

# General formalism for Fourier-based wave front sensing

OLIVIER FAUVARQUE,<sup>1,\*</sup> BENOIT NEICHEL,<sup>1</sup> THIERRY FUSCO,<sup>1,2</sup> JEAN-FRANCOIS SAUVAGE,<sup>1,2</sup> AND ORION GIRAULT<sup>1</sup>

<sup>1</sup>Aix Marseille Université, CNRS, LAM (Laboratoire d'Astrophysique de Marseille) UMR 7326, 13388 Marseille, France

<sup>2</sup>ONERA—the French Aerospace Laboratory, F-92322 Chatillon, France

\*Corresponding author: olivier.fauvarque@lam.fr

Received 29 July 2016; revised 27 October 2016; accepted 30 October 2016 (Doc. ID 272627); published 2 December 2016

We introduce in this paper a general formalism for Fourier-based wave front sensing. To do so, we consider the filtering mask as a free parameter. Such an approach allows us to unify sensors like the pyramid wave front sensor (PWFS) and the Zernike wave front sensor (ZWFS). In particular, we take the opportunity to generalize these two sensors in terms of sensors' class, where optical quantities such as the apex angle for the PWFS or the depth of the Zernike mask for the ZWFS become free parameters. In order to compare all the generated sensors of these two classes thanks to common performance criteria, we first define a general phase-linear quantity that we call meta-intensity. Analytical developments allow us to then split the perfectly phase-linear behavior of a WFS from the nonlinear contributions, making robust and analytic definitions of the sensitivity and the linearity range possible. Moreover, we define a new quantity called the SD factor, which characterizes the trade-off between these two antagonistic quantities. These developments are generalized for a modulation device and polychromatic light. A nonexhaustive study is finally conducted on the two classes, allowing us to retrieve the usual results and also make explicit the influence of the optical parameters introduced above. © 2016 Optical Society of America

**OCIS codes:** (010.1080) Active or adaptive optics; (010.7060) Turbulence; (010.7350) Wave-front sensing; (110.6770) Telescopes.

<https://doi.org/10.1364/OPTICA.3.001440>

## 1. INTRODUCTION

By placing amplitude or phase masks in a focal plane, it is possible to filter the light from one pupil plane to another. Those masks are able, in particular, to transform incoming phase fluctuations into intensity variations on a detector.

Such optical designs are thus particularly relevant in the context of wave front sensing, especially for the adaptive optics (AO). Moreover, Fourier-based wave front sensors (WFSs) have many advantages compared to other WFSs, such as the Shack–Hartmann, in terms of, for instance, noise propagation or sampling flexibility.

The historical example of those Fourier-based WFSs, dating from 1858, is the famous Foucault's knife. Ragazzoni generalized this physical concept with the pyramid WFS [1], which consists of a 4-face pyramidal mask in the focal plane. Variations of this concept regularly appear in the literature and define new WFSs: Akondi *et al.* [2] and Vohnsen *et al.* [3] tested new pyramidal masks by changing the number of the faces of the pyramid. We recently proposed the flattened pyramid WFS [4], which introduces a new way to use the pyramid mask by reducing its apex angle.

Using another physical concept based on the phase contrast method, Zernike introduced the Zernike WFS [5], where a Fourier mask, completely transparent, has a circular depression in its midst, allowing one to create interference between the spatial frequencies of the incoming phase.

The essential purpose of this paper is to merge all these Fourier-based WFSs under the same mathematical formalism in order to build robust and relevant criteria allowing researchers to compare their performance in the context of AO wave front sensing. Such an approach has been described, for instance, in the context of phase masks in astronomy regarding interferometry and coronagraphy (see [6]).

In the first section of this paper, we will present an original interpretation of the Fourier filtering technique thanks to the focal plane tessellation formalism. Such an approach will allow us to describe all the masks involved in AO wave front sensing in a unique mathematical framework. We will then define a unified post-processing in order to create, from the detector output signal, a quantity called the meta-intensity, which will be linear with the phase. Thanks to an analytical development of the meta-intensity for any Fourier mask, we will, in the third part, define rigorously the sensitivity and the linearity range depending only on the choice of the Fourier mask used to do wave front sensing. Moreover, we will show why these two performance criteria are inevitably antagonistic thanks to a third criterion, the SD factor, which will quantify the trade-off between them. We will also give some hints regarding the optimization of WFSs by considering masks as free-form objects. We will then show, in the fourth section, how it is possible to generalize these results in the context of a modulation in the pupil plane upstream to the filtering mask. In

the fifth and sixth sections, we will apply all these mathematical developments to the PWFS and the ZWFS. If, in the four first sections, the light will be considered as monochromatic, the last section will be dedicated to the case of polychromatic light. In particular, we will show how our formalism allows us to define the chromaticity of a sensor.

## 2. WAVE FRONT SENSING AND FOURIER FILTERING

### A. Optical System

We consider the optical system shown in Fig. 1. The first plane contains the pupil and a focusing device corresponding to a perfect telescope. The associated spatial variables of this plane are  $x_p$  and  $y_p$ ; the  $p$  index refers to the “pupil plane.” The second plane contains the mask and an imaging lens. The spatial variables are  $f_x$  and  $f_y$ ; this plane corresponds to the space of spatial frequencies. The detector is placed in the third plane. This one is conjugated with the first pupil plane. The associated spatial variables are  $x_d$  and  $y_d$  for “detector.”

The incoming electromagnetic (EM) field can be written as

$$\psi_p(x_p, y_p, \lambda) = \sqrt{n(\lambda)} \mathbb{I}_p(x_p, y_p) \exp\left(\frac{2i\pi}{\lambda} \Delta(x_p, y_p)\right),$$

where  $\lambda$  is the wavelength of the incoming light,  $n(\lambda)$  is the number of photons by unit area at the wavelength  $\lambda$ , and  $\mathbb{I}_p$  is the indicator function of the pupil.  $\Delta$  is the optical path difference created by atmospheric turbulence or any other sources of perturbation.

The mask has for its transparency function  $m(f_x, f_y, \lambda)$ . As a complex quantity, it can be decomposed into two terms:

$$m(f_x, f_y, \lambda) = a(f_x, f_y, \lambda) \exp\left(\frac{2i\pi}{\lambda} \text{OS}(f_x, f_y, \lambda)\right).$$

The function  $a$  codes the amplitude filtering. Since the masks are passive,  $a \in [0, 1]$ . The phase term is directly coded by the “optical shape” of the mask. This quantity “OS” depends on the optical indexes (and consequently, on the wavelength of the refractive material) and on the geometric shape of the mask. Via the Fresnel optical formalism, it is possible to write the EM field in the detector plan as

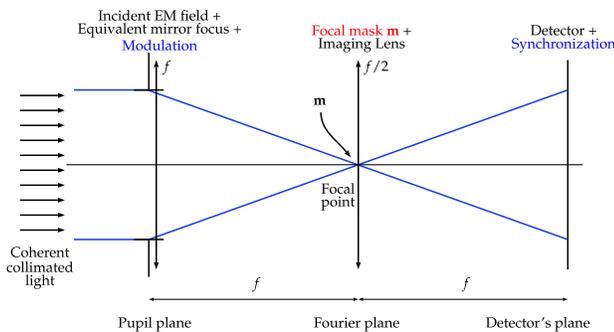


Fig. 1. Schematic view (in 1D) of a Fourier filtering optical system.

$$\psi_d(x_d, y_d, \lambda) \propto \iint \frac{dx_p dy_p}{f^2} \psi_p(x_p - x_d, y_p - y_d, \lambda), \quad (1)$$

$$\iint \frac{df_x df_y}{\lambda^2} m(f_x, f_y, \lambda) \exp\left(-\frac{2i\pi}{f\lambda} [x_p f_x + y_p f_y]\right).$$

For the sake of clarity, we will consider in the first sections that the incoming light is monochromatic. The wavelength will thus be set at  $\lambda_0$ . Section 8 will be dedicated to the polychromatic light case. The incoming EM field and the mask have the following expressions:

$$\psi_p(x_d, y_d) = \sqrt{n} \mathbb{I}_p(x_d, y_d) \exp(i\phi(x_d, y_d)),$$

$$m(f_x, f_y) = a(f_x, f_y) \exp\left(\frac{2i\pi}{\lambda_0} \text{OS}(f_x, f_y)\right), \quad (2)$$

where  $\phi = 2\pi\Delta/\lambda_0$  is the perturbed phase at the considered wavelength. The monochromatic assumption and the associated notations allow us to simplify Eq. (1) into the following form:

$$\psi_d = \psi_p \star \mathcal{F}[m], \quad (3)$$

where  $\mathcal{F}$  is the 2D Fourier transform (2DFT) defined in Appendix A. The interesting point about Eq. (3) is that the contributions of the pupil and the mask are clearly split: the EM field in the detector's plane  $\psi_d$  is the convolution of the incoming field  $\psi_p$  with the Fourier transform of the mask  $\mathcal{F}[m]$ .

Moreover, since the Fourier transform is a bijective application, the quantity  $\mathcal{F}[m]$  completely characterizes the mask. In other words, looking at  $\mathcal{F}[m]$  or at  $m$  itself is mathematically equivalent. Fraunhofer's diffraction also says that the  $\mathcal{F}$  operator allows us to go from a focal plane to a pupil plane. As a consequence,  $\mathcal{F}[m]$  may be seen as the propagated EM field when the diffractive object is the mask  $m$ .

One notes that Eq. (3) is not exact. Indeed, the optical design described above has a magnification equal to  $-1$ . This appears clearly when no mask is inserted in the focal plane. As a consequence, the term  $\psi_d$  should be replaced by its symmetric, i.e.,  $\mathcal{S}[\psi_d]$  where  $\mathcal{S}$  is the symmetric operator defined in Appendix A. However, since this symmetric operation does not impact the wave front sensing, we will not consider it.

The signal that is effectively obtained on the detector is the intensity associated with the EM field in the detector's plane  $\psi_d$ ,

$$I = |\psi_p \star \mathcal{F}[m]|^2. \quad (4)$$

### B. Considered Masks in the Context of Wave Front Sensing

In this part, we will give a general interpretation of the effect of Fourier filtering masks in the context of wave front sensing. The essential objective is to code the phase by using the incoming flux. Subsequently, the mask has to split and extract this information. These operations are done in the focal plane that is the Fourier space associated with the pupil plane. The area close to the optical axis contains the low spatial frequencies of the incoming EM field, whereas the remote zones contain the high spatial frequencies.

These physical results have been widely exploited to remove some spatial frequencies from images. We can, for example, mention the Abbe experiment and the Schlieren photography introduced by Toepler [7]. In our case, the use of opaque masks is not relevant since the reference sources are often faint. We therefore

consider only masks having pure phase transparency functions. Mathematically, this means that the function  $a$  in Eq. (2) equals 1 in the entire focal plane. Having said that, we will now see how it is possible to split and then extract the spatial frequency information thanks to a Fourier mask.

