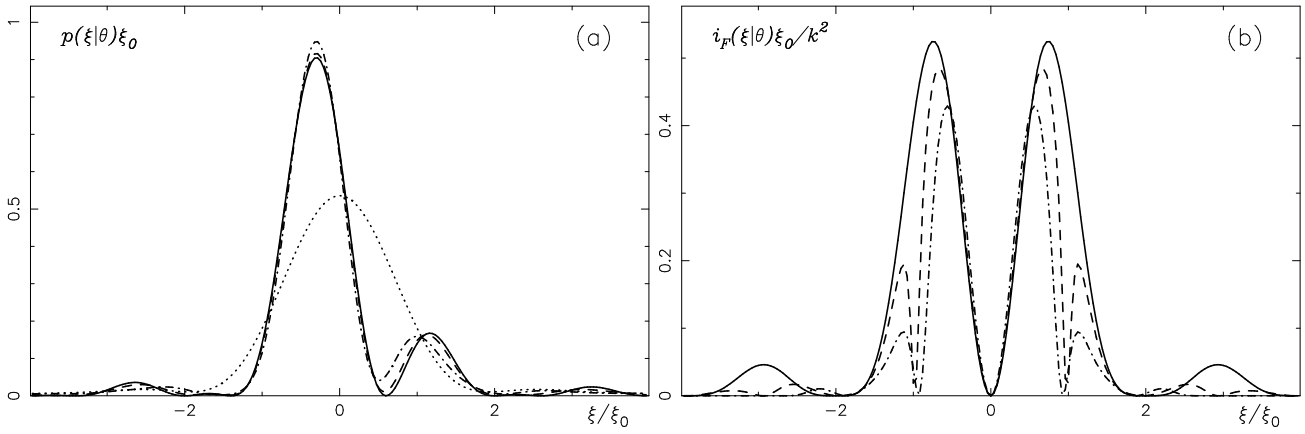

Figure 14: **Shack-Hartmann sensor:** (a) Likelihood functions for a phase step $\theta = \lambda/5$, (b) normalised Fisher information densities for $\theta = 0$, integrated over different bandpass ratios $\Delta\lambda/\lambda_c$: Monochromatic (solid), $\Delta\lambda/\lambda_c = 0.25$ (dashed), $\Delta\lambda/\lambda_c = 0.5$ (dashed-dotted). The dotted curve in (a) is the limit to which the likelihood functions converge for very large phase steps, that is $k\theta \gg 1$, and a bandpass ratio of $\Delta\lambda/\lambda_c = 0.5$.

The number of available photons can be increased with broadband measurements, that is by using a larger bandpass of the light. The corresponding gain in the information, being proportional to the number of photons, is, however, partially compensated by a loss of information due to a modification of the likelihood function. As shown below, the Cramér-Rao bound decreases with an increase of the bandpass, strongly at small bandpasses, but only slowly at large bandpasses.

The likelihood function for a broadband measurement is obtained by integrating the monochromatic likelihood function over the wavenumber k around a central wavenumber k_c with a suitable prior probability distribution or spectral density $p_k(k)$ as a weight function. In the following calculations $p_k(k) = \text{const}$ is assumed for any bandpass. Figure 14a shows the likelihood function for three bandpasses and a phase step $\theta = \lambda/5$. Apparently,

for small phase steps, the position of the maximum of the likelihood function and the width of the central core are only weakly affected by the bandpass.

The dotted line in figure 14 shows the limit to which all curves converge for very large phase steps. For such large phase steps, that is for $k\theta \gg 1$, the sine and cosine functions in equation (A-4) containing $k\theta$ vary rapidly with k. Each term containing $k\theta$ can therefore be replaced by its average value which is $1/2$ for the two quadratic term and $(1 - \cos(ka\xi))/2$ for the mixed term, respectively. The dotted curve is then given by

$$p_{k\theta\gg1}(\xi) = \frac{1}{(\pi k a)^2} \, \frac{1}{k_2 - k_1} \int_{k_1}^{k_2} \mathrm{d}k \, \frac{1}{\xi^2} \left[1 - \cos(ck\xi)\cos(k\xi a)\right] \quad , \tag{31}$$

where k_1 and k_2 are the integration limits. Figure 14b shows the curves for the Fisher information densities for

