A&A 500, 647–655 (2009)
DOI: 10.1051/0004-6361/200811061
© ESO 2009

Astronomy & Astrophysics

Optimal point spread function modeling for weak lensing: complexity and sparsity

S. Paulin-Henriksson[1], A. Refregier[1], and A. Amara[2,3]

[1] Service d'Astrophysique, CEA Saclay, Bâtiment 709, 91191 Gif–sur–Yvette Cedex, France
e-mail: paulin@in2p3.fr
[2] Department of Physics, ETH Zürich, Wolfgang-Pauli-Strasse 16, 8093 Zürich, Switzerland
[3] Department of Physics and Center for Theoretical and Computational Physics, University of Hong Kong, Pok Fu Lam Road, Hong Kong

Received 30 September 2008 / Accepted 23 December 2008

ABSTRACT

Context. We address the issue of controling the systematic errors in shape measurements for weak gravitational lensing.
Aims. We make a step to quantify the impact of systematic errors in modeling the point spread function (PSF) of observations, on the determination of cosmological parameters from cosmic shear.
Methods. We explore the impact of PSF fitting errors on cosmic shear measurements using the concepts of complexity and sparsity. Complexity, introduced in a previous paper, characterizes the number of degrees of freedom of the PSF. For instance, fitting an underlying PSF with a model of low complexity produces small statistical errors on the model parameters, although these parameters could be affected by large biases. Alternatively, fitting a large number of parameters (i.e. a high complexity) tends to reduce biases at the expense of increasing the statistical errors. We attempt to find a trade-off between scatters and biases by studying the mean squared error of a PSF model. We also characterize the model sparsity, which describes how efficiently the model is able to represent the underlying PSF using a limited number of free parameters. We present the general case and give an illustration for a realistic example of a PSF modeled by a shapelet basis set.
Results. We derive a relation between the complexity and the sparsity of the PSF model, the signal-to-noise ratio of stars and the systematic errors in the cosmological parameters. By insisting that the systematic errors are below the statistical uncertainties, we derive a relation between the number of stars required to calibrate the PSF and the sparsity of the PSF model. We discuss the impact of our results for current and future cosmic shear surveys. In the typical case where the sparsity can be represented by a power-law function of the complexity, we demonstrate that current ground-based surveys can calibrate the PSF with few stars, while future surveys will require hard constraints on the sparsity in order to calibrate the PSF with 50 stars.

Key words. gravitational lensing – cosmology: dark matter – cosmology: cosmological parameters – cosmology: observations

1. Introduction

Studying spatial correlations between galaxy shapes induced by gravitational lensing of the large scale structure ("Cosmic Shear") is a powerful probe of dark energy and dark matter. A number of current and planned surveys are dedicated to cosmic shear, such as: the Canada-France-Hawaii-Telescope Legacy Survey[1] (CFHTLS), the KIlo Degree Survey and the VISTA Kilo-Degree Infrared Galaxy Survey[2] (KIDS/VIKING), the Dark Energy Survey[3] (DES), the Panoramic Survey Telescope & Rapid Response System[4] (Pan-STARRS), the SuperNovae Acceleration Probe[5] (SNAP), the Large Synoptic Survey Telescope[6] (LSST) and the Dark UNiverse Explorer[7] (DUNE/Euclid).

The most efficient way of improving the statistical precision of cosmic shear analyses is by enlarging the surveys. As long as the median redshift is sufficiently high ($z \gtrsim 0.7$), Amara & Refregier (2007a) demonstrate that, it is more advantageous to perform the cosmic shear surveys in fields as wide as possible, rather than deep. This in order to minimize the error bars in cosmological parameters. To date, the largest data optimized for cosmic shear is the Wide field of the CFHTLS, which covers 50 deg^2 (Fu et al. 2008) and will eventually reach 170 deg^2. Another analysis has also been published that combines 4 surveys that, together, cover an area of 100 deg^2 (Benjamin et al. 2007). In a few years, the KIDS/VIKING survey will cover 1500 deg^2. Eventually, projects currently being planned, such as DUNE/Euclid, planned for 2017, will be able to perform cosmic shear measurements over the entire observable extragalactic sky (~20 000 deg^2).

To let cosmic shear surveys reaching their full potential, it is necessary to ensure that systematic errors are sub-dominant relative to statistical uncertainties. In particular, a tight control on all the effects associated with shape measurements is required. To illustrate the difficulty in compiling accurate shear measurements, we begin with an overview of the "forward process" that illustrates how the original image of a galaxy is distorted in forming the final image that we measure. In the forward process,

[1] http://www.cfht.hawaii.edu/Science/CFHLS/
[2] http://www.eso.org/sci/observing/policies/PublicSurveys/sciencePublicSurveys.html
[3] http://www.darkenergysurvey.org
[4] http://pan-starrs.ifa.hawaii.edu
[5] http://snap.lbl.gov
[6] http://www.lsst.org
[7] http://www.dune-mission.net and http://www.esa.int/esaCP/index.html

a galaxy image is: (i) sheared by gravitational lensing; (ii) convolved with a point spread function (PSF) originating in a number of sources (e.g. instruments and atmosphere); (iii) pixelated at the detector; and finally (iv) affected by noise.