The first step is to choose a tessellation ( $\Omega_i$ ) of the Fourier plane. Mathematically, this means

$$\cup_i \Omega_i = \mathbb{R}^2 \quad \text{and} \quad \Omega_i \cap \Omega_j = \emptyset \quad \text{if } i \neq j.$$

The index  $i$  goes in a finite or a countable set. Each element of this tessellation allows us to select a certain part of the spatial frequencies. The second step consists of separating the spatial frequencies contained in each element  $\Omega_i$  thanks to a tip and a tilt generated via a slope (or an angle) in the optical shape of the mask. If several  $\Omega_i$  have the same rejection angle, it is possible to do interference between them thanks to the local pistons. This corresponds, in terms of optical shape, to local thickness differences. (Obviously, it is possible to add the next shapes (focus, astigmatism, etc.) by playing with the optical shape of each  $\Omega_i$ , but we choose to not explore this level of complexity in this paper.) Mathematically, every considered mask can thus be written as

$$m(f_x, f_y) = \sum_i \mathbb{I}_{\Omega_i}(f_x, f_y) \exp\left(\frac{2i\pi}{\lambda_0}(\delta_i + f_x \alpha_i + f_y \beta_i)\right), \quad (5)$$

where  $\delta_i$  codes the local piston, and  $(\alpha_i, \beta_i)$  the local tip/tilt corresponding to the rejection angles.  $\mathbb{I}_{\Omega_i}$  is the indicator function of  $\Omega_i$ . To summarize, it is possible to describe any kind of mask used in the context of wave front sensing via the following set of parameters:

$$\{\Omega_i, \delta_i, \alpha_i, \beta_i\}_i. \quad (6)$$

This set will be called the tessellation parameters of the considered mask. With this formalism, it is now possible to write the 2DFT corresponding to the expression of the general transparency function [Eq. (5)],

$$\mathcal{F}[m](x_d, y_d) = \sum_i \exp\left(\frac{2i\pi\delta_i}{\lambda_0}\right) \mathcal{F}[\mathbb{I}_{\Omega_i}](x_d - f\alpha_i, y_d - f\beta_i), \quad (7)$$

where  $f$  is the focal point of the imaging lens.

### C. 2D Fourier Transform of Classical Tessellations

Although all the tessellations of the Fourier plan are *a priori* acceptable, two of them play a significant role for the existing WFSs.

The first one consists of splitting the plane into 4 elements following the Cartesian coordinate system (left insert of Fig. 2). The second one is a polar splitting of the Fourier plane. This tessellation has a parameter  $\rho$  that codes the size of the central circle (right insert of Fig. 2).

From these definitions, one can write the indicator functions of each part of the tessellation and then get their 2DFTs. We detail this indicator function for the element  $\Omega^{++}$ ,

$$\mathbb{I}_{\Omega^{++}}(f_x, f_y) = \Theta(f_x)\Theta(f_y),$$

where  $\Theta$  is the Heaviside function. Hence, its 2DFT equals

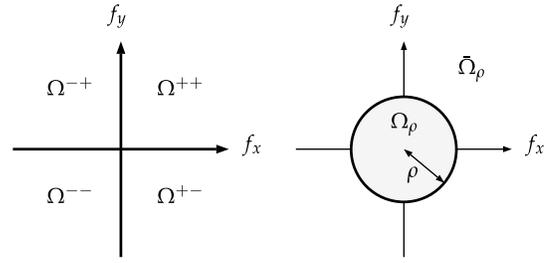


Fig. 2. Two classical tessellations. Cartesian splitting (left insert) and polar splitting (right insert).

$$\begin{aligned} \mathcal{F}[\mathbb{I}_{\Omega^{++}}](x_d, y_d) &= \frac{1}{4} \left( \delta(x_d)\delta(y_d) - \frac{1}{\pi^2 x_d y_d} \right) \\ &\quad - \frac{1}{4} \left( \frac{\delta(x_d)}{\pi y_d} + \frac{\delta(y_d)}{\pi x_d} \right). \end{aligned} \quad (8)$$

For the other parts of the Cartesian tessellation, we get

$$\begin{aligned} \mathcal{F}[\mathbb{I}_{\Omega^{+-}}](x_d, y_d) &= \frac{1}{4} \left( \delta(x_d)\delta(y_d) + \frac{1}{\pi^2 x_d y_d} \right) \\ &\quad - \frac{1}{4} \left( \frac{\delta(x_d)}{\pi y_d} - \frac{\delta(y_d)}{\pi x_d} \right), \end{aligned} \quad (9)$$

$$\begin{aligned} \mathcal{F}[\mathbb{I}_{\Omega^{-+}}](x_d, y_d) &= \frac{1}{4} \left( \delta(x_d)\delta(y_d) - \frac{1}{\pi^2 x_d y_d} \right) \\ &\quad + \frac{1}{4} \left( \frac{\delta(x_d)}{\pi y_d} + \frac{\delta(y_d)}{\pi x_d} \right), \end{aligned} \quad (10)$$

$$\begin{aligned} \mathcal{F}[\mathbb{I}_{\Omega^{--}}](x_d, y_d) &= \frac{1}{4} \left( \delta(x_d)\delta(y_d) + \frac{1}{\pi^2 x_d y_d} \right) \\ &\quad + \frac{1}{4} \left( \frac{\delta(x_d)}{\pi y_d} - \frac{\delta(y_d)}{\pi x_d} \right). \end{aligned} \quad (11)$$

For the polar tessellation, the indicator function  $\mathbb{I}_{\Omega_\rho}(f_r, f_\theta)$  equals  $\Theta(\rho - f_r)$ . Its 2DFT is

$$\mathcal{F}[\mathbb{I}_{\Omega_\rho}](r_d, \theta_d) = \frac{\rho}{r_d} J_1(2\pi\rho r_d), \quad (12)$$

where  $J_\alpha$  are the first-kind Bessel functions.

### D. Other WFSs

We mention here that other tessellations may be relevant in the wave front sensing context, especially regarding the case of the point diffraction interferometer introduced by Smart and Steel [8]. Moreover, this formalism easily extends to the optical differentiation WFS introduced by Oti *et al.* [9] and also describes the first stage of coronagraphic systems.

## 3. META-INTENSITY DEFINITIONS

In this section, we introduce the basic elements needed to construct, from the intensity on the detector, a numerical quantity called the meta-intensity and written  $mI$ , which consists of the linear response to an incoming turbulent phase.

**A. Reference Phase**

The phase seen by the WFS is the sum of the turbulent phase induced by the atmosphere and the static aberrations of the wave front sensing path. Mathematically, we can split the incoming phase into two terms,

$$\phi = \phi_r + \phi_t, \tag{13}$$

where  $\phi_t$  is the turbulent phase, and  $\phi_r$  the static reference phase.  $\phi_r$  may also be seen as the operating point of the WFS.

**B. Toward the Linearity**

The goal of the meta-intensity is to transform the intensity on the detector into a quantity that will be linear with the turbulent phase around the reference phase. Moreover,  $mI$  has also to be independent of the flux  $n$ . Mathematically, such conditions are written as

$$mI(\phi_t + a\Phi_t) = mI(\phi_t) + amI(\Phi_t) \\ \forall \phi_t, \Phi_t \in \text{Phase space and } \forall a \in \mathbb{R}.$$

We make use of the power series of exponential functions and the Cauchy product laws to develop the squared module of Eq. (4) and get an expression of the intensity depending on the successive powers of the turbulent phase:

$$I(\phi, n) = I(\phi_r + \phi_t, n) = n \sum_{q=0}^{\infty} \frac{(-1)^q}{q!} \sum_{k=0}^q (-1)^k \binom{q}{k} \phi_t^{k*} \overline{\phi_t^{q-k*}} \\ \text{where } \phi_t^{k*} \triangleq (\mathbb{I}_p e^{i\phi_r} \star \mathcal{F}[m])^k. \tag{14}$$

The complex quantity  $\phi_t^{k*}$  may be considered as a  $k$ th-moment of the turbulent phase through the mask around the static reference phase  $\phi_r$ . By looking at Eq. (14), we note that one way to make  $I$  independent on the flux is to divide it by the spatial average incoming flux  $n$ . This operation is easy in practice, since  $n$  is proportional to the total flux on the detector. The first step of the post-processing to build the meta-intensity is thus to normalize the intensity by the spatial average flux  $n$ .

Regarding the linearity with respect to the phase, one can explicitly develop Eq. (14). This is done below for  $q = 0$  to 2:

Constant term,  $q = 0$ :

$$I_c \triangleq |\mathbb{I}_p e^{i\phi_r} \star \mathcal{F}[m]|^2. \tag{15}$$

Linear term,  $q = 1$ :

$$I_l(\phi_t) \triangleq 2\Re[(\mathbb{I}_p e^{i\phi_r} \star \mathcal{F}[m]) \overline{(\mathbb{I}_p e^{i\phi_r} \phi_t \star \mathcal{F}[m])}]. \tag{16}$$

Quadratic term,  $q = 2$ :

$$I_q(\phi_t) \triangleq |\mathbb{I}_p e^{i\phi_r} \phi_t \star \mathcal{F}[m]|^2 \\ - \Re[(\mathbb{I}_p e^{i\phi_r} \star \mathcal{F}[m]) \overline{(\mathbb{I}_p e^{i\phi_r} \phi_t^2 \star \mathcal{F}[m])}]. \tag{17}$$

The only term linear with the turbulent phase  $\phi_t$  is the  $q = 1$  term, which we call the linear intensity,  $I_l$ . Ideally, and in order to maximize the linearity, one would want to minimize the other terms, i.e.,  $q = 0$  and  $q \geq 2$ . The constant intensity  $I_c$  ( $q = 0$ ) corresponds to the normalized intensity on the detector when the phase equals the reference phase, i.e.,  $I(\phi_r, n)/n$ . Removing such a term can thus be done thanks to a calibration path. With regard to the  $q \geq 2$  terms, they are unfortunately impossible to remove, but we still consider that the second step

to build  $mI$  consists of this return-to-reference. Mathematically, the meta-intensity  $mI$  is thus defined as

$$mI(\phi_t) = \frac{I(\phi_r + \phi_t, n) - I(\phi_r, n)}{n}. \tag{18}$$

This definition is the easiest way to define a linear quantity for any type of sensor. Finally, it might be relevant to mention that it avoids a lot of other processes that can be applied to the meta-intensities after these two fundamental first steps (normalization, return-to-reference), but it is a vast topic that will be tackled in a forthcoming paper (see [10]).