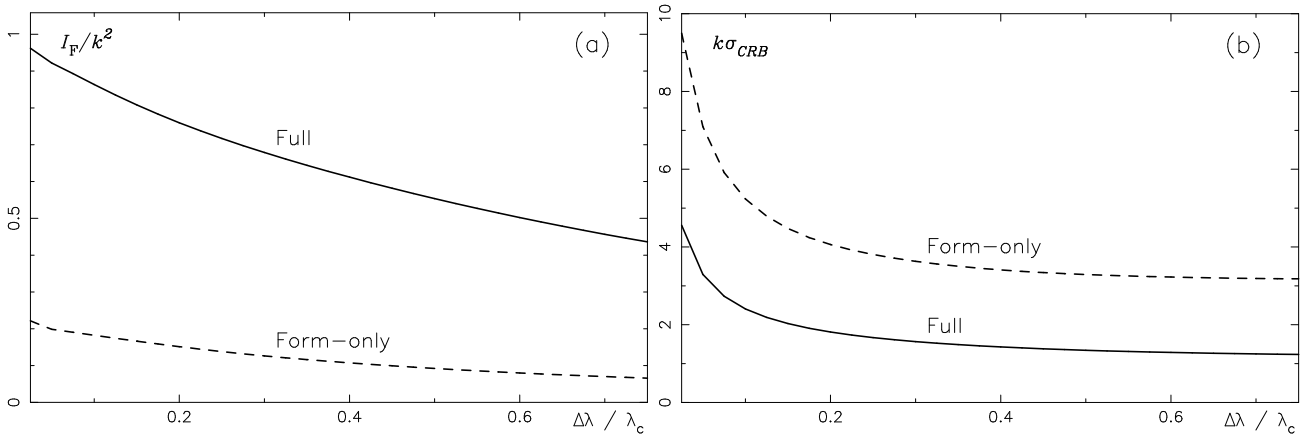

Figure 15: **Shack-Hartmann sensor:** (a) Normalised Fisher information I_F/k^2 and (b) normalised Cramér-Rao bound σ_CRB as functions of the ratio of the width $\Delta\lambda$ of the bandpass to the central wavelength λ_c. *Solid*: full information, *dashed*: form-only information.

the same bandpasses as in figure 14a, but for $\theta = 0$. In the ideal case the peaks in the Fisher information density are, according to equation (1), due to the zeroes or small values of the likelihood function at those locations. The peaks are now reduced because the zeroes and the small values of the likelihood functions are washed out by the averaging process. The peaks of the Fisher information density at $\pm 3\zeta/\zeta_0$ for $\theta = 0$ disappear nearly completely for larger bandpasses, and the Fisher information density is therefore concentrated closer to the origin.

Figure 15a shows that, due to the weak dependencies of the form of the core and of the position of the maximum on the bandpass, the Fisher information per photon as a function of the bandpass decreases only weakly with the bandpass. This is true for the full information as well as for the form-only information. See section 4.1 for the definition of form. The decrease of the Fisher information per photon with an increase of the bandpass is slower than the proportional increase with the number of photons. Therefore the Cramér-Rao bound σ_CRB of the error in the estimate of the phase step decreases continuously with an increase of the bandpass. This can also be seen in figure 15b which shows σ_CRB in units of $\lambda/(2\pi)$ for one photon per nanometer of the bandpass.

6.2 Effects of atmospheric turbulence

Turbulence and temperature inhomogeneities in the atmosphere lead to a non-planarity of the incoming wavefront. These disturbances reduce the Fisher information, and consequently reduce the precision with which the phase step can be estimated.

The effects of the atmosphere can be taken into account by writing the complex amplitude $U(x)$ in equation (A-2) as a product of the complex amplitude $U_\mathrm{s}(x)$ related to the relative segment phase and a stochastic

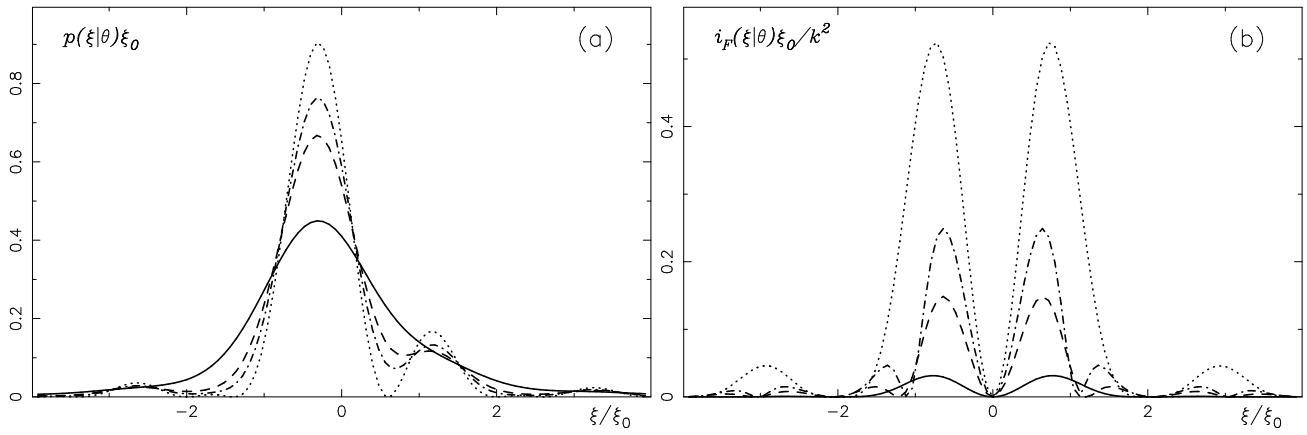

Figure 16: **Shack-Hartmann sensor:** (a) Likelihood functions for a phase step $\theta = 0.1\lambda$ under different seeing conditions, (b) corresponding normalised Fisher information densities for $\theta = 0$. Seeing conditions: *Dotted*: no atmosphere, *Dashed-dotted*: $r_0/2a = 2.25$, *Dashed*: $r_0/2a = 1.5$, *Solid*: $r_0/2a = 0.75$.