Cosmic shear analyses involve the reverse process: we begin with the final image and move backward from step (iv) to (i) in recovering the original lensing effect. A detailed and illustrated description of the forward and inverse processes is given in the GREAT08 Challenge Handbook (Bridle et al. 2008). The GREAT08 Challenge aims to provide a wide range of expertise into gravitational lensing by presenting the relevant issues in a clear way so as to make it accessible to a broad range of disciplines, including the machine-learning and signal-processing communities. Other similar challenges have also been performed within the weak lensing community, as part of the STEP collaboration (Heymans et al. 2006; Massey et al. 2007; Rhodes et al. 2008), which focused mainly on understanding the systematic errors at play in current shear measurement methods. These challenges focus mainly on reducing the errors originating in the shape measurement method. However, even with a perfect method there are fundamental limits due the statistical potential of a data set.

In Paulin-Henriksson et al. (2008) (P1 hereafter), we investigated the link between systematic errors in the power spectrum and uncertainties in the PSF correction phase. The framework is the following. Since the PSF of an instrument varies on all scales, the PSF needs to be measured using the stars that surround the lensed galaxy. Each star provides an image of the PSF that is pixelated and noisy, which means that to reach a given accuracy in the knowledge of the PSF, a number of stars is required. We estimate the number of stars N_* required to calibrate the PSF to a given accuracy, according to the stellar signal-to-noise ratio (SNR), the minimum galaxy size, the complexity of the PSF and the tolerated variance in the systematics $\sigma^2_{\rm sys}$. On the other hand, Amara & Refregier (2007b) estimated the upper limit to $\sigma^2_{\rm sys}$ when estimating cosmological parameters. By combining both papers together, we derive the minimum number of stars required to reach a given accuracy. For instance, analyses completed to date, that allow us to constrain $\sigma_8\,\Omega_{\rm M}$ with an accuracy of 0.05, require $\sigma^2_{\rm sys}$ lower than a few 10^{-6} and the PSF to be calibrated by using 5 stars; while for future ambitious surveys that will allow us to constrain w_0 and w_a to an accuracy of 0.02 and 0.1, respectively, $\sigma^2_{\rm sys}$ must be lower than 10^{-7}, which requires at least 50 stars (for stars with signal-to-noise ratio of 500 and a PSF described by a few degrees of freedom, as can be typically achieved in space).

In P1 we use the same functional form for both the underlying PSF as well as the model used to fit it. This means that the PSF model is able to describe perfectly the underlying PSF. The errors in the fit due to noise causes a scatter of the fitted parameters around the truth. For instance, if the model is an orthogonal basis set, then the fitted parameters follow a Gaussian distribution around the truth. In this paper, we extend this investigation by studying the impact of fitting a PSF with a model that has a different form. This addresses the case in which the underlying PSF (unknown in practice) is estimated by fitting the parameters of an arbitrary model. This can lead to both a scatter in the fitting parameters and an offset in the average value relative to the true value, i.e. a bias in the fitting parameters. We can therefore model a given PSF using either a complex model of small biases but large scatters, or a simpler model that would lead to smaller scatters but larger biases. To quantify these effects, we revisit the concept of complexity proposed in P1 and introduce the concept of sparsity.

This paper is organised as follow: first, in Sect. 2, we discuss the concepts of complexity and sparsity, which are the key concepts of this paper; Sect. 3 presents our notation; Sect. 4 presents the definition of optimal complexity, illustrates our formalism with a PSF example and uses the sparsity as a tool for optimizing the complexity; Sect. 5 derives the minimum number of stars required to calibrate the PSF, extending results of P1; and finally, Sect. 6 summarizes our conclusions.

2. Complexity and sparsity

In P1, we introduced the concept of complexity; we demonstrated that a few complexity factors characterize the amount of information that needs to be collected about the PSF. This is summarized and revisited in Sect. 2.1. In Sect. 2.2, we introduce the concept of sparsity, which measures the ability of a PSF model to represent the underlying PSF with a small number of free parameters. This allows us to explore how an optimal PSF model can be constructed to minimize $\sigma_{\rm sys}$.

2.1. Complexity

In P1, we define the complexity factors of the PSF, which represent the number of degrees of freedom (DoF) that are estimated from stars (in the limit of infinite resolution, i.e. infinitely small pixels): the higher the number of DoF, the larger the complexity factors. Each PSF shape parameter is associated with a complexity factor that is related to the rms of its estimator. In the simple formalism where we consider unweighted quadrupole moments, the PSF is characterized by only two complexity factors ψ_ϵ and ψ_{R^2} associated with the 2 component PSF ellipticity $\epsilon_{\rm PSF}$ and the square PSF rms radius $R^2_{\rm PSF}$ respectively (as defined in P1). For a given star, one has:

$$\begin{pmatrix} \psi_{R^2} \\ \psi_\epsilon \end{pmatrix} \equiv \mathcal{S}_* \begin{pmatrix} \sigma[R^2_{\rm PSF}]/R^2_{\rm PSF} \\ \sigma[\epsilon_{\rm PSF}] \end{pmatrix} \tag{1}$$

where $\mathcal{S}_*$ is the photometric SNR of the star, $\sigma[R^2_{\rm PSF}]$ is the rms error of the PSF size estimator and $\sigma[\epsilon_{\rm PSF}]$ is the rms error of the PSF ellipticity estimator. As in P1, we assume that the small ellipticity regime holds (i.e. $|\epsilon_{\rm PSF}| \lesssim 0.1$), implying that the measure of $\epsilon_{\rm PSF}$ is isotropic and the 2 components of the ellipticity have the same rms uncertainty, i.e. $\sigma[\epsilon_{{\rm PSF},i}] \equiv \sigma[\epsilon_{\rm PSF}]$.