**C. Linear and Effective Meta-Intensities**

The previous post-processing allows us to build, in practice, the meta-intensity from the intensity on the detector and a reference intensity. Unfortunately, we saw that Eq. (18) is not perfectly linear with the turbulent phase, since  $mI$  still contains the quadratic and next high-order moments of the phase. In the following, we will distinguish the effective behavior of a WFS from its ideal one. Equation (18) will define the effective meta-intensity, while the restriction of the phase power series development to its first term, i.e., to the linear intensity  $I_l$ , will define the linear meta-intensity [see Eq. (16)]. We note that this linear meta-intensity cannot be strictly deduced from the intensity on the detector, but it can be easily computed by numerical simulations. It corresponds to the behavior a perfectly linear WFS would have.

As a side comment, another interpretation of the linear and quadratic meta-intensities can be deduced from the derivatives of the normalized intensity with respect to the amplitude phase around the reference phase, i.e.,

$$I_l(\phi_t) = \frac{1}{n} \left. \frac{dI(\phi_r + a\phi_t, n)}{da} \right|_{a=0} \\ I_q(\phi_t) = \frac{1}{2n} \left. \frac{d^2 I(\phi_r + a\phi_t, n)}{da^2} \right|_{a=0}.$$

This result is not surprising regarding the calculation of Eq. (14), which is the Taylor's development of the intensity around  $\phi_r$ .

**4. SENSITIVITY AND LINEARITY RANGE**

In this section, we define the sensitivity and the linearity range of a WFS regarding an incoming turbulent phase. From now, this turbulent phase will be normalized regarding the 2-norm (i.e., the RMS norm).  $\phi_t$  is used in the light of the previous developments.

Due to the fact that the sensitivity only makes sense in the linear regime of the WFS, we will use the linear meta-intensity  $I_l$  to define it. On the other hand, the linearity range will be calculated thanks to the study of the variation of the distance between the effective and the linear meta-intensities when the input phase amplitude varies.

**A. Sensitivity**

The perfectly linear response to the normalized incoming phase  $\phi_t$  is obtained thanks to the expression of the linear intensity [Eq. (16)]  $I_l(\phi_t)$ .

We choose to define as “the sensitivity regarding to  $\phi_t$ ” the 2-norm of this linear response:

$$s(\phi_t) = \|I_l(\phi_t)\|_2. \tag{19}$$

Such a definition is in fact consistent with the method introduced by Rigaut and Gendron [11] to estimate the noise propagation

associated with a WFS. The Rigaut and Gendron method consists of looking at the diagonal terms of the matrix  $({}^t\mathcal{M}\mathcal{C}_B^{-1}\mathcal{M})^{-1}$ , where  $\mathcal{M}$  is the interaction matrix, i.e., the matrix that contains the output meta-intensities for an input phase basis, and  $\mathcal{C}_B$  is the noise covariance matrix. Physically, these values correspond to the noise propagation coefficients for a detector or photon noise regime, depending the nature of  $\mathcal{C}_B$ .

In the case of a diagonal and constant noise covariance matrix, i.e.,  $\mathcal{C}_B = \sigma_{\text{noise}}^2 \times Id$ , the noise propagation coefficients for the turbulent phase  $\phi_t$ , called  $\sigma_{\text{WFS}}^2(\phi_t)$  in Rigaut and Gendron are related to the sensitivity definition introduced above  $s(\phi_t)$  via the following equation:

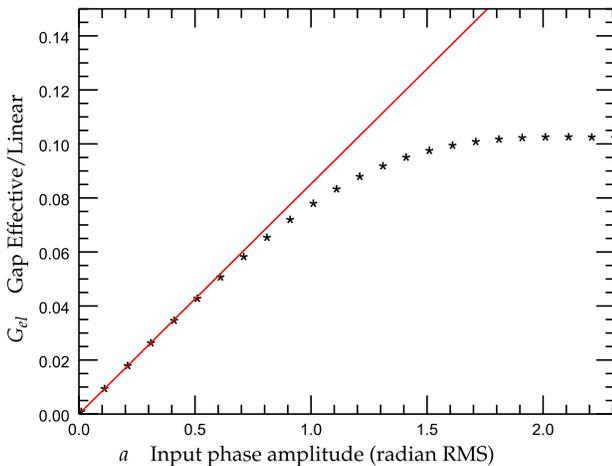
$$\sigma_{\text{WFS}}^2(\phi_t) = \sigma_{\text{noise}}^2 s(\phi_t)^{-2}.$$

## B. Linearity Range

The linearity range quantifies the gap between the perfectly linear behavior of a WFS and its effective behavior. Indeed, when the incoming phase amplitude increases, the meta-intensity defined in Eq. (18) differs more and more from the linear intensity [Eq. (16)]. We call the mathematical function that quantifies this deviation from linearity  $G_{el}$  for “gap between effective and linear behaviors,” and we define it by

$$G_{el}(\phi_t, a) = \left\| \frac{mI(a\phi_t)}{a} - I_l(\phi_t) \right\|_2, \quad (20)$$

where the variable  $a$  codes the amplitude of the incoming phase. We note that  $\phi_t$  is, once again, normalized regarding the RMS norm. The black curve of Fig. 3 shows the typical evolution of  $G_{el}$  when  $a$  increases. Two distinct regimes are observed: a linear growth for the lowest amplitude and, then, a saturation regime. In the saturation regime, the intensity on the detector does not change anymore as the phase amplitude continues to grow. The linear regime can be explained by looking at the analytic expression of  $G_{el}$ ,



**Fig. 3.** Typical graph (\*) of the gap between the effective meta-intensity and the linear intensity depending on  $a$ , and the input phase amplitude [Eq. (20)]. The slope of the red line equals the 2-norm of the quadratic intensity. The input phase is the vertical coma, and the considered WFS is the Zernike WFS.

$$\begin{aligned} \frac{mI(a\phi_t)}{a} - I_l(\phi_t) &= \sum_{q=2}^{\infty} a^{q-1} \frac{(-1)^q}{q!} \sum_{k=0}^q (-1)^k \binom{q}{k} \phi_t^{k*} \overline{\phi_t^{q-k*}} \\ &= aI_q(\phi_t) + a^2(\dots). \end{aligned}$$

The linear growth is thus directly linked to the quadratic intensity; more precisely, the associated slope equals the 2-norm of  $I_q$  (see the red curve of Fig. 3):

$$G_{el}(\phi_t, a) = a\|I_q(\phi_t)\|_2 + a^2(\dots).$$

In others words, a small slope means that the effective meta-intensities slowly differ from the linear behavior for an increasing incoming phase amplitude; the associated linear range is thus large.

We thus choose to define the linearity range regarding the normalized phase mode  $\phi_t$  as the inverse of the 2-norm of the quadratic term,

$$d(\phi_t) = (\|I_q(\phi_t)\|_2)^{-1}, \quad (21)$$

where the letter  $d$  refers to the word “dynamic,” which is a synonym of “linearity range.”

Such a definition allows us to calculate the linear range in an analytical way thanks to the expression of the quadratic term [Eq. (17)].

## C. SD Factor

The previous developments allow us to give an explanation about the fact that the sensitivity and the linearity range are antagonistic quantities. By looking at their product, which we call the SD factor,

$$s(\phi_t)d(\phi_t) = (\|I_l(\phi_t)\|_2)(\|I_q(\phi_t)\|_2)^{-1}, \quad (22)$$

we can see that such a quantity corresponds to the ratio between the norms of the first and the second derivatives of the intensity regarding the turbulent phase. One can easily understand that it is difficult to increase the numerator while decreasing the denominator at the same time. The SD factor is thus a relevant indicator of the trade-off between the sensitivity and the linearity range.

The expressions of the linear and quadratic intensities [Eqs. (16) and (17)] and the 2-norm definition allow us to numerically optimize a WFS, i.e., its mask, by maximizing its SD factor. For a given turbulent phase  $\phi_t$ , we can thus consider a mask depending on a scalar optical parameter that we call  $p$ . The only constraint on this parameter is to define a mask that is transparent, i.e.,  $|m(f_x, f_y)| = 1$  (or at least a passive one:  $|m(f_x, f_y)| \leq 1$ ). As examples, the optical parameters of the tessellation formalism ( $\delta, \alpha, \beta$ ) are appropriate since they are variables for the geometrical shape of the mask. The next step just consists of finding numerically the maximum of the function  $s \cdot d(\phi_t, m(p))$  regarding the parameter  $p$ .

Finally, one can modify the SD factor to give more or less importance to the sensitivity or the linearity range by introducing a power exponent  $\eta$  in it,

$$s^\eta d^{1/\eta} = (\|I_l(\phi_t)\|_2)^\eta \cdot (\|I_q(\phi_t)\|_2)^{-1/\eta},$$

with  $\eta \in \mathbb{R}_+^*$ . For instance,  $\eta = 0.5$  confers more importance to the linearity range than to the sensitivity, whereas  $\eta = 2$  represents the opposite case. The power exponent  $\eta$  allows one to consider different requirement specifications.

## 5. MODULATION

The Fourier filtering may be coupled with an additional optical stage placed in the entrance pupil plane. Such a device creates an

oscillating aberration and changes the shape and the size of the focal spot on the Fourier mask. The detector is synchronized with this regular oscillation in order to have one image for each aberration cycle. This system, called modulation, allows us to adjust the WFS's performance.

Historically, such a device has been introduced by Ragazzoni [1] in order to improve the linearity range of the classical PWFS. Even if, to our knowledge, modulation is only used for the PWFS, we give in this section a definition for any kind of WFSs and show how to define a generalized sensitivity and linearity range with such a device.

### A. General Definition of Modulation

The oscillating aberration introduced by the modulation device defines a closed path (or loop) in phase space. Mathematically, it means that modulation may be defined in the following way:

$$\phi_m(s) = \sum_{k=0}^{\infty} m_k(s)\Phi_k \quad \text{with} \quad \forall k \quad m_k(0) = m_k(1)$$

$$w(s) \quad \text{with} \quad w(s) \geq 0, w(0) = w(1), \int_0^1 w(s)ds = 1,$$

where  $s$  is the temporal variable normalized with respect to the duration of the modulation cycle  $\tau$ , i.e.,  $s = t/\tau$ . The phase polynomials  $\Phi_k$  describe a phase basis, typically the Zernike basis. The functions  $m_k(s)$  indicate the amplitudes of the phase modes used during modulation.  $w(s)$  is the weighting function: it codes the time spent for each modulation phase  $\phi_m(s)$ .

### B. Intensity on the Detector

Thanks to these definitions, it is possible to write the modulated intensity, called  $I_m$ , on the detector. In particular, this one is the integral during a cycle of the intensity with an additional phase corresponding to the local modulation phase,

$$I_m(\phi, n) = \int_0^1 I(\phi + \phi_m(s), n \cdot w(s))ds.$$

Such an equation is linked to the fact that the modulation handles the light as an incoherent quantity. Moreover, remembering Eq. (13), it appears that the reference phase may be seen as a static modulation. In the following, we will assume that the modulation phase contains this reference phase. The generalized phase power series of the intensity becomes

$$\frac{I_m(\phi, n)}{n} = \sum_{q=0}^{\infty} \frac{(-1)^q}{q!} \sum_{k=0}^q (-1)^k \binom{q}{k} \int_0^1 w(s)ds \phi_i^{k*}(s) \overline{\phi_i^{q-k*}(s)}$$

where  $\phi_i^{k*}(s) = (\mathbb{I}_p e^{i\phi_m(s)} \phi_i^k) \star \mathcal{F}[m]$ .