function $U_{\rm atm}(x)$ describing the disturbance introduced by the atmosphere. The likelihood function in equation (A-3) can then be written as

$$p(\xi \,|\, \theta) = \frac{k}{4}\pi a \int_{-a}^{+a} \mathrm{d}x_{\rm p}{}' \int_{-a-x_{\rm p}}^{+a-x_{\rm p}} \mathrm{d}x_{\rm p}\, U_{\rm s}^*(x_{\rm p}{}')U_{\rm s}(x_{\rm p}{}' + x_{\rm p}) \; < U_{\rm atm}^*(x_{\rm p}{}' + x_{\rm p})U_{\rm atm}(x_{\rm p}{}') > \; e^{ik\xi x_{\rm p}} \quad . \tag{32}$$

The expression in the brackets is the atmospheric structure function. For Kolmogorov turbulence and neglecting the effects of the inner and outer scales, it is given by

$$< U_{\rm atm}(x_{\rm p}{}' + x_{\rm p})U_{\rm atm}(x_{\rm p}{}') >= \exp\left[-3.44\left(\frac{x_{\rm p}}{r_0}\right)^{5/3}\right] \quad , \tag{33}$$

where r_0 is the atmospheric coherence length. It can be shown that the likelihood function for a given wavelength depends on ξ and ξ_0 only through the ratio ξ/ξ_0 and on r_0 and a only through the ratio r_0/a.

Figure 16a shows, for a phase step of $\theta = 0.1\lambda$, the influence of the atmosphere on the likelihood function for four different seeing conditions, expressed as ratios $r_0/(2a)$. As in case of the integration over a finite bandpass, the location of the maximum does not depend on the seeing conditions. However, the secondary maxima are washed out and the width of the central core increases strongly with deteriorating seeing conditions.

The Fisher information densities for a phase step $\theta = 0$ and four different seeing conditions are shown in figure 16b. As in case of larger bandpasses and for similar reasons as given in the corresponding section 6.1, under the influence of the atmosphere, the Fisher information density is more concentrated near the origin than the Fisher information density under ideal conditions. For the design of the wavefront sensor one would usually assume a given seeing, that is a given value for r_0, and then define the width of the lens. Therefore, figure 17a shows the normalised Fisher information as a function of the ratio $2a/r_0$, the full information as the solid line and the form-only information as the dashed line. The strong dependence of the width of the core on the seeing, already mentioned above, also explains the strong decrease of the Fisher information with an increase of the ratio $2a/r_0$. If the diameter of the aperture is equal to r_0, the Fisher information is already reduced by a factor of approximately 10 compared to the Fisher information under ideal conditions.

On the one hand, the increase of the width of the lens increases the Fisher information due to the larger number of photons. On the other hand, the Fisher information per photon decreases rapidly with the size of the lens. The combined effect is given by the product of the Fisher information per photon times the width of the lens. The corresponding curves for the Cramér-Rao bounds in units of $\lambda/(2\pi)$ are plotted in figure 17b for a flux density of one photon per millimeter of the subaperture. Contrary to the situation for broadband measurements, the Cramér-Rao bound has a minimum, which is attained for a ratio $2a/r_0 \approx 0.4$.

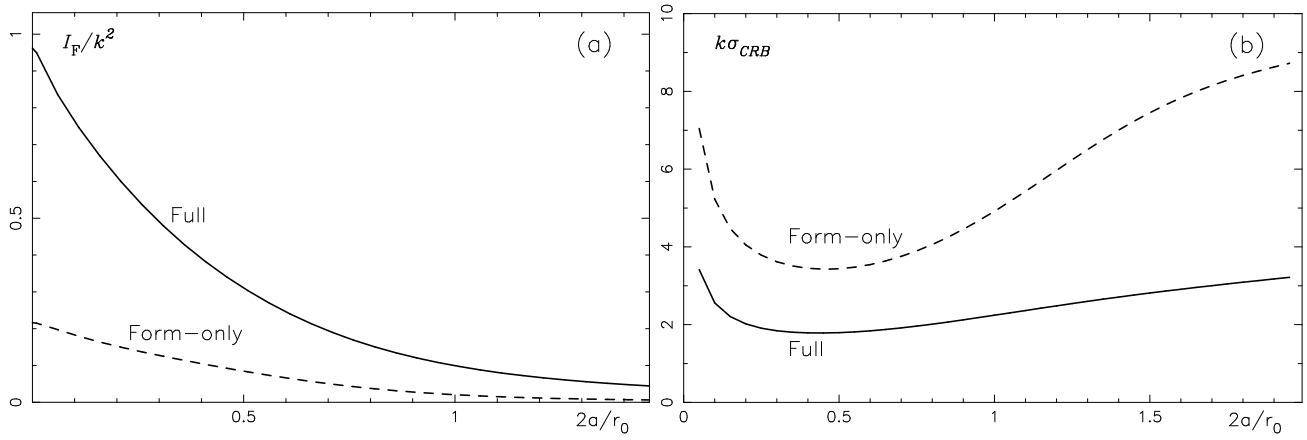

Figure 17: **Shack-Hartmann sensor:** (a) Normalised Fisher information I_F/k^2 and (b) normalised Cramér-Rao bound σ_{CRB} as functions of the ratio of the aperture width $2a$ to the atmospheric coherence length r_0.