If the PSF can be considered constant over several stars, or for particular representations of the PSF (for example with shapelet basis sets in the small ellipticity regime, see P1), ψ_{R^2} and ψ_ϵ are spatially constant and Eq. (1) can be extended to a set of several stars. For a combination of several stars, $\mathcal{S}_*$ becomes $\sqrt{n_*}\mathcal{S}_{\rm eff}$:

$$\begin{pmatrix} \psi_{R^2} \\ \psi_\epsilon \end{pmatrix} \equiv \sqrt{n_*}\,\mathcal{S}_{\rm eff} \begin{pmatrix} \sigma[R^2_{\rm PSF}]/R^2_{\rm PSF} \\ \sigma[\epsilon_{\rm PSF}] \end{pmatrix}, \tag{2}$$

where $\mathcal{S}_{\rm eff}$ is the effective stellar SNR and n_* is the effective number of stars k used in the PSF calibration, as defined by (see P1):

$$n_*\,\mathcal{S}^2_{\rm eff} \equiv \sum_k \mathcal{S}^2_k\,. \tag{3}$$

We also show in P1 that the polar shapelet basis set, proposed by Massey & Refregier (2005), tested on simulated data in Massey et al. (2007) and used on real data by Bergé et al. (2008), is particularly convenient for modeling the PSF. For example, in

the small ellipticity regime, ψ_ϵ and ψ_{R^2} depend only on the polar shapelet basis set over which the PSF is decomposed, not on the PSF itself. For this reason we use shapelets in this paper when illustrating our discussions by an example. Note that our results and conclusions are not restricted to shapelets but remain valid whatever the PSF model. For convenience, we choose the shapelet "diamond" option (described in details in P1) that imposes a lower limit to the scales described, implying a link between ψ_ϵ and ψ_{R^2}:

$$\psi_\epsilon^2 = \psi_{R^2}^2 - \mathcal{N}, \tag{4}$$

with $\mathcal{N}$ the highest even integer lower than or equal to the order $n_{\max}$ of the basis set. We then consider the overall complexity Ψ defined in the following according to ψ_ϵ, ψ_{R^2} and the variance in the galaxy ellipticity distribution.

2.2. Sparsity

In this paper, we introduce the concept of sparsity of the PSF model, which describes how efficiently a model can represent the underlying PSF with a limited number of DoF (i.e. with a limited complexity). Specifically, the sparsity quantifies how the residuals between the estimated and the underlying PSF decrease as the complexity of the PSF model increases. With a high number of DoF, i.e. a high complexity, one might expect small residuals but large scatters ointhe fitted parameters. On the other hand, with a small number of DoF, i.e. a low complexity, one might expect large residuals but small scatters in the fitted parameters. The sparsity characterizes the slope in this relation and thus is an estimate of the amount of information that can be contained in a given number of DoF. We show how to use sparsity in optimizing the complexity and minimizing $\sigma_{\rm sys}$.

Consider the shape parameters $R^2_{\rm PSF}$ and $\epsilon_{\rm PSF}$ of the underlying PSF, as defined previously. The differences $\delta(R^2_{\rm PSF})$ and $\delta\epsilon_{\rm PSF}$ between the underlying PSF ("true" index) and its estimation ("est" index) can be written:

$$\begin{aligned} \delta(R^2_{\rm PSF}) &\equiv R^2_{\rm est} - R^2_{\rm true}, \\ \delta\epsilon_{\rm PSF} &\equiv \epsilon_{\rm est} - \epsilon_{\rm true}. \end{aligned} \tag{5}$$

These differences are of two types: the statistical scatter relative to the average σ and the bias-offset b of the average relative to the true value. The mean square errors (MSE) of $R^2_{\rm PSF}$ and $\epsilon_{\rm PSF}$ are:

$$\mathrm{MSE}[R^2_{\rm PSF}] \equiv \sigma^2[R^2_{\rm PSF}] + b^2[R^2_{\rm PSF}], \tag{6}$$

$$\mathrm{MSE}[\epsilon_{{\rm PSF},i}] \equiv \sigma^2[\epsilon_{{\rm PSF},i}] + b^2[\epsilon_{{\rm PSF},i}] \text{ with } i = 1, 2. \tag{7}$$

In P1, we address the zero bias case $b[\epsilon_i] = b[R^2] = 0$, that is equivalent to consider the PSF model is able to describe the underlying PSF perfectly. However, this nulling of the biases is not necessarily the optimal PSF modeling. It can be advantageous to work with a simplistic PSF model that is unable to describe all the PSF features and has some biases but low statistical scatter (see our PSF example in Sect. 4.1 and Fig. 1). This paper proposes another approach that consists in optimizing the PSF model in order to minimize $\sigma_{\rm sys}$. We do this by searching for the optimal trade-off between the systematic errors ($b[\epsilon_1]$, $b[\epsilon_2]$, and $b[R^2]$) and the statistical errors ($\sigma[\epsilon]$, and $\sigma[R^2]$), which is equivalent to searching the optimal complexity Ψ of the PSF model. The sparsity allows us to perfom this optimization because it characterizes the decrease in the biases b as Ψ increases.