### C. Modulated Meta-Intensity

The first terms of the previous equation correspond once again to the modulated constant, linear, and quadratic terms:

$$I_{mc} = \int_0^1 |\mathbb{I}_p e^{i\phi_m(s)} \star \mathcal{F}[m]|^2 w(s)ds$$

$$I_{mi}(\phi_t) = \int_0^1 2\Im[(\mathbb{I}_p e^{i\phi_m(s)} \star \mathcal{F}[m]) \overline{(\mathbb{I}_p e^{i\phi_m(s)} \phi_t \star \mathcal{F}[m])}] w(s)ds$$

$$I_{mq}(\phi_t) = \int_0^1 [|\mathbb{I}_p e^{i\phi_m(s)} \phi_t \star \mathcal{F}[m]|^2] w(s)ds$$

$$- \int_0^1 \left[ \Re[(\mathbb{I}_p e^{i\phi_m(s)} \star \mathcal{F}[m]) \overline{(\mathbb{I}_p e^{i\phi_m(s)} \phi_t^2 \star \mathcal{F}[m])}] \right] w(s)ds.$$

The method used previously to construct a phase-linear quantity (at least in the small-phase approximation) is thus still valid. The effective modulated meta-intensity, called  $mI_m$ , will be subsequently defined in the same way as Eq. (18), i.e.,

$$mI_m(\phi_t) = \frac{I_m(\phi_t, n) - I_m(0, n)}{n}. \tag{23}$$

The definitions of the sensitivity and the dynamic stay as they were in Eqs. (19) and (21) under the condition they must now use the modulated linear and quadratic intensities,  $I_{mi}$  and  $I_{mq}$ .

## 6. APPLICATION TO THE PYRAMID WFS

In this section, we apply the theoretical formalism developed above to the pyramid WFSs. Without loss of generality, we assume in the next developments that the operating phase equals the null phase. Ragazzoni [1] introduced this sensor, which is a generalization of Foucault's knife. The corresponding Fourier mask placed in the focal plane is a transparent squared pyramid. Its apex angle is called  $\theta$ . In its historical configuration, this mask creates in the detector's plane 4 pupil images, each containing only one part of the spatial frequency.

### A. Tessellation Formalism

The pyramidal mask allows us to illustrate the spatial frequency separation induced by local tip/tilt OPDs. The tessellation parameters (i.e.,  $\Omega_i; \delta_i, \alpha_i, \beta_i$ ), associated with the transparency function of this mask that we note as  $m_{\Delta}$  are

$$\Omega^{-+}:0, -\theta, \theta \quad \Omega^{++}:0, \theta, \theta$$

$$\Omega^{--}:0, -\theta, -\theta \quad \Omega^{+-}:0, \theta, -\theta.$$

The associated 2DFT is

$$\mathcal{F}[m_{\Delta}](x_d, y_d) = \mathcal{F}[\mathbb{I}_{\Omega^{++}}](x_d - f\theta, y_d - f\theta)$$

$$+ \mathcal{F}[\mathbb{I}_{\Omega^{-}}](x_d + f\theta, y_d + f\theta)$$

$$+ \mathcal{F}[\mathbb{I}_{\Omega^{+}}](x_d + f\theta, y_d - f\theta)$$

$$+ \mathcal{F}[\mathbb{I}_{\Omega^{+-}}](x_d - f\theta, y_d + f\theta). \tag{24}$$

Considering the 2DFT of each quadrant  $\Omega_i$  [see Eqs. (8)–(11)], we have with Eq. (24) an analytic formulation of the 2DFT of the pyramid WFS mask. We can see in Fig. 4 the module and the argument of this quantity [Eq. (24)]. Such quantities allows us to determine the imaging on the detector (top right insert of Fig. 5), since this intensity corresponds to the convolution between the entrance pupil and the 2DFT of the mask. It may be interesting to study the different cases of this optical design with respect to the apex angle parameter:

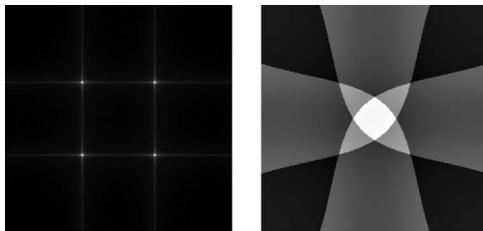
– The first one is when  $\theta$  tends to infinity. It corresponds to the case of a reflective pyramid that creates 4 pupil images on

4 different detectors. This design was introduced by Wang *et al.* [12]. It allows one to completely separate the 4 pupil images. Concretely, it means the four 2DFTs are considered as independent: there is no crosstalk between the spatial frequencies of the 4 quadrants  $\Omega_i$ . This hypothesis of no interference between the pupils is usually done in theoretical approaches and in the usual simulation algorithms.

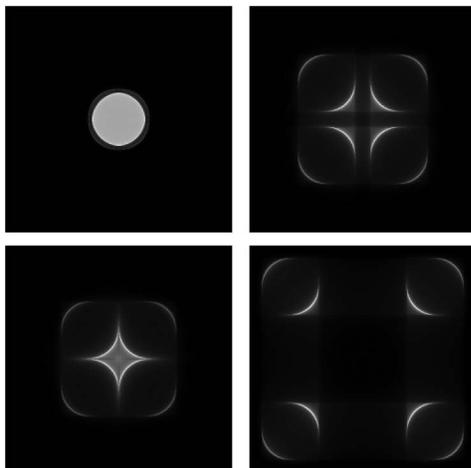
– The second case is the more common one. It corresponds to an achromatic refractive pyramid. The 4 pupil images are created on a unique detector and  $D/2 \leq f\theta$ , where  $D$  is the diameter of the entrance pupil and  $f$  is the focal point of the imaging lens. It means there is no overlap between these 4 images, even if each pupil image contains a part of the spatial frequencies of the other quadrants due to the fact that the 2DFTs of the four  $\Omega_i$  do not have a compact support. This interference between the 4 pupil images is particularly marked in the middle area (see top right insert and bottom inserts of Fig. 5).

– We introduced the last case in [4] with the flattened pyramid WFS. The idea of this sensor is to overlap the 4 pupil images by using a small angle:  $0 < \theta f < D/2$ . (The case  $\theta = 0$  is obviously useless since it corresponds to the trivial mask.) This optical configuration is shown in the top left insert of Fig. 5.

Before getting into more details on these WFSs, it would be relevant to now consider the pyramid not as a unique and “static” optical system, but more like a class, including an infinite number of different pyramids. Indeed, the angle of the apex  $\theta$  is one optical parameter, and it is possible to envisage another ones as, for example, the number of faces or the modulation parameters.



**Fig. 4.** Modulus (left) and argument (right) of the 2DFT with an apex angle that equals  $\frac{1.5D}{2f}$ , where  $D$  is the entrance pupil diameter.



**Fig. 5.** Intensity on the detector for a circular pupil and a flat incoming phase.  $\theta$  equals 0.1, 1, 1.5, and  $3\frac{D}{2f}$ .

## B. Sensitivity, Linearity Range, and SD Factor

### 1. Apex Angle

Figure 6 shows the sensitivity, the linearity range, and the SD factor (red, black, and blue curves) with respect to the first 24 Zernike radial orders, i.e., the spatial frequencies, for different apex angles. Physically, the optical parameter  $\theta$  sets the overlap rate of the pupil images. First of all, one can note that as soon as  $2f\theta/D > 1.5$ , the 3 curves ( $s$ ,  $d$ , and  $s \cdot d$ ) do not evolve anymore, and this is a general behavior observed for all the configurations tested in this paper. In other words,  $\theta$  has an influence only when the pupil images overlap.

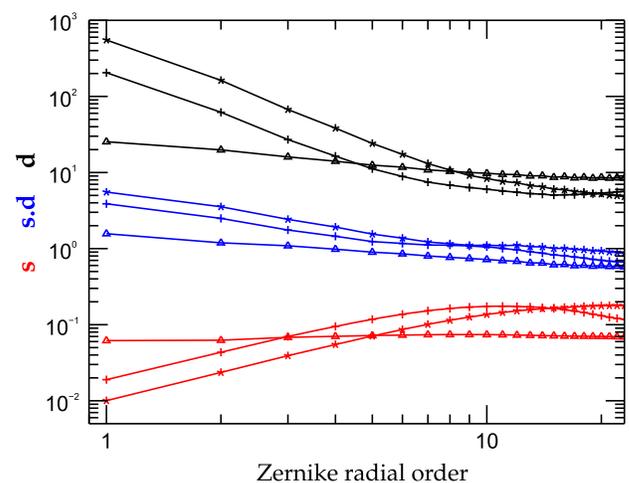
Second, as mentioned in [4], Fig. 6 shows that an optical recombination induced by a small angle provides a better sensitivity in high spatial frequencies, while it decreases for low frequencies. Moreover, it is possible to choose where the sensitivity is maximum by changing the  $\theta$  value. It comes as no surprise that the linearity range has an inverse behavior: it improves for the low frequencies and decreases at the high ones. The curve of the SD factor is more interesting, since for small angles, this curve stays above the classical pyramid one for all spatial frequencies. In terms of the trade-off between the sensitivity and the linearity range, there is thus a real gain to using small angles.

### 2. Modulation Radius

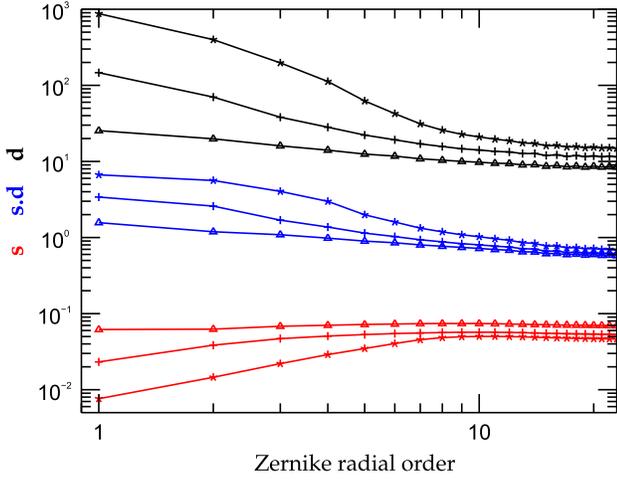
The pyramid WFS is usually coupled with a modulation stage. With the formalism introduced in section 5, the classical modulation used with Ragazzoni’s pyramid, i.e., a circular and uniform tip/tilt modulation with a modulation radius equal to  $r_m$ , corresponds to

$$\phi_m(s) = r_m[\cos(2\pi s)Z_1^{-1} + \sin(2\pi s)Z_1^1] \quad \text{and} \quad w(s) = 1.$$

We consider here that the apex angle  $\theta$  equals  $3\frac{2\theta f}{D}$ , and the pupil images are thus widely separated (right insert of Fig. 5). Figure 7 shows the sensitivity, the linearity range, and the SD factor with respect to the first 24 Zernike radial orders for different modulation radii. One can observe the classical influence of the modulation radius on the sensitivity and the linearity range: there is a loss a sensitivity for the low spatial frequencies with a slope sensor



**Fig. 6.** Sensitivity, linearity range, and SD factor of the pyramid WFS with respect to the spatial frequencies. The apex angle equals 0.05 (\*), 0.1 (+), and  $3\frac{2\theta f}{D}$  ( $\Delta$ ). The phase basis corresponds to the 24 first Zernike radial orders.