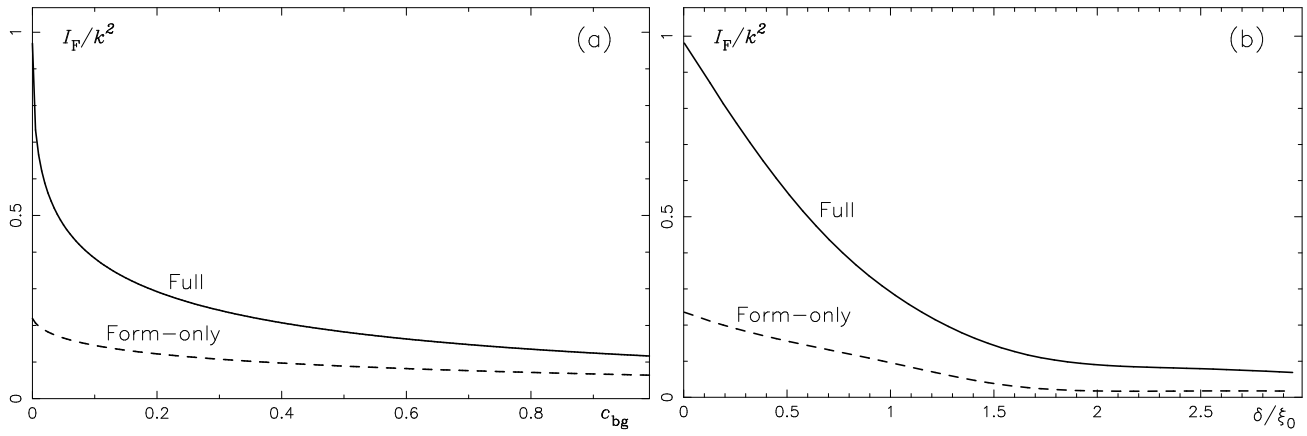

Figure 18: **Shack-Hartmann sensor:** (a) Normalised Fisher information as a function of the readout flux in units of the flux from the sky source. (b) Normalised Fisher information as a function of the pixel half-width δ.

6.2.1 Effects of detector readout noise

The effect of the readout noise of the detector can be modelled by adding a constant background c_{bg} to the likelihood function in equation (A-4):

$$p(\xi \,|\, \theta) = \frac{c_{\mathrm{bg}}}{\xi_0} + p_0(\xi \,|\, \theta) \quad . \tag{34}$$

Here $p_0(\xi \,|\, \theta)$ is the likelihood function for one photon without background noise. The value of the constant c_{bg} is the ratio of the constant background to the peak in the ideal likelihood function for $\theta = 0$.

Figure 18a shows for a narrowband measurement the full Fisher information and the form-only Fisher information as functions of c_{bg}. The full Fisher information declines initially much faster than the form-only Fisher information. The reason for this is that the zeroes in the likelihood function, which are, according to equation (1), responsible for the strong peaks in the full Fisher information density, disappear because of the constant background. Since, for large backgrounds, the ratio of the full to the form-only Fisher information only equals approximately two, the full Fisher information is also more strongly affected than the form-only Fisher information for a large backgrounds. Regarding figure 18b, see next section.