In the following, we define a "sparsity parameter" α in the particular case where the biases are modeled as a power-law function of the PSF model complexity ($B \propto 1/\Psi^\alpha$), and we study the impact of α on the number of stars required to calibrate the PSF. We thus revisit the main result of P1 by deriving N_* (the number of stars required to calibrate the PSF) according to α instead of Ψ (the complexity of the underlying PSF). Moreover, this new relation is optimized to minimize $\sigma_{\rm sys}$.

We emphasize that, in this paper, we propose to optimize the complexity of the PSF model within a given the basis set. We do not address the issue of choosing the basis set itself. There is no doubt that, to optimize the PSF modeling, it is necessary to select carefully this basis set. For instance, generic basis sets such as shapelets, wavelets, or Fourier modes, although they have enormous advantages, are not optimal. This issue will be addressed in forthcoming works.

3. PSF calibration for shear measurement

When deconvolving the observed galaxy with the estimated PSF, $\delta(R^2_{\rm PSF})$ and $\delta\epsilon_{\rm PSF}$ propagate into an error $\delta\epsilon_{\rm gal}$ in the estimation of the galaxy ellipticity. We denote $R_{\rm gal}$ and $\epsilon_{\rm gal}$, the rms radius and the two-component ellipticity of the galaxy. When $R_{\rm PSF}$, $\epsilon_{\rm PSF}$, $R_{\rm gal}$, and $\epsilon_{\rm gal}$ are defined using the unweighted moments of the flux, this propagates to (see P1):

$$\delta\epsilon^{\rm sys}_{\rm gal} \simeq \left(\epsilon_{\rm gal} - \epsilon_{\rm PSF}\right) \frac{\delta\left(R^2_{\rm PSF}\right)}{R^2_{\rm gal}} - \left(\frac{R_{\rm PSF}}{R_{\rm gal}}\right)^2 \delta\epsilon_{\rm PSF}. \tag{8}$$

The spatial average of $|\delta\epsilon^{\rm sys}_{\rm gal}|^2$ is related to the variance in the systematic errors in the shear measurements $\sigma^2_{\rm sys}$ (Amara & Refregier 2007b), defined by the integral of the systematics $C^{\rm sys}_\ell$ in the power spectrum:

$$\sigma^2_{\rm sys} \equiv \frac{1}{2\pi} \int d(\ln \ell)\, \ell(\ell+1) \left|C^{\rm sys}_\ell\right| \tag{9}$$

$$\equiv (P^\gamma)^{-2} \left\langle \left|\delta\epsilon^{\rm sys}_{\rm gal}\right|^2 \right\rangle \tag{10}$$

with P^γ the calibration factor between the gravitational shear and the ellipticity. Its value depends on the distribution of galaxy ellipticities and is typically about 1.84 (Rhodes et al. 2000). The brackets $<>$ denote a spatial average over the entire field. As in P1, we substitute Eq. (8) into Eq. (10) with the following simplifying assumptions:

1. The galaxy is not correlated with the PSF.
2. The error on the PSF ellipticity ($\delta\epsilon_{\rm PSF}$) and the PSF ellipticity itself ($\epsilon_{\rm PSF}$) are not correlated. This is warranted by the fact that, in the assumed small ellipticity regime, $\delta\epsilon_{\rm PSF}$ does not have any preferred direction, implying that $\langle \epsilon_{\rm PSF}.\delta\epsilon_{\rm PSF}\rangle = 0$.
3. We neglect correlations between the ellipticity and the inverse squared radius of the galaxy. This is reasonable for the PSF calibration in the small ellipticity regime.

With these simplifications, one can substitute Eq. (8) into Eq. (10) to obtain:

$$\sigma^2_{\rm sys} = (P^\gamma)^{-2} \left\langle \left(\frac{R_{\rm PSF}}{R_{\rm gal}}\right)^4 \right\rangle \times \left[\left\langle |\delta\epsilon_{\rm PSF}|^2 \right\rangle + \left[\left\langle \left|\epsilon_{\rm gal}\right|^2 \right\rangle + \left\langle |\epsilon_{\rm PSF}|^2 \right\rangle \right] \frac{\left\langle \left|\delta R^2_{\rm PSF}\right|^2 \right\rangle}{\left\langle R^4_{\rm PSF} \right\rangle} \right]. \tag{11}$$