**Fig. 7.** Sensitivity, linearity range, and SD factor of the modulated pyramid WFS with respect to the spatial frequencies. The modulation radii equal 0 ( $\Delta$ ), 1 (+), and 3 ( $*$ )  $\lambda/D$ . The apex angle equals  $3\frac{2\theta_f}{D}$ , i.e., the 4 pupil images are widely separated.

behavior on this range. The cut-off frequency is growing linearly with the modulation radius. The linearity range is also improving with the modulation radius. We thus observe that our definition of the sensitivity and the linearity range allows us to get, in an analytical way, the usual behaviors of the modulated pyramid WFS. By looking at the SD factor, which characterizes the trade-off sensitivity/linearity range, it appears that the gain is particularly pronounced for low spatial frequencies but tends to be null for the highest ones.

### C. Mathematical Developments Applied to Ragazzoni's PWFS

The aim of this subsection is to apply the analytical developments to the nonmodulated Pyramid WFS. In particular, we show how our formalism may be linked to the classical meta-intensities used for the PWFS, which are usually called “slope maps” [1,13].

We assume that the apex angle tends to infinity. As a consequence, the 4 pupil images are completely separated and do not interfere. Such a theoretical framework allows us to study the 4 quadrants independently. Subsequently, each pupil image has its own intensity:

$$\begin{aligned} I^{-+} &= |\psi_p \star \mathcal{F}[\mathbb{I}_{\Omega^+}]|^2 & I^{++} &= |\psi_p \star \mathcal{F}[\mathbb{I}_{\Omega^{++}}]|^2 \\ I^{--} &= |\psi_p \star \mathcal{F}[\mathbb{I}_{\Omega^-}]|^2 & I^{+-} &= |\psi_p \star \mathcal{F}[\mathbb{I}_{\Omega^+}]|^2. \end{aligned}$$

The first step consists of getting the linear and quadratic intensities associated with the general definition of the meta-intensities [see Eq. (18)]. Thanks to Eq. (16), it is possible to determine the phase-linear dependence of the meta-intensity associated with one of the 4 pupil images. Let us take the case of the pupil image containing the  $x$ -positive and  $y$ -positive spatial frequencies, i.e., the  $\Omega^{++}$  quadrant. The linear intensity is

$$I_l^{++}(\phi_t) = 2\Im[(\mathbb{I}_p \star \text{CIR}[\mathbb{I}_{\Omega^{++}}]) \overline{(\mathbb{I}_p \phi_t \star \text{CIR}[\mathbb{I}_{\Omega^{++}}])}].$$

The development of this equation gives

$$\begin{aligned} I_l^{++}(\phi_t) &= \frac{1}{8}[(\mathcal{I} + \mathcal{H}_{xy}^2)[\mathbb{I}_p](\mathcal{H}_x + \mathcal{H}_y)[\mathbb{I}_p \phi_t]] \\ &\quad - \frac{1}{8}[(\mathcal{I} + \mathcal{H}_{xy}^2)[\mathbb{I}_p \phi_t](\mathcal{H}_x + \mathcal{H}_y)[\mathbb{I}_p]], \end{aligned}$$

where the operators  $\mathcal{I}$ ,  $\mathcal{H}_x$ ,  $\mathcal{H}_y$ , and  $\mathcal{H}_{xy}^2$  are defined in Appendix A. The linear intensities associated with the other quadrants are

$$\begin{aligned} I_l^{+-}(\phi_t) &= \frac{1}{8}[(\mathcal{I} - \mathcal{H}_{xy}^2)[\mathbb{I}_p](\mathcal{H}_x - \mathcal{H}_y)[\mathbb{I}_p \phi_t]] \\ &\quad - \frac{1}{8}[(\mathcal{I} - \mathcal{H}_{xy}^2)[\mathbb{I}_p \phi_t](\mathcal{H}_x - \mathcal{H}_y)[\mathbb{I}_p]] \\ I_l^{-+}(\phi_t) &= \frac{1}{8}[(\mathcal{I} - \mathcal{H}_{xy}^2)[\mathbb{I}_p](\mathcal{H}_x + \mathcal{H}_y)[\mathbb{I}_p \phi_t]] \\ &\quad - \frac{1}{8}[(\mathcal{I} - \mathcal{H}_{xy}^2)[\mathbb{I}_p \phi_t](\mathcal{H}_x + \mathcal{H}_y)[\mathbb{I}_p]] \\ I_l^{--}(\phi_t) &= \frac{1}{8}[(\mathcal{I} + \mathcal{H}_{xy}^2)[\mathbb{I}_p](\mathcal{H}_x - \mathcal{H}_y)[\mathbb{I}_p \phi_t]] \\ &\quad - \frac{1}{8}[(\mathcal{I} + \mathcal{H}_{xy}^2)[\mathbb{I}_p \phi_t](\mathcal{H}_x - \mathcal{H}_y)[\mathbb{I}_p]]. \end{aligned}$$

These equations allow us to have an analytic expression of the sensibility associated with the turbulent phase mode  $\phi_t$ ,

$$\begin{aligned} s(\phi_t) &= \|I_l^{+-}(\phi_t)\|_2 + \|I_l^{-+}(\phi_t)\|_2 \\ &\quad + \|I_l^{--}(\phi_t)\|_2 + \|I_l^{++}(\phi_t)\|_2. \end{aligned}$$

Concerning the quadratic intensities range, we choose not to explicitly show their expressions due to the fact that they may seem “abstruse,” but, obviously, it is possible to get them thanks to Eq. (17). We deduce from them the linearity range:

$$\begin{aligned} d(\phi_t) &= (\|I_q^{+-}(\phi_t)\|_2 + \|I_q^{-+}(\phi_t)\|_2 \\ &\quad + \|I_q^{--}(\phi_t)\|_2 + \|I_q^{++}(\phi_t)\|_2)^{-1}. \end{aligned}$$

The general meta-intensities we introduced above [Eq. (18)] do not correspond to the classical processing done on the intensity in order to create a quantity linear with the incoming turbulent phase. Indeed, it is the customary to use the slope maps  $S^x$  and  $S^y$ . These quantities are directly calculated from the 4 intensities on the detector. There are two ways to define them (Ragazzoni [1] and V erinaud [13]), depending on the normalization, but we are interested here in the V erinaud's one:

$$S^x = \frac{I^{++} + I^{+-} - I^{-+} - I^{--}}{n}, \quad (25)$$

$$S^y = \frac{I^{++} - I^{-+} + I^{+-} - I^{--}}{n}. \quad (26)$$

One can note that this definition does not require a return-to-reference. However, it is only valid if the reference phase is the null phase. In that particular case, the slope maps are related to our meta-intensities via the linear transformations:

$$S^x = mI^{++} + mI^{+-} - mI^{-+} - mI^{--}, \quad (27)$$

$$S^y = mI^{++} - mI^{-+} + mI^{+-} - mI^{--}. \quad (28)$$

Subsequently, it is possible to get the linear and quadratic slope maps:

$$S_l^x = I_l^{++} + I_l^{+-} - I_l^{-+} - I_l^{--} \quad S_l^y = I_l^{++} - I_l^{-+} + I_l^{+-} - I_l^{--}$$

$$S_q^x = I_q^{++} + I_q^{+-} - I_q^{-+} - I_q^{--} \quad S_q^y = I_q^{++} - I_q^{-+} + I_q^{+-} - I_q^{--}$$

It is thus possible to have the analytic expression of the linear slope maps:

$$S_l^x(\phi_t) = \frac{1}{2} (\mathcal{I}[\mathbb{I}_P] \mathcal{H}_x[\mathbb{I}_P \phi_t] + \mathcal{H}_{xy}^2[\mathbb{I}_P] \mathcal{H}_y[\mathbb{I}_P \phi_t])$$

$$- \frac{1}{2} (\mathcal{I}[\mathbb{I}_P \phi_t] \mathcal{H}_x[\mathbb{I}_P] - \mathcal{H}_{xy}^2[\mathbb{I}_P \phi_t] \mathcal{H}_y[\mathbb{I}_P])$$

$$S_l^y(\phi_t) = \frac{1}{2} (\mathcal{I}[\mathbb{I}_P] \mathcal{H}_y[\mathbb{I}_P \phi_t] + \mathcal{H}_{xy}^2[\mathbb{I}_P] \mathcal{H}_x[\mathbb{I}_P \phi_t])$$

$$- \frac{1}{2} (\mathcal{I}[\mathbb{I}_P \phi_t] \mathcal{H}_y[\mathbb{I}_P] - \mathcal{H}_{xy}^2[\mathbb{I}_P \phi_t] \mathcal{H}_x[\mathbb{I}_P]).$$

These results already appeared in Shatokhina *et al.* [14]. They correspond to the 2D generalization of Vérinaud's calculation in [13] and allow us to derive the sensitivity associated with the slope maps, which we call  $s_S(\phi_t)$ ,

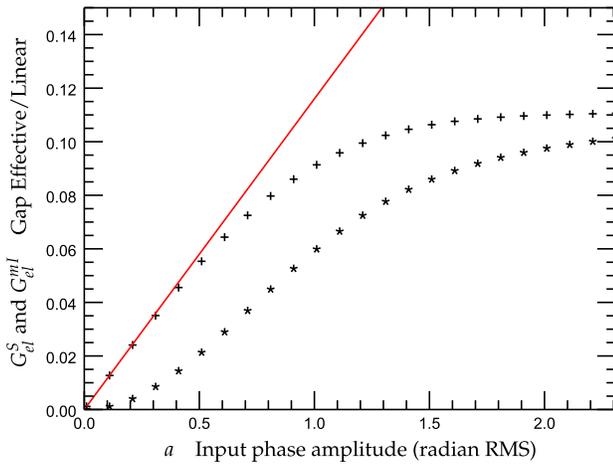
$$s_S(\phi_t) = \|S_l^x(\phi_t)\|_2 + \|S_l^y(\phi_t)\|_2.$$

Very fortunately, the loss of sensitivity due to the computation of the slope maps is insignificant, and the equality  $s_S(\phi_t) = s(\phi_t)$  may be considered as true. This is surprising, since a part of information is lost during Eqs. (27) and (28). This remarkable fact will be discussed in a paper dedicated to the numerical handling of the meta-intensities (see [10], in preparation).