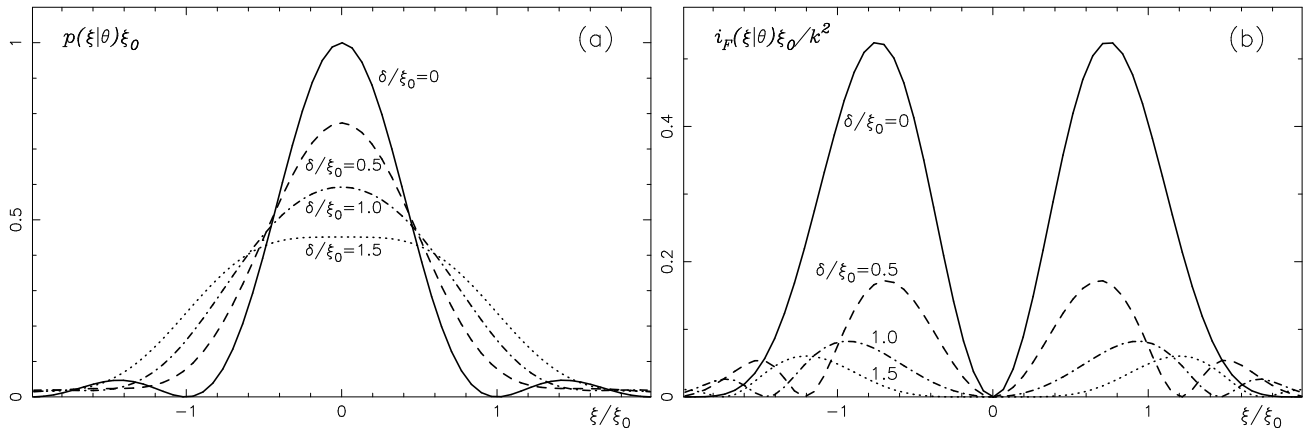

Figure 19: **Shack-Hartmann sensor:** (a) Normalised likelihood functions for pixel half-width of $\delta \to 0$ (solid), $\delta = 0.5\xi_0$ (dashed), $\delta = \xi_0$ (dashed-dotted), and $\delta = 1.5\xi_0$ (dotted). (b) Corresponding normalised Fisher information densities.

6.2.2 Effects of the detector pixelization

The effect of finite pixel sizes δ can be modelled by a product of a comb function, describing the positions of the centres of the pixels, and a continuous likelihood function which is obtained by averaging, for each ξ, the initial likelihood function over the interval $[\xi - \delta/2, \xi + \delta/2]$. Figure 19a shows the likelihood functions for $\delta \to 0$ (solid), $\delta = 0.5\xi_0$ (dashed), $\delta = \xi_0$ (dashed-dotted), and $\delta = 1.5\xi_0$ (dotted). The corresponding Fisher information densities are shown in figure 19b for the full information. The strong decline of the Fisher information densities with the pixel size is also apparent from figure 18b, which shows that the decrease is linear for small pixel widths. The effect of the finite pixels size is similar for the full Fisher information and the form-only Fisher information.

7 Conclusions

This paper has shown that the quantum statistical information relating to a relative segment phase depends quadratically on the ratio of the segment width to the width of the aperture. Its maximum, given by the square of the wavenumber per photon passing through the aperture, is attained for the case where the segment covers half of the aperture. In this special case the limiting precision for the measurement of the relative segment phase error is equal $\lambda/(2\pi)$, where λ is the wavelength of the light. Under ideal conditions, that is with narrowband measurements and for negligible atmospheric or detector noise, both the Shack-Hartmann and the Foucault sensor can, for segment sizes that are small compared with the diameter of the full aperture, potentially extract approximately half of the ultimate quantum statistical information. For the special case where the aperture is only filled by two segments with the phase step at the centre of the aperture, the Shack-Hartmann device even attains the ultimate precision limit, and the precision obtained by the Foucault sensor is smaller by a factor of $\sqrt{2}$. The other two sensors, Mach-Zehnder and curvature, require in general approximately ten to twenty times the number of photons in order to reach the same precision as the first two sensors. In case of the Shack-Hartmann sensor, broadband measurements always increase the precision, whereas in the presence of atmospheric disturbances the optimum width for the subaperture corresponding to one lenslet approximately equals half of the atmospheric coherence length. Also detector readout noise and the effects of a finite sampling strongly reduce the precision. The effects of these disturbances on the other sensors may in detail be different from the effects on the Shack-Hartmann sensor and remain to be investigated.

Acknowledgements

The authors would like to thank R.N. Wilson, R. Mazzoleni, and N. Yaitskova for useful suggestions on the manuscript and a critical reading of the final version. This work is embedded in the Active Phasing Experiment

which is supported by by the Sixth European Framework Program for Research and Technological Development under the contract number 011863.

Appendix A : Expressions for the likelihood function and Fisher information density for the four sensors

In this appendix the expressions for the likelihood functions and Fisher information densities for the four sensors discussed in the main text are derived.

A1: Shack-Hartmann sensor

For monochromatic light with wavelength λ the Fraunhofer diffraction pattern in the focal plane of an infinitely long cylindrical lens covering a one-dimensional subaperture of width $2a$ on the mirror can be calculated with a Fourier transformation. The diffraction pattern may be viewed as a likelihood function $p(\xi \mid \theta)$ for the detection of a photon at a coordinate ξ in the focal plane. ξ is a unitless position variable, related to a physical position variable x in the focal or detector plane by $\xi = x/f$, where f is the focal length of the lens. If x_p and x_p' are coordinates perpendicular to the axis of the cylindrical lens, and U the complex field in the plane of the cylindrical lens, the likelihood function is given by

$$p(\xi \mid \theta) = \frac{k}{2\pi} < \int_{-a}^{+a} \mathrm{d}x_\mathrm{p} U^*(x_\mathrm{p}'') e^{-ik\xi x_\mathrm{p}''} \int_{-a}^{+a} \mathrm{d}x_\mathrm{p}' U(x_\mathrm{p}') e^{ik\xi x_\mathrm{p}'} > \quad \text{(A-1)}$$

$$= \frac{k}{2\pi} \int_{-a}^{+a} \mathrm{d}x_\mathrm{p}' \int_{-a}^{+a} \mathrm{d}x_\mathrm{p} \; < U^*(x_\mathrm{p}'') U(x_\mathrm{p}') > e^{ik\xi(x_\mathrm{p}' - x_\mathrm{p}'')} \quad . \quad \text{(A-2)}$$

The bracket denotes a statistical average over a possibly stochastic ensemble of complex amplitudes U. In this paper such a stochastic ensemble is used for the treatment of atmospheric effects in section 6.2.