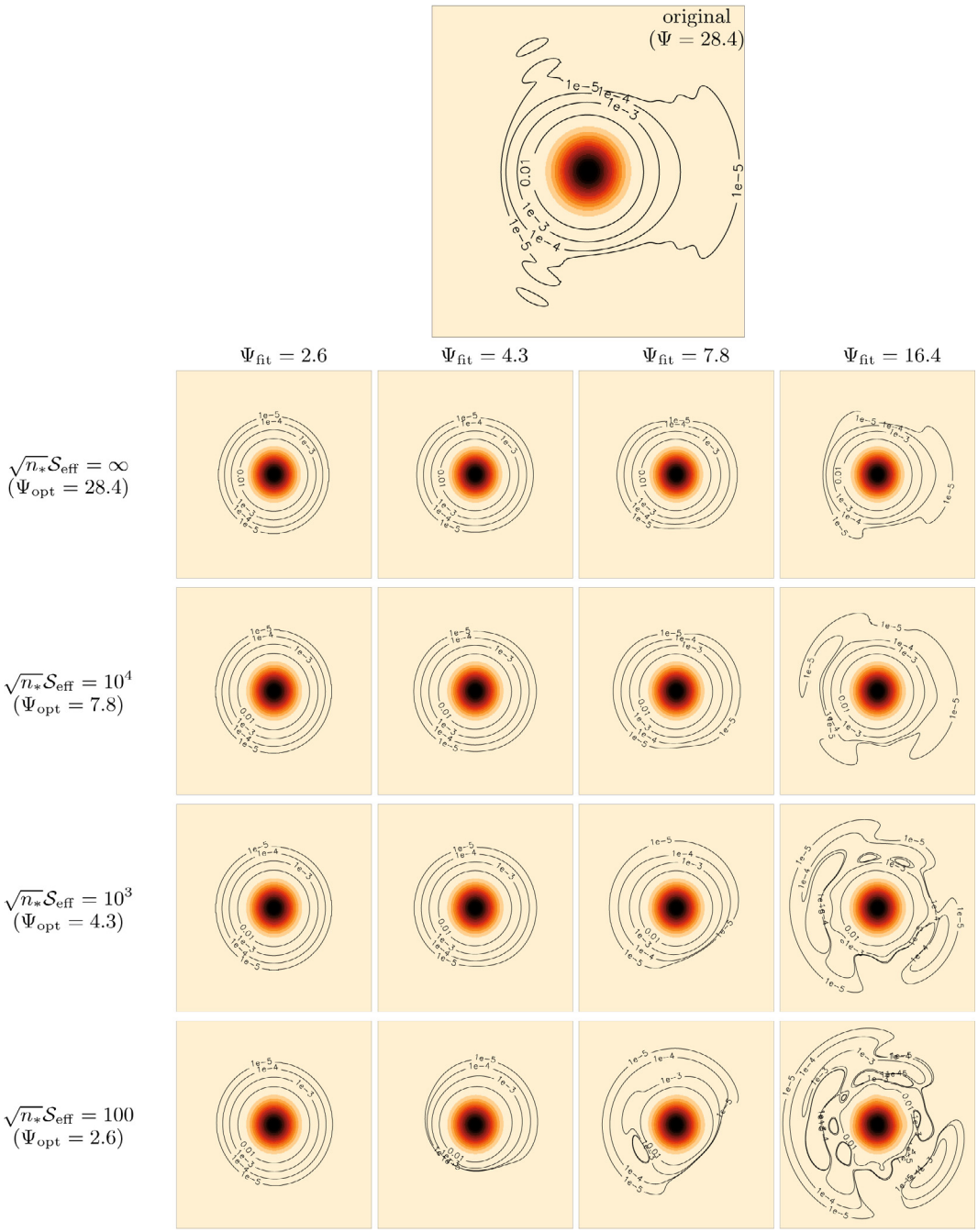

Fig. 1. PSF example (*top panel*) adopted in this paper and best fits of it (other panels) with 4 shapelet basis sets (corresponding to $\Psi_{\rm fit}$ = 2.6, 4.3, 7.8 and 16.4, i.e. $n_{\rm max}$ = 4, 6, 10 and 20 with the diamond option) and for $\sqrt{n_*}\mathcal{S}_{\rm eff}$ equal to 100, 10^3, 10^4 and ∞ (n_* is the numbers of stars used for the fit and $\mathcal{S}_{\rm eff}$ is the effective SNR of stars, see Eq. (3); infinity corresponds to the ideal case of no background). For a given value of $\sqrt{n_*}\mathcal{S}_{\rm eff}$, all fits are performed with the same realization of the noise. Colors show the flux (darker colors indicate brighter regions) and show that this PSF is almost circular at the center. Contours show some isophotes not visible with the color scale and reveal the complex structure of the tails. The original (i.e. underlying) PSF in the top panel was built using a model with Ψ = 24.8. The optimal value $\Psi_{\rm opt}$ of the fitted complexity (in order to minimize $\sigma_{\rm sys}$) is indicated under brackets for each value of $\sqrt{n_*}\mathcal{S}_{\rm eff}$. This figure illustrates that, for a given value of $\sqrt{n_*}\mathcal{S}_{\rm eff}$, the simpler the model (i.e. the lower $\Psi_{\rm fit}$), the poorer the description of the tails and the larger the bias. On the other hand, for a given $\Psi_{\rm fit}$, the lower the amount of information available in stars (i.e. the smaller $\sqrt{n_*}\mathcal{S}_{\rm eff}$), the noisier the description of the tails.