On the other hand, the calculation of the quadratic slopes maps shows that they are equal to zero:

$$S_q^x(\phi_t) = 0 \quad \text{and} \quad S_q^y(\phi_t) = 0.$$

The linear Eqs. (27) and (28) thus have a positive influence on the linearity range without any significant loss of sensitivity! This result constitutes a strong argument in favor of the slope maps. Unfortunately, it does not mean that the PWFS associated with the slope maps has an infinite linearity range: if the quadratic intensity is null, it is not the case for the next phase powers. In other words, the linearity ranges associated with the slope maps



**Fig. 8.** Distance between the effective and linear meta-intensities (+,  $G_{el}^S$ ) and the effective and linear slope maps (\*,  $G_{el}^{mI}$ ) as functions of the input phase amplitude. The slope of the red line equals the 2-norm of the quadratic intensity associated with the meta-intensities. The input phase is the vertical coma, and the considered WFS is the PWFS with an infinite apex angle.

are determined by the cubic intensity, i.e., the third term ( $q = 3$ ) of Eq. (14). Such a fact is visible in Fig. 8: the slope at the origin of the distance between the effective and linear slopes maps, which is defined in Eq. (29), is null, but this distance does not equal zero:

$$G_{el}^S(\phi_t, a) = \left\| \frac{S_x(a\phi_t)}{a} - S_l^x(\phi_t) \right\|_2 + \left\| \frac{S_y(a\phi_t)}{a} - S_l^y(\phi_t) \right\|_2, \quad (29)$$

$$G_{el}^{mI}(\phi_t, a) = \sum_{4\text{ pupils}} \left\| \frac{mI^{\pm\pm}(a\phi_t)}{a} - I_l^{\pm\pm}(\phi_t) \right\|_2. \quad (30)$$

## 7. APPLICATION TO THE ZERNIKE WFS

### A. Tessellation Formalism

This section focuses on the Zernike WFS initially introduced by Zernike himself [5]. This WFS allows us to illustrate the piston presence in the tessellation formalism. Indeed, this sensor works by splitting the incoming energy into two parts thanks to polar tessellations  $\Omega_\rho$  and  $\tilde{\Omega}_\rho$  and by creating an optical path difference between these two contributions. Note that there are no rejection tip/tilt angles for the Zernike WFS. Hence, we can write

$$\Omega_\rho: \delta, 0, 0 \quad \tilde{\Omega}_\rho: 0, 0, 0.$$

As a consequence, the 2DFT of the Zernike mask with optical parameters  $\delta$  and  $\rho$  is

$$\mathcal{F}[m_Z](r_d) = \mathcal{F}[\mathbb{I}_{\tilde{\Omega}_\rho}](r_d) + \exp\left(\frac{2i\pi}{\lambda_0} \delta\right) \mathcal{F}[\mathbb{I}_{\Omega_\rho}](r_d), \quad (31)$$

$$= \delta(r_d) + \left( \exp\left(\frac{2i\pi}{\lambda_0} \delta\right) - 1 \right) \frac{\rho}{r_d} J_1(2\pi\rho r_d). \quad (32)$$

Historically, the size of the central mask  $\rho$  was calculated in order to have the same amount of energy in  $\Omega_\rho$  and  $\tilde{\Omega}_\rho$ , i.e.,  $\rho = 1.06\lambda_0/D$  for a circular pupil; the depth  $\delta$  associated with the piston was set to create a  $\pi/2$  phase gap between the two waves coming from  $\Omega_\rho$  and  $\tilde{\Omega}_\rho$ . Nevertheless, the previous set of optical parameters is not the only one that generates a WFS.  $\rho$  provides the energetic partition between the two parts of the polar tessellation and an exact equality is actually not strictly needed: a certain leeway exists and allows us to modify some properties of the sensor. In the same way, recent works by N'Diaye *et al.* [15] show that  $\delta$  may range in the interval  $[-\lambda_0/8; 3\lambda_0/8]$  and still defines a WFS. Consequently, as for the pyramid WFS, the historical Zernike sensor is only a member of a vaster class of sensors generated by the polar tessellation and the two optical parameters  $\rho$  and  $\delta$ .

By looking at Eq. (32), it appears that the entrance pupil is convoluted with the  $J_1$  Bessel function. As a consequence, we define the following operator, which we call the Zernike operator  $\mathcal{Z}_\rho$ , as

$$\mathcal{Z}_\rho[f] = f \star \frac{\rho}{r_d} J_1(2\pi\rho r_d).$$

Thanks to the general definition of the linear intensity  $I_l$  [see Eq. (16)], it is possible to get the purely linear behavior of the ZWFS:

$$I_l(\phi_t) = 2 \sin\left(\frac{2\pi}{\lambda_0}\delta\right) \overbrace{\mathbb{I}_\rho(\phi_t \mathcal{Z}_\rho[\mathbb{I}_\rho] - \mathcal{Z}_\rho[\mathbb{I}_\rho \phi_t])}^a. \quad (33)$$

The spatial variability of the linear intensity  $I_l$  is coded by the term  $b$  of Eq. (33). One notes that this variability only depends on the optical parameter  $\rho$ , which is the size of the Zernike mask. The depth of this mask coded by the optical parameter  $\delta$  allows us to adjust the global scalar factor  $a$ .

The quadratic intensity  $I_q$  is calculated via its definition in Eq. (17):

$$I_q(\phi_t) = \left[ 1 - \cos\left(\frac{2\pi}{\lambda_0}\delta\right) \right] (2\mathcal{Z}_\rho^2[\phi_t] - 2\phi_t \mathcal{Z}_\rho[\phi_t] + \mathcal{Z}_\rho[\phi_t^2] + \phi_t^2 \mathcal{Z}_\rho[\mathbb{I}_\rho] - 2\mathcal{Z}_\rho[\mathbb{I}_\rho] \mathcal{Z}_\rho[\phi_t^2]).$$

Once again, it appears that  $\rho$  only influences the spatial variability of the quadratic intensity, whereas  $\delta$  adjusts a global scalar factor.

### B. Sensitivity, Linearity Range, and SD Factor

From the expression of the linear intensity, we can get the sensitivity associated with the input phase  $\phi_t$ :

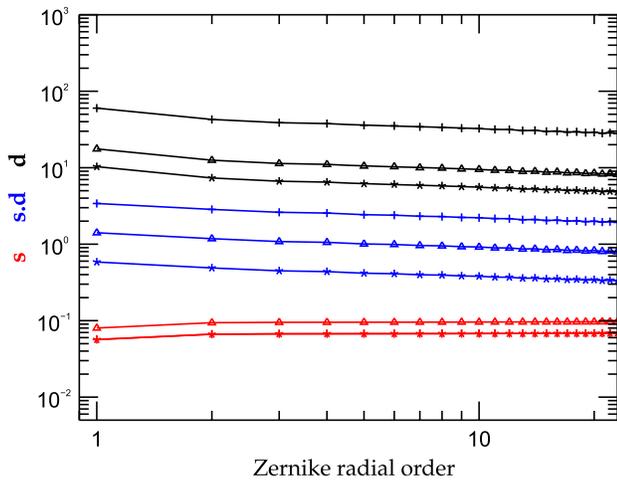
$$s(\phi_t) = 2 \left| \sin\left(\frac{2\pi}{\lambda_0}\delta\right) \right| \left\| \phi_t \mathcal{Z}_\rho[\mathbb{I}_\rho] - \mathcal{Z}_\rho[\mathbb{I}_\rho \phi_t] \right\|_2. \quad (34)$$

The linearity range is directly related to the inverse of the 2-norm of the quadratic intensity:

$$d(\phi_t) = \frac{1}{1 - \cos\left(\frac{2\pi}{\lambda_0}\delta\right)} \left\| 2\mathcal{Z}_\rho^2[\phi_t] - 2\phi_t \mathcal{Z}_\rho[\phi_t] + \mathcal{Z}_\rho[\phi_t^2] + \phi_t^2 \mathcal{Z}_\rho[\mathbb{I}_\rho] - 2\mathcal{Z}_\rho[\mathbb{I}_\rho] \mathcal{Z}_\rho[\phi_t^2] \right\|_2^{-1}. \quad (35)$$

#### 1. Depth of the Zernike Mask

The size of the Zernike mask  $\rho$  is set at  $1.06\lambda_0/D$ . We study the influence of its depth on the sensitivity, the linearity range, and the SD factor. Equation (34) shows that the sensibility is maximum for  $\delta = \lambda_0/4$ . That corresponds to the historical Zernike WFS. From  $\delta = 0$  to  $\delta = \lambda_0/2$ , we then observe, thanks to Eq. (35), that the linearity range is decreasing. When  $\delta$  tends



**Fig. 9.** Sensitivity, linearity range, and SD factor of the Zernike WFS with respect to the spatial frequencies. Depth of the Zernike mask  $\delta$  equals  $1/8$  (+),  $1/4$  ( $\Delta$ ), and  $3/8$  (\*)  $\lambda_0$ . The sensitivity is identical for  $\delta$  equals  $1/8$  and  $3/8$   $\lambda_0$ .

to 0, the linearity range is infinite, but such a configuration corresponds to the trivial mask and the sensitivity is then null. Figure 9 confirms these results. Moreover, the SD factor curve shows that deep Zernike masks, e.g.,  $3/8\lambda_0$ , allow us to get a significant gain in terms of the linearity range without too much sensitivity loss. Finally, we effectively observe that the shapes of the curves do not change when  $\delta$  varies. This parameter only translates them by the global factors of Eqs. (34) and (35).

#### 2. Size of the Zernike Mask

We set in this paragraph the depth of the Zernike mask at  $\lambda_0/4$  and study the influence of its size on the performance criteria. We observe in Fig. 10 that this time, the shapes of curves change with  $\rho$ . This is fully consistent with Eqs. (34) and (35). Moreover, a larger Zernike mask slightly improves the sensitivity, while the linearity range decreases. The SD factor approximately stays constant. We conclude this part about the Zernike WFS class by remarking that the first optical stage of the Roddier and Roddier coronagraph is also a member of this class.

## 8. POLYCHROMATIC LIGHT

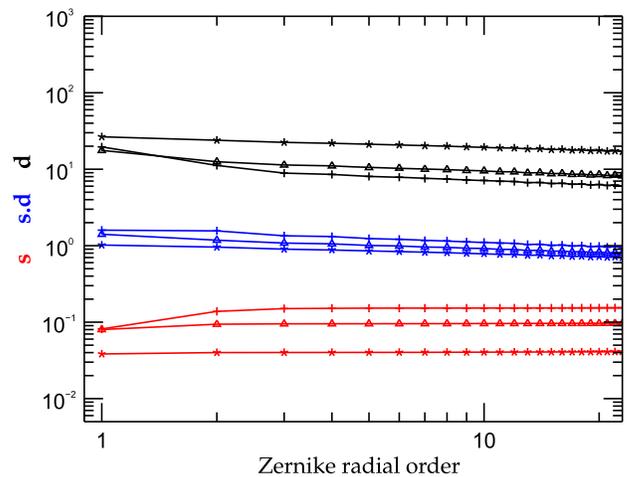
In this section, we explore the influence of incoming polychromatic light on the previous mathematical developments.