After the substitution $x_\mathrm{p} = x_\mathrm{p}' - x_\mathrm{p}''$ one gets

$$p(\xi \mid \theta) = \frac{k}{2\pi} \int_{-a}^{+a} \mathrm{d}x_\mathrm{p}' \int_{-a-x_\mathrm{p}}^{+a-x_\mathrm{p}} \mathrm{d}x_\mathrm{p} < U^*(x_\mathrm{p}') U(x_\mathrm{p}' + x_\mathrm{p}) > \; e^{ik\xi x_\mathrm{p}} \quad . \quad \text{(A-3)}$$

Next, the non-stochastic wave function U_s related to the relative segment phase as defined in equation (9), is introduced for U in equation (A-3). With $\xi_0 = \lambda/(2a)$ as the location of the first zero in the diffraction pattern for $\theta = 0$ one gets

$$p(\xi \mid \theta) = \frac{1}{\xi_0} \frac{1}{(\pi\xi/\xi_0)^2} \Big[\sin^2(\pi\xi/\xi_0) + 2\sin^2(\pi\frac{w}{a}\xi/\xi_0)(1 - \cos(2\pi\theta/\lambda)) + 2\sin(\pi\xi/\xi_0)\sin(\pi\frac{w}{a}\xi/\xi_0)\left(\cos(\pi\frac{c}{a}\xi/\xi_0 - 2\pi\theta/\lambda) - \cos(\frac{c}{a}\xi/\xi_0)\right)\Big] \quad . \quad \text{(A-4)}$$

The Fisher information can easily be computed by introducing the amplitude $q(\xi \mid \theta)$, calculated from equation (A-4) following equation (4), into equation (8). The lengthy expression for arbitrary values of c, w, and θ is not quoted here. For the special case of $\theta \to 0$ one obtains the following much shorter expression:

$$i_\mathrm{F}(\xi \mid \theta \to 0) = \frac{4}{\pi^2} \frac{k^2}{\xi_0} \frac{1}{(\xi/\xi_0)^2} \sin^2(\pi\frac{w}{a}\xi/\xi_0) \sin^2(\pi\frac{c}{a}\xi/\xi_0) \quad . \quad \text{(A-5)}$$

A2: Mach-Zehnder sensor

The amplitude U_1 in the first, unmodified beam in figure 4 is given by equation (9). After the light in the second beam passed through the phase plate and the spatial filter, the amplitude U_2 is given by the convolution

$$U_2(x) = \frac{1}{\sqrt{\lambda}} \int_{-\infty}^{\infty} \mathrm{d}x' \, U_1(x') T(x - x') \quad , \quad \text{(A-6)}$$

where $T(x)$ is the inverse Fourier transform of the transmission function (29) of the pinhole. Since the detectors are in image planes of the pupil, the same coordinates can be used as for the pupil, that is $x = x_\mathrm{p}$. The likelihood functions on the two detectors are then

$$p_{1,2}(x\,|\,\theta) = \frac{1}{2}\left\{|U_1(x\,|\,\theta)|^2 + |U_2(x\,|\,\theta)|^2 \pm 2\,\mathrm{Re}[U_1^*(x\,|\,\theta)U_2(x\,|\,\theta)e^{i\phi}]\right\} \quad , \tag{A-7}$$

where the plus sign refers to p_1 and the minus sign to p_2. With the special choice of the phase shift of $\phi = \pi/2$ and the abbreviations

$$b = k\zeta \tag{A-8}$$

$$\Psi_\mathrm{L} = \Phi(-b(a+x), b(c-w-x)) \tag{A-9}$$

$$\Psi_\mathrm{M} = \Phi(b(c-w-x), b(c+w-x)) \tag{A-10}$$

$$\Psi_\mathrm{H} = \Phi(b(c+w-x), b(a-x)) \tag{A-11}$$

$$\Phi(u,v) = \left\{ \begin{array}{ll} (2/\pi)\int_u^v \mathrm{d}z\ \sin z\,/\,z & \text{sharp edge pinhole} \\ (2/\sqrt{\pi})\int_u^v \mathrm{d}z\ \exp(-z^2) & \text{Gaussian pinhole} \end{array} \right\} \quad , \tag{A-12}$$

one obtains for the two likelihood functions

$$\begin{aligned} p_{1,2}(x\,|\,\theta) = \frac{1}{8a}\{&W(x,a,-a) + \frac{1}{4}\left[(\Psi_\mathrm{L}+\Psi_\mathrm{H})^2 + \Psi_\mathrm{M}^2 + 2\cos(k\theta)(\Psi_\mathrm{L}+\Psi_\mathrm{H})\Psi_\mathrm{M}\right] \\ &\pm \sin(k\theta)\,[W(x,c-w,c+w)(\Psi_\mathrm{L}+\Psi_\mathrm{H}) \\ &\qquad -(W(x,-a,c-w)+W(x,c+w,a))\Psi_\mathrm{M}]\} \end{aligned} \tag{A-13}$$