We develop a more compact expression by adopting the following notation:

$$C \equiv (P^\gamma)^{-2}\left\langle\left(\frac{R_{\rm PSF}}{R_{\rm gal}}\right)^4\right\rangle , \quad (12)$$

$$\mathcal{E} \equiv \left\langle\left|\epsilon_{\rm gal}\right|^2\right\rangle + \left\langle\left|\epsilon_{\rm PSF}\right|^2\right\rangle , \quad (13)$$

which leads to:

$$\sigma^2_{\rm sys} = C\left[\left\langle\left|\delta\epsilon_{\rm PSF}\right|^2\right\rangle + \frac{\mathcal{E}}{R^4_{\rm PSF}}\left\langle\left|\delta R^2_{\rm PSF}\right|^2\right\rangle\right] . \quad (14)$$

In P1, we considered only the scatters $\sigma[R^2_{\rm PSF}]$ and $\sigma[\epsilon_{\rm PSF}]$ (i.e. in the zero bias case: $b[\epsilon_i] = b[R^2] = 0$) and we approximate the statistical averages with spatial averages: $\sigma^2[R^2_{\rm PSF}] \simeq \left\langle\left|\delta R^2_{\rm PSF}\right|^2\right\rangle$

and $\sigma^2[\epsilon_{\rm PSF,i}] \simeq \left\langle \left|\delta\epsilon_{\rm PSF,i}\right|^2 \right\rangle$. In this paper, with the introduction of biases, the scatter becomes the MSE:

$$\begin{aligned} {\rm MSE}[R^2_{\rm PSF}] &\simeq \left\langle \left|\delta R^2_{\rm PSF}\right|^2 \right\rangle, \\ {\rm MSE}[\epsilon_{\rm PSF,i}] &\simeq \left\langle \left|\delta\epsilon_{\rm PSF,i}\right|^2 \right\rangle. \end{aligned} \tag{15}$$

This leads to:

$$\begin{aligned} \sigma^2_{\rm sys} \simeq C\Big[& b^2[\epsilon_{\rm PSF,1}] + \sigma^2[\epsilon_{\rm PSF,1}] \\ & + b^2[\epsilon_{\rm PSF,2}] + \sigma^2[\epsilon_{\rm PSF,2}] \\ & + \tfrac{\mathcal{E}}{R^4_{\rm PSF}}\left(b^2[R^2_{\rm PSF}] + \sigma^2[R^2_{\rm PSF}]\right)\Big]. \end{aligned} \tag{16}$$

We can see that $\sigma^2_{\rm sys}$ is proportional to the quadratic sum of 6 terms: three bias terms and three statistical ones. Collecting terms of similar type using the following notation:

$$B \equiv b^2\left[\epsilon_1^{\rm PSF}\right] + b^2\left[\epsilon_2^{\rm PSF}\right] + \mathcal{E}\frac{b^2[R^2_{\rm PSF}]}{R^4_{\rm PSF}}, \tag{17}$$

$$\Sigma \equiv \sigma^2\left[\epsilon_1^{\rm PSF}\right] + \sigma^2[\epsilon_2^{\rm PSF}] + \mathcal{E}\frac{\sigma^2[R^2_{\rm PSF}]}{R^4_{\rm PSF}}, \tag{18}$$

gives:

$$\sigma^2_{\rm sys} \simeq C\left[B + \Sigma\right]. \tag{19}$$

Although only Σ depends on the SNR of the stars, B and Σ both depend on the complexity of the modeling. They can be denoted $B(\Psi)$ and $\Sigma(\Psi, \mathcal{S}_{\rm eff})$. In P1, we show that it is given by:

$$\Sigma(\Psi, \mathcal{S}_{\rm eff}) = \frac{\Psi^2}{n_* \mathcal{S}^2_{\rm eff}}, \tag{20}$$

where the overall complexity of the model Ψ is given by the complexities ψ_ϵ and ψ_{R^2} associated with the ellipticity and the squared radius of the model, respectively (see Eq. (2)):

$$\Psi^2 = 2\psi^2_\epsilon + \mathcal{E}\psi^2_{R^2}. \tag{21}$$

Equations (19) and (20) then infer that:

$$\sigma^2_{\rm sys} \simeq C\left[B(\Psi) + \frac{\Psi^2}{n_*\,\mathcal{S}^2_{\rm eff}}\right]. \tag{22}$$

From this equation, we can see that increasing the complexity Ψ by adding degrees of freedom in the PSF model can reduce B but also increases the statistical errors. Minimizing $\sigma^2_{\rm sys}$ thus implies the search for the optimal trade-off in the value of Ψ.

4. Optimal PSF model

In Sect. 4.1, we present a PSF example that we use in the remainder of this paper to illustrate our discussion. In Sect. 4.2, we show the optimal complexity of the PSF model (that minimizes $\sigma_{\rm sys}$) and apply this to the PSF example. We then explore this optimization in more details in Sect. 4.3 by examining a particular case in which the bias can be described by a power-law function of the complexity.