### A. Polychromatic Intensity

The first step consists of writing all the previous formulas by making explicit the dependence on the wavelength  $\lambda$ . The incident EM field is

$$\psi_\rho(x_p, y_p, \lambda) = \sqrt{n(\lambda)\mathbb{I}_\rho(x_p, y_p)} \exp\left(\frac{2i\pi}{\lambda} \Delta(x_p, y_p)\right),$$

with  $n(\lambda)$  defined in such a way that  $n(\lambda)d\lambda$  is the number of photons by unit area in the wavelength range  $[\lambda - d\lambda/2; \lambda + d\lambda/2]$ . In other words,  $n(\lambda)$  is the spectrum of the studied source. The optical path difference  $\Delta$  is not dependent on the wavelength, since we assume that the atmosphere is not a dispersive medium. If we choose a particular wavelength  $\lambda_0$  as a reference, it is possible to write the phase of the incoming field as



**Fig. 10.** Sensitivity, linearity range, and SD factor of the Zernike WFS with respect to the spatial frequencies. Size of the Zernike mask  $\rho$  equals  $0.5$  (\*),  $1$  ( $\Delta$ ), and  $1.5$  (+)  $1.06\lambda_0/D$ .

$$\phi = \frac{2\pi}{\lambda} (\Delta_t + \Delta_r) = \frac{2\pi \lambda_0}{\lambda_0 \lambda} (\Delta_t + \Delta_r) \triangleq \frac{\lambda_0}{\lambda} (\phi_t + \phi_r), \quad (36)$$

where  $\phi_t$  (resp.  $\phi_r$ ) is the turbulent phase corresponding to the turbulent OPD  $\Delta_t$  (resp. reference OPD  $\Delta_r$ ) at the wavelength  $\lambda_0$ . In other words,  $\lambda_0$  is the analysis wavelength. Equation (36) also means that the only impact of polychromatism on the incoming turbulent phase corresponds in a scale factoring. Polychromatism does not mix two different turbulent phase modes together. This fact is essential for the continuation of our study.

The transparency function of the mask does not change, but we now keep its  $\lambda$ -dependency:

$$m(f_x, f_y, \lambda) = \exp\left(\frac{2i\pi}{\lambda} \text{OS}(f_x, f_y, \lambda)\right).$$

The optical shape of the mask depends, *a priori*, on the wavelength, since this quantity is the product of the refractive index  $n_r$  (which depends on  $\lambda$  for dispersive material) and the geometrical shape of the mask, GS,

$$\text{OS}(f_x, f_y, \lambda) = n_r(\lambda) \text{GS}(f_x, f_y).$$

The last source of  $\lambda$  dependency comes from the Fresnel formalism. Indeed, the light propagation itself includes the wavelength, as we can see in the expression of the EM field in the detector plane,

$$\psi_d(x_d, y_d) \propto \iint \frac{dx_p dy_p}{f^2} \psi_p(x_p - x_d, y_p - y_d, \lambda) \iint \frac{df_x df_y}{\lambda^2} m(f_x, f_y, \lambda) \exp\left(-\frac{2i\pi}{f\lambda} [x_p f_x + y_p f_y]\right).$$

Concretely, it means that the right-side integral of this equation corresponds to a  $\lambda$ -dependent 2DFT of the mask. In other words, without particular assumptions about the mask, its Fourier transform depends on the wavelength. The fundamental Eq. (4) becomes

$$I_\lambda(\phi_t) = |\psi_p(\phi_t, \lambda) \star \mathcal{F}_\lambda[m]|^2.$$

Under these new assumptions, we can write the polychromatic intensity on the detector, called  $I_p$ , as the integral of the monochromatic intensities on the whole spectrum of the light source:

$$I_p(\phi_t) = \int d\lambda |\psi_p(\phi_t, \lambda) \star \mathcal{F}_\lambda[m]|^2. \quad (37)$$

### B. Substitution Test

The chromatic 2DFT equals

$$\mathcal{F}_\lambda[m](x_p, y_p) = \iint \frac{df_x df_y}{\lambda^2} m(f_x, f_y, \lambda) e^{\frac{2i\pi}{\lambda} [x_p f_x + y_p f_y]}.$$

By looking at this equation, it appears that the chromatic 2DFT may be decoupled from the wavelength on the condition that the substitution

$$(u, v) = \left(\frac{f_x}{\lambda}, \frac{f_y}{\lambda}\right), \quad (38)$$

makes the  $\lambda$  dependency disappear. This condition requires in fact that the transparency function of the mask is a function of only two variables  $u$  and  $v$  instead of three:  $f_x$ ,  $f_y$ , and  $\lambda$ . In terms of physics, such a condition means that the point-spread functions for every wavelength always see the same “mask.” In other words,

the optical shape of this mask is invariable by scale change. In particular, one can note that as soon as a mask does not need any characteristic length quantity in its transparency function, the substitution of Eq. (38) is possible.

Under the substitution of Eq. (38) assumption, the 2DFT becomes independent of the wavelength. The total polychromatic intensity [Eq. (37)] may be simplified into

$$I_p(\phi_t) = \int n(\lambda) d\lambda \left| \mathbb{I}_p \exp\left(1 \frac{\lambda_0}{\lambda} (\phi_t + \phi_r)\right) \star \mathcal{F}[m] \right|^2.$$

The phase power series of such an intensity becomes

$$I_p(\phi_t) = \sum_{q=0}^{\infty} \frac{(-1)^q}{q!} \int n(\lambda) \left(\frac{\lambda_0}{\lambda}\right)^q \sum_{k=0}^q (-1)^k \binom{q}{k} \phi_t^{k\star} \overline{\phi_t^{q-k\star}} d\lambda$$

where  $\phi_t^{k\star} \triangleq \left(\mathbb{I}_p e^{\frac{2i\pi\lambda_0}{\lambda} \phi_r} \phi_t^k\right) \star \mathcal{F}[m]. \quad (39)$

We assume, from now on, that  $\phi_r = 0$ . As a consequence, the  $k$ th moment does not depend on  $\lambda$  anymore. Subsequently, the phase power series in Eq. (39) may be simplified into

$$I_p(\phi_t) = \sum_{q=0}^{\infty} \frac{(-1)^q}{q!} \int n(\lambda) \left(\frac{\lambda_0}{\lambda}\right)^q d\lambda \sum_{k=0}^q (-1)^k \binom{q}{k} \phi_t^{k\star} \overline{\phi_t^{q-k\star}}$$

where  $\phi_t^{k\star} \triangleq (\mathbb{I}_p \phi_t^k) \star \mathcal{F}[m]. \quad (40)$

### C. Polychromatic Meta-Intensities

The first terms of the previous equation correspond once again to the constant, linear, and quadratic terms,  $I_{pc}$ ,  $I_{pl}$ , and  $I_{pq}$ :

$$I_{pc} = |\mathbb{I}_p \star \mathcal{F}[m]|^2$$

$$I_{pl}(\phi_t) = \frac{1}{n} \left( \int n(\lambda) \frac{\lambda_0}{\lambda} d\lambda \right) 2\Re\left[\mathbb{I}_p \star \mathcal{F}[m] \overline{(\mathbb{I}_p \phi_t \star \mathcal{F}[m])}\right]$$

$$I_{pq}(\phi_t) = \frac{1}{n} \left( \int n(\lambda) \left(\frac{\lambda_0}{\lambda}\right)^2 d\lambda \right) [|\mathbb{I}_p \phi_t \star \mathcal{F}[m]|^2 - \Re\left[\mathbb{I}_p \star \mathcal{F}[m] \overline{(\mathbb{I}_p \phi_t^2 \star \mathcal{F}[m])}\right]],$$

where  $n$  is the total flux,

$$n = \int n(\lambda) d\lambda.$$

The method used previously to construct a phase-linear quantity (at least in the small-phase approximation) is once again valid. The effective polychromatic meta-intensity, called  $mI_p$ , will be subsequently defined in the same way as Eqs. (18) and (23), i.e.,

$$mI_p(\phi_t) = \frac{I_p(\phi_t) - I_p(0)}{n}. \quad (41)$$

Subsequently, it is this time easy to do the link between the polychromatic linear and quadratic intensities to the monochromatic ones:

$$I_{pl}(\phi_t) = \frac{1}{n} \left( \int n(\lambda) \frac{\lambda_0}{\lambda} d\lambda \right) I_l(\phi_t)$$

$$I_{pq}(\phi_t) = \frac{1}{n} \left( \int n(\lambda) \left(\frac{\lambda_0}{\lambda}\right)^2 d\lambda \right) I_q(\phi_t).$$

It appears thus that under the assumption that the substitution of Eq. (38) is valid and that the reference phase equals the null phase, the polychromatic linear and quadratic intensities only differ from the monochromatic ones by global gain factors, which

only depend on the spectrum  $n(\lambda)$  and on the reference wavelength  $\lambda_0$  but not on the incoming phase mode.

In particular, it means that the polychromatism does not induce a change in the spatial structure of the meta-intensities, but only introduces a global blurring factor. In other words, an interaction matrix done with a monochromatic reference source is still valid for a polychromatic source if such a global factor is taken into account.

The substitution of Eq. (38) is thus a robust way to define what is an achromatic sensor. Indeed, if this is not possible, the factorization in the power series in Eq. (39) is not possible anymore. Concretely, it means that there is crosstalk between the phase modes due to the polychromatism of the source. The interaction matrix will be, subsequently, changed in its spatial structure.

Finally, one can notice the impact of a non-null reference phase: a WFS that does not work around the null phase is irremediably chromatic.

#### D. Sensitivity, Linear Range, and SD Factor

We can finally note that, for an achromatic sensor, the polychromatic sensitivity and linearity range (called  $s_p$  and  $d_p$ ) may easily be linked to the monochromatic ones:

$$s_p(\phi_t) = \frac{1}{n} \left( \int n(\lambda) \frac{\lambda_0}{\lambda} d\lambda \right) s(\phi_t)$$

$$d_p(\phi_t) = n \left( \int n(\lambda) \left( \frac{\lambda_0}{\lambda} \right)^2 d\lambda \right)^{-1} d(\phi_t).$$

#### E. Applications to Classical WFSs

As we have just seen in the previous paragraph, as soon as the reference phase equals zero, the chromatic behavior of a WFS only depends on the expression of the transparency function of its Fourier mask. In this section, we browse the previous examples of WFSs and study their chromatic behavior in light of the substitution test defined in Eq. (38).