One of the assumptions for the derivation of equation (2) was that the integrated intensity is independent of θ. In case of the Mach-Zehnder interferometer this is strictly only true for $\theta \to 0$. For both types of pinholes the integrated intensity is a symmetric function of θ with, according to numerical calculations, a quasi parabolic behaviour around $\theta = 0$. Around the maximum at $\theta = 0$ the derivative of the integrated intensity with respect to θ is therefore zero.

The Fisher information is identical in the two beams behind the second beamsplitter. With the definition

$$\Psi^2 = \Psi_\mathrm{L}^2 + \Psi_\mathrm{M}^2 + \Psi_\mathrm{H}^2 \quad , \tag{A-14}$$

the total Fisher information as a sum of the Fisher information from both beams behind the second beamsplitter is, in the limit $\theta \to 0$, given by

$$I_\mathrm{F,MZ} = \frac{k^2}{4a}\left[\int_{-a}^{+a} \mathrm{d}x\ \Psi_\mathrm{M}^2/\left(1+\Psi^2/4\right) + \int_{c-w}^{c+w} \mathrm{d}x\ \left((\Psi_\mathrm{L}^2+\Psi_\mathrm{H}^2)^2 - \Psi_\mathrm{M}^2\right)/\left(1+\Psi^2/4\right)\right] \quad . \tag{A-15}$$

A3: Foucault sensor

The Fourier transform $T(x)$ of a Heaviside step function representing the effect of the knife edge in the focal plane is given by

$$T(x) = \frac{\sqrt{\lambda}}{2\pi}\left(\pm\frac{1}{ix} + \pi\delta(x)\right) \quad , \tag{A-16}$$

where the signs + and − refer to the two channels with the transmission equal to zero for $\xi < 0$ and $\xi > 0$, respectively. Introducing equation (A-16) into equation (A-6) gives with the definitions

$$Y_1 = \ln\frac{(c+w-x)^2}{(c-w-x)^2} \tag{A-17}$$

$$Y_2 = \ln\frac{(a-x)^2}{(a+x)^2} \tag{A-18}$$

the expression for the likelihood function on the two detectors

$$\begin{aligned} |U_{2,\mathrm{FC}}(x\,|\,\theta)|^2 = \frac{1}{8a\pi^2}\{&\frac{1}{4}Y_1^2 + \pi^2 W(x,-a,a) + \frac{1}{2}(1-\cos(k\theta))\,Y_2(Y_2-Y_1) \\ &\pm \sin(k\theta)\pi(W(x,-a,a)Y_2 - W(x,c-w,c+w)Y_1)\} \quad . \end{aligned} \tag{A-19}$$

The inversion of the sign in front of the sine-term can be understood from the fact that the intensity distribution must be symmetric with respect to a change of the transmission from $\xi < 0$ to $\xi > 0$ and a simultaneous change of the sign of the phase step θ. A similar expression for the centred single-border configuration has been derived in [14].

For one channel alone the requirement for the integrated intensity to be independent of θ is not fulfilled, since in the vicinity of $\theta = 0$ the centre of gravity of the diffraction pattern in the focal plane depends linearly on θ, as shown by equation (26). However, the requirement is obviously fulfilled for the sum of the integrated intensities in the two channels. Therefore, the formalism has to use the sum of the Fisher information densities in the two channels, which are, however, in the limit $\theta \to 0$, independent of the channel. For $\theta \to 0$ the sum of the Fisher information in the two channels is given by

$$I_{\mathrm{F,FC}} = \frac{k^2}{4a} \int_{-a}^{+a} \left(Y_2 - W(x, c-w, c-w) Y_1\right)^2 \Big/ \left(\pi^2 + Y_1^2/4\right) \tag{A-20}$$

A4: Curvature sensor

The one-dimensional equivalent of the two-dimensional expression for the complex amplitude in the plane of the detector given in [5] [7] is

$$U_{\mathrm{CV}}(x \,|\, \theta) = \frac{1}{\sqrt{\lambda(f-l)}} e^{ikf} e^{i\pi x^2/(\lambda f)} \int_{-a}^{+a} \mathrm{d}x_{\mathrm{p}}\, U(x_{\mathrm{p}} \,|\, \theta)\, e^{i\frac{\pi}{2}(x_{\mathrm{p}} - fx/l)^2/\rho^2} \quad , \tag{A-21}$$

where U is the complex amplitude in the pupil and

$$\rho = \sqrt{\frac{\lambda f(f-l)}{2l}}. \tag{A-22}$$

With the definition

$$\eta = (x_{\mathrm{p}} - \frac{f}{l} x)/\rho \tag{A-23}$$

equation (A-21) can be written as

$$U_{\mathrm{CV}}(x \,|\, \theta) = \sqrt{\frac{i}{2}} \sqrt{\frac{f}{l}}\, e^{ikf}\, e^{i\pi x^2/(\lambda f)} \int_{(-a-fx/l)/\rho}^{(+a-fx/l)/\rho} \mathrm{d}\eta\, U_{\mathrm{s}}(\eta \,|\, \theta)\, e^{i\pi\eta^2/2} \quad . \tag{A-24}$$