4.1. PSF example

To illustrate our discussion, we study a realistic example of a PSF with complex features in the tails, and investigate what happens when fitting it with various shapelet basis sets as function of the SNR of the available stars. We also use a shapelet basis set (which differs from that used in the fits) for describing the underlying PSF. This use of shapelets for both the PSF model and the underlying PSF was chosen for three reasons:

- first, it allows pixelation issues to be ignored, which are beyond the scope of this paper. Indeed, the description of the underlying PSF is performed by using the continuous shapelet functions and the fits are performed at high resolution;
- second, it considerably simplifies both the calculations and the fitting process, due to the orthogonality of shapelet functions (the average estimation of a fitted coefficient is the true value, independently of the other coefficients);
- third, it is a simple and convenient framework for illustrating the use of sparsity as a tool in optimizing the complexity of the PSF modeling.

Our example of an underlying PSF is constructed using $n_{\rm max}$ = 34 (with the diamond option), as shown in Fig. 1.

In Fig. 1, we also show the 16 fits performed with the 4 shapelet basis sets (corresponding to $n_{\rm max}$ = 4, 6, 10, and 20) and $\sqrt{n_*}\mathcal{S}_{\rm eff}$ (the stellar signal-to-noise ratio, see Eq. (3)) equal to 100, 10^3, 10^4, or ∞ (the latter is the ideal case of no background). To determine the overall complexity Ψ, which depends on the rms of galaxy ellipticities in terms of the parameter $\mathcal{E}$ (see Eqs. (13) and (21)), we adopt the typical value $\mathcal{E} = 0.2$, for which $n_{\rm max}$ = 4, 6, 10, 20, 34 correspond to Ψ = 2.6, 4.3, 7.8, 16.4, 28.4 respectively. In the following, we also adopt the value $C = 0.066$, that corresponds to the typical values $P^\gamma = 1.84$ and $\left[\left\langle\left(\frac{R_{\rm gal}}{R_{\rm PSF}}\right)^4\right\rangle\right]^{1/4} = 1.5$. Figure 1 illustrates that:

- when $\sqrt{n_*}\mathcal{S}_{\rm eff}$ is sufficiently high, complex basis sets are required to model the complex tails, i.e. the amount of bias B decreases as the complexity Ψ of the model increases. For instance, a fit with $\mathcal{S}_{\rm eff} = \infty$ and Ψ = 28.4 would allow one to recover our PSF example exactly, with $B = 0$;
- a higher complexity requires a higher number of DoF to be fitted. Consequently, for a given value of $\sqrt{n_*}\mathcal{S}_{\rm eff}$, increasing the complexity of the model also increases the scatter in the estimated shape. Therefore, it is not always appropriate to use a complex fit model; it may be more robust to use a simplified (but more biased) fit model.

4.2. Optimizing the complexity of the PSF model

The optimal PSF model is that for which $\sigma_{\rm sys}$ is minimized, varying Ψ. We define the optimal value $\sigma^{\rm opt}_{\rm sys}$ to be the minimal value of $\sigma_{\rm sys}$ and in the same spirit, we note that $\Psi_{\rm opt}$ is the corresponding value of Ψ:

$$\sigma^{\rm opt}_{\rm sys} \equiv \sigma_{\rm sys} \text{ such that } \left.\frac{\partial(\sigma^2_{\rm sys})}{\partial\Psi}\right|_{\Psi_{\rm opt}} = 0. \tag{23}$$

For instance, Fig. 2 illustrates the search for the optimal shapelet basis when our PSF example (see the previous section) is estimated with 50 stars (and with $\mathcal{S}_{\rm eff}$ = 1000). The optimal model is that corresponding to $\Psi_{\rm opt} \simeq 6$ (i.e. $n_{\rm max}$ = 8 with the diamond option) and $(\sigma^{\rm opt}_{\rm sys})^2 \simeq 10^{-7}$, shown by the red diamond.