##### 1. Pyramid WFS Class

Perfect pyramidal mask. We first consider the class of pyramid WFSs. The general expression of the transparency function of the associated mask is

$$m_{\Delta}(f_x, f_y, \lambda) = \exp \left( \frac{2i\pi}{\lambda} n_r(\lambda) \theta(|f_x| + |f_y|) \right).$$

There are two cases, depending on whether we consider a dispersive pyramid (e.g., a transparent glass pyramid) or a reflective pyramid (see [12]). In the first case,  $n_r(\lambda)$  follows the usual empirical Cauchy's equation,

$$n_r(\lambda) = B + \frac{C}{\lambda^2} + \dots$$

where  $B$  and  $C$  depend on the nature of the propagation medium. Note that in this case, it is impossible to make  $m$  only a function of the two variables  $f_x/\lambda$  and  $f_y/\lambda$ ; it means that the associated WFS is not achromatic. Fortunately, a classical solution to solve this problem exists which consists of using two transparent pyramids attached by their base (see [16]). Such a device allows us, in substance, to nullify the first  $\lambda$ -dependent term  $C$  of the refractive index. In this way, variable substitution becomes possible, which makes the WFS achromatic. In the case of a reflective pyramid,

the propagation medium is air, which is not significantly dispersive on the visible spectrum. Subsequently, the substitution can be made and the simplified transparency function is

$$m_{\Delta}(u, v) = \exp(2i\pi n_{\text{air}} \theta(|u| + |v|)).$$

These results prove that the reflective pyramid and the two attached pyramids are achromatic sensors. Moreover, the apex angle  $\theta$  stays a free parameter. Subsequently, both the classical PWFS introduced by Ragazzoni [1] where the pupil images are completely separated and the flattened PWFS we proposed in [4] are achromatic sensors.

Modulation and Chromaticity. We already saw (paragraph 5.B) that the modulated phase  $\phi_m(s)$  may be seen as a variable reference phase. Moreover, the results of paragraph 8.C showed that a non-null reference phase irremediably induces a structural dependence of the output meta-intensities regarding the spectrum of the source. Consequently, the modulated PWFS (and all its variations) is not an achromatic sensor.

##### 2. Zernike Mask

The transparency function of the mask of the Zernike WFS class needs, unfortunately, 2 characteristic length quantities: the first one is the size of the central disk, allowing us to define the two parts of the polar tessellation. This size is set in order to set the energy partition in the 2 parts of the tessellation. For a circular pupil of a diameter  $D$  and a reference wavelength  $\lambda_0$ , the size of the pupil equals  $1.06\lambda_0/D$ . The second characteristic quantity is the depth of the central disk well. It is set in order to create an OPD between the two parts of the polar tessellation. In other words, the optical design of the Zernike mask is optimized only for one wavelength. Moreover, this mask is manufactured in a transparent medium, which is, *a priori*, dispersive as well, but we do not consider such a chromatic effect for the following study. Such considerations allow us to write the transparency function as

$$m_Z(f_x, f_y, \lambda) = 1 + (e^{\frac{w_0}{2\lambda}} - 1) \Theta \left( 1.06 \frac{\lambda_0}{D} - \sqrt{f_x^2 + f_y^2} \right).$$

It is clearly seen that the substitution of Eq. (38) does not allow us to cancel the wavelength dependence of the mask. Subsequently, the Zernike WFS cannot be an achromatic sensor. Nevertheless, N'Diaye *et al.* [17] showed that chromaticity does not play a significant role in terms of the error budget, making the ZWFS practically achromatic.

## 9. CONCLUSIONS

Thanks to the original approach of the Fourier plane tessellation, we managed to unify all the optical designs based on Fourier filtering. The essential point is to consider any kind of Fourier mask as a spatial frequency splitter.

From this approach, we showed that the 2D Fourier transform of the mask plays a significant role all along the mathematical developments, allowing us, in particular, to efficiently describe the imaging for any kind of entrance pupils or incoming phases. A phase power series development of the intensity on the detector allows us to identify the linear term. In order to extract this one from the effective signal obtained on the detector, we introduced a numerical process that may be seen as the simplest way to get a phase-linear response in the small-phase regime. The quantity obtained after such a calculation has been called the meta-intensity

$mI$  and is defined in the same way for any kind of Fourier-based WFSs.

These analytical developments have then allowed us to define the sensitivity and the linearity ranges of all the WFSs studied here. They are directly based on the linear and quadratic terms of the phase power series of the intensity. Moreover, we defined the SD factor, which quantifies the trade-off between these two antagonistic performance criteria. Such a factor gave us some practical ideas to optimize the Fourier masks depending on the required specifications. In other words, it opens the way to free-form filtering masks.

We then extended all these results to the modulation. This is particularly useful when considering unusual modulations, such as nonuniform or non-tip/tilt modulations.

The last theoretical result was about the effect of a polychromatic incoming light on the efficiency of a WFS. We gave a robust criterion depending only on the transparency function of the mask in order to characterize the achromatic property of a WFS.

As a second goal for this paper, we applied the theoretical developments to existing and new sensors. Indeed, the developed formalism allowed us to unify (and this was its main purpose) all the Fourier-based WFSs. We showed in particular that the pyramid WFS and the Zernike WFS may be considered not as unique designs but more as classes, where optical characteristics become flexible parameters.

It appeared, for instance, that the flattened PWFS and the classical PWFS were only two types of the pyramid WFSs class, where the apex angle is considered as a parameter. An exhaustive study that explores in great detail all the parameters of the pyramid class, i.e., the apex angle, the number of faces, and the modulation parameters, is conducted in [18].

We also showed that the slope maps usually associated with the PWFS were directly linked to the general meta-intensities we introduced. Moreover, we showed that the linearity range was improved without significant loss of sensitivity thanks to this slope maps computation.

Finally, new WFSs have been studied by considering the depth and the size of the central well of the Zernike WFS as a free parameter. Moreover, we isolated, analytically, the role played by each of these two parameters in the response of the sensors of the ZWFS class.

## APPENDIX A: NOTATIONS

$$\text{2D Fourier Transform: } \mathcal{F}[f](\mu, \nu) = \iint f(x, y) e^{-2i\pi(x\mu + y\nu)} dx dy$$

$$\text{Identity transform: } \mathcal{I}[f](x, y) = f(x, y)$$

$$\text{Symmetric transform: } \mathcal{S}[f](x, y) = f(-x, -y)$$

$$\text{Hilbert transforms: } \mathcal{H}_x[f](x, y) = p.v. \left\{ \frac{1}{\pi} \int \frac{f(t, y)}{x-t} dt \right\} -$$

$$\mathcal{H}_y[f](x, y) = p.v. \left\{ \frac{1}{\pi} \int \frac{f(x, t)}{y-t} dt \right\} -$$

$$\mathcal{H}_{xy}^2[f](x, y) =$$

$$p.v. \left\{ \frac{1}{\pi^2} \iint \frac{f(t, t')}{(x-t)(y-t')} dt dt' \right\}$$

$$\text{2-Norm: } \|f\|_2 = \left( \iint |f(x, y)|^2 dx dy \right)^{1/2}$$

**Funding.** European Commission (EC) (312430); Agence Nationale de la Recherche (ANR); French Aerospace Lab (ONERA).

## REFERENCES

1. R. Ragazzoni, "Pupil plane wavefront sensing with an oscillating prism," *J. Mod. Opt.* **43**, 289–293 (1996).
2. V. Akondi, S. Castillo, and B. Vohnsen, "Multi-faceted digital pyramid wavefront sensor," *Opt. Commun.* **323**, 77–86 (2014).
3. B. Vohnsen, S. Castillo, and D. Rativa, "Wavefront sensing with an axicon," *Opt. Lett.* **36**, 846–848 (2011).
4. O. Fauvarque, B. Neichel, T. Fusco, and J.-F. Sauvage, "Variation around a pyramid theme: optical recombination and optimal use of photons," *Opt. Lett.* **40**, 3528–3531 (2015).
5. F. Zernike, "Diffraction theory of the knife-edge test and its improved form, the phase-contrast method," *Mon. Not. R. Astron. Soc.* **94**, 377–384 (1934).
6. K. Dohlen, "Phase masks in astronomy: from the Mach-Zehnder interferometer to coronagraphs," *EAS Pub. Ser.* **12**, 33–44 (2004).
7. A. Toepler, *Beobachtung nach einer neuen optischen Method* (1864).
8. R. N. Smartt and W. H. Steel, "Theory and application of point-diffraction interferometers," *Jpn J. Appl. Phys.* **14**, 351 (1975).
9. J. E. Oti, V. F. Canales, and M. P. Cagigal, "Analysis of the signal-to-noise ratio in the optical differentiation wavefront sensor," *Opt. Express* **11**, 2783–2790 (2003).
10. O. Fauvarque, B. Neichel, T. Fusco, and J.-F. Sauvage, "Optimization of the imaging processing in the context of Fourier based wave front sensing," in preparation.
11. F. Rigaut and E. Gendron, "Laser guide star in adaptive optics, the tilt determination problem," *Astron. Astrophys.* **261**, 677–684 (1992).
12. A. Wang, J. Yao, D. Cai, and H. Ren, "Design and fabrication of a pyramid wavefront sensor," *Opt. Eng.* **49**, 073401 (2010).
13. C. Vérinaud, "On the nature of the measurements provided by a pyramid wave-front sensor," *Opt. Commun.* **233**, 27–38 (2004).
14. I. Shatokhina, A. Obereder, and R. Ramlau, "Fast algorithm for wavefront reconstruction in XAO/SCAO with pyramid wavefront sensor," *Proc. SPIE* **9148**, 91480P (2014).
15. M. N'Diaye, K. Dohlen, A. Caillat, and A. Costille, "Design optimization and lab demonstration of ZELDA: a Zernike sensor for near-coronagraph quasi-static measurements," *Proc. SPIE* **9148**, 91485H (2014).
16. S. Esposito, A. Tozzi, D. Ferruzzi, M. Carbillet, A. Riccardi, L. Fini, C. Vérinaud, M. Accardo, G. Brusa, D. Gallieni, R. Biasi, C. Baffa, V. Billotti, I. Foppiani, A. Puglisi, R. Ragazzoni, P. Ranfagni, P. Stefanini, P. Salinari, W. Seifert, and J. Storm, "First-light adaptive optics system for large binocular telescope," *Proc. SPIE* **4839**, 164–173 (2003).
17. M. N'Diaye, K. Dohlen, T. Fusco, and B. Paul, "Calibration of quasi-static aberrations in exoplanet direct-imaging instruments with a Zernike phase-mask sensor," *Astron. Astrophys.* **555**, A94 (2013).
18. O. Fauvarque, B. Neichel, T. Fusco, J.-F. Sauvage, and O. Giraut, "A general formalism for Fourier based wave front sensing: application to the pyramid wave front sensors," *Proc. SPIE* **9909**, 990960 (2016).