The expression for U_{CV} has been normalised such that

$$\int_{\infty}^{+\infty} \mathrm{d}x\, |U_{\mathrm{CV}}(x \,|\, \theta)|^2 = 1 \tag{A-25}$$

Introducing U_{s} from equation (9) for U in equation (A-24) one obtains with the abbreviations

$$\begin{array}{lllllll} C_1 & = & \mathcal{C}((-a - fx/l)/\rho) & S_1 & = & \mathcal{S}((-a - fx/l)/\rho \\ C_2 & = & \mathcal{C}((c - w - fx/l)/\rho) & S_2 & = & \mathcal{S}((c - w - fx/l)/\rho \\ C_3 & = & \mathcal{C}((c + w - fx/l)/\rho) & S_3 & = & \mathcal{S}((c + w - fx/l)/\rho \quad , \\ C_4 & = & \mathcal{C}((a - fx/l)/\rho) & S_4 & = & \mathcal{S}((a - fx/l)/\rho \\ C_{ij} & = & C_j - C_i & S_{ij} & = & S_i - S_j \end{array} \tag{A-26}$$

where $\mathcal{C}$ and $\mathcal{S}$ denote the Fresnel integrals, the following expression for the likelihood function:

$$\begin{aligned} |U_{\mathrm{CV}}(x \,|\, \theta)|^2 \;=\; & \frac{f}{4al} \left(C_{21}^2 + S_{21}^2 + C_{32}^2 + S_{32}^2 + C_{43}^2 + S_{43}^2\right. \\ & +2\, \cos(k\theta)\, [C_{21}C_{32} + S_{21}S_{32} + C_{32}C_{43} + S_{32}S_{43}] \\ & +2\, \sin(k\theta)\, [S_{12}C_{32} - C_{21}S_{32} + S_{32}C_{43} - C_{32}C_{43}] \\ & \left. +2C_{21}C_{43} + 2S_{21}S_{43}\right) \quad . \end{aligned} \tag{A-27}$$

The corresponding expression for the Fisher information density can readily be calculated from the equations (5) and (A-27), and is not presented here.

References

[1] G. Chanan, M. Troy, F. Dekens, J. Nelson, T. Mast, and D. Kirkman, 1998, Applied Optics, **37** 140 (1998)

[2] G. Chanan, Ohara, C., and Troy, M., *Applied Optics*, **39** 4706 (2000)

[3] G. Chanan and D.G. MacMartin, Applied Optics **43** 608 (2004)

[4] N. Yaitskova, K. Dohlen, P. Dierickx, L. Montoya, 2005, J. Opt. Soc. Am. A **22** 1093 (2005)

[5] F. Roddier, *Curvature sensing: a diffraction theory*, NOAO Advanced Development Program, 87-3

[6] G. Chanan, M. Troy and E. Sirko, Applied Optics **38** 704 (1999)

[7] J.M.R. Gonzalez, *Sensor de curvatura en telescopios con espejo segmentado*, Thesis, Departamento de Astrofisica, Universidad de La Laguna, Spain (2001)

[8] A. Schumacher, N. Devaney, Mon. Not. R. Astron. Soc **366** 537 (2005)

[9] K. Dohlen, in *Astronomy with High Contrast Imaging II*, edited by C. Aime and R. Soummer, (EAS Publication Series **12**, 2004), pp. 33–44

[10] S. Esposito, N. Devaney, in *Beyond Conventional Adaptive Optics*, 2001, (ESO Proc. vol. 58, 2001), pp. 161–166

[11] S. Esposito,, E. Pinna, A. Tozzi, P. Stafanini, N. Devaney, Proc. SPIE **5169** 72 (2003)

[12] F. Gonte, N. Yaitskova, P. Dierickx, *et al.*, Proc. SPIE **5489** (2004), 1184

[13] F. Gonte, N. Yaitskova, F. Derie, *et al.*, Proc. SPIE **5894** (2005), 306

[14] E. Pinna, *Applicazione del sensore a piramide alla misura di discontinuita del fronte d'onda*, Ph.D. thesis, Universita degli Studi di Firenze, Italy (2004)

[15] S. Esposito, E. Pinna, A. Puglisi, A. Pozzi and P. Stefanini, Optics Letters **30**(19), 2772

[16] M. Kendall, A. Stuart, *The Advanced Theory of Statistics, Vol. 2*, (Macmillan, New York, 1979)

[17] S.L. Braunstein, C.M. Caves, and G.J. Milburn, Annals of Physics **247** 135 (1996)