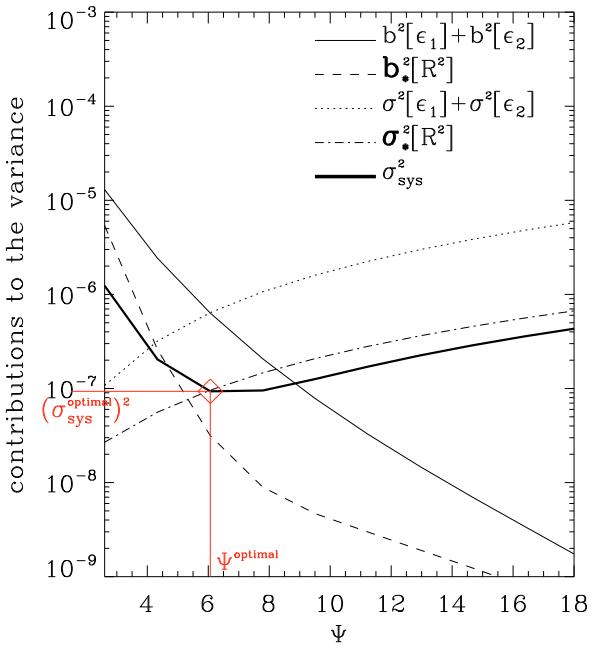

Fig. 2. Total variance $\sigma^2_{\rm sys}$ and its contributions (see Eq. (16)) with respect to the fit model complexity Ψ, for our PSF example (see Sect. 4.1 and Fig. 1) fitted with 50 stars (i.e. $n_* = 50$), and $\mathcal{S}_{\rm eff} = 1000$. We note $\mathbf{b}_*[R^2] = \sqrt{\mathcal{E}}\, b[R^2]/R^2$ and $\sigma_*[R^2] = \sqrt{\mathcal{E}}\,\sigma[R^2]/R^2$. The optimal basis set is that for which $\sigma_{\rm sys}$ is minimum. At this point, shown by the red diamond, $\sigma_{\rm sys} \equiv \sigma^{\rm opt}_{\rm sys} \simeq 10^{-7}$ and $\Psi \equiv \Psi_{\rm opt} \simeq 6$.

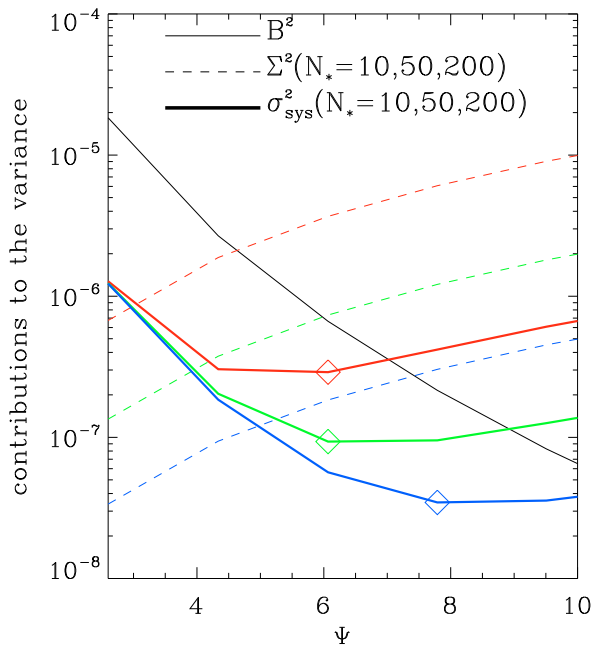

Fig. 3. Total variance $\sigma^2_{\rm sys}$ (3 thick and solid colored curves) and its contributions to B (one thin and solid black curve) and Σ (3 dashed colored curves), as defined by Eqs. (17) to (19), with respect to the PSF model complexity Ψ, for our PSF example (presented in Sect. 4.1 and Fig. 1) fitted with 10 (red curves), 50 (green curves) or 200 (blue curves) stars (i.e. $n_* = 10, 50, 200$) with an effective SNR of $\mathcal{S}_{\rm eff} = 1000$. Σ depends on n_* (see Eq. (20)), implying a different curve for each value of n_*, while B does not. This illustrates that $\sigma^{\rm opt}_{\rm sys}$ (i.e. the minimum value of $\sigma_{\rm sys}$ shown by the diamonds) increases with n_*.

For a given fit model, increasing n_* reduces the scatters but not the biases in the model fitting (see Eq. (22)). Therefore, as n_* increases, $\Psi_{\rm opt}$ increases and $\sigma^{\rm opt}_{\rm sys}$ decreases. This is illustrated in Fig. 3, which shows B, Σ, and $\sigma_{\rm sys}$ (see Eqs. (17) to (19)) when our PSF example is estimated with $n_* = 10$, 50, or 200 (still $\mathcal{S}_{\rm eff} = 1000$).

Figure 4 shows $\left(\sigma^{\rm opt}_{\rm sys}\right)^2$ as a function of n_*. The diamonds represent the curve for our PSF example illustrated in all previous plots, while the bold-straight line without any diamond shows the ideal case (addressed in P1) of a PSF described perfectly by the model (i.e. $B = 0$). Thus, $\left(\sigma^{\rm opt}_{\rm sys}\right)^2$ varies with $1/n_*$ as predicted by our scaling relation presented in P1. The dotted and dashed lines are discussed in Sect. 4.3.

4.3. Example of optimal complexity in the case of a power-law function

In this section, we derive the optimal complexity when the bias B is a power-law function of the complexity written as $B \propto 1/\Psi^\alpha$. We investigate $\left(\sigma^{\rm opt}_{\rm sys}\right)^2$ as a function of n_* and α. We normalize the power-law function, such as:

$$B(\Psi) \equiv B_0 \left(\frac{\Psi_0}{\Psi}\right)^\alpha . \tag{24}$$

In our example of a PSF fitted with a shapelet basis set (see Sect. 4.1 and Fig 1), the smallest value of Ψ that we consider is 2.6 (which corresponds to $n_{\rm max} = 4$ with the "diamond" configuration, see P1). This explains why, for this example, we choose to normalize the power-law function to this value $\Psi_0 = 2.6$, implying $B_0 = 2 \times 10^{-5}$. This model is illustrated in Fig. 5 for α equal to 2, 4, 6, and 8. This is superimposed on the B versus Ψ relation that we obtain when fitting our PSF example with different shapelet basis-sets, as decribed in Sect. 4.1. We see that in this case B is reasonably described by $\alpha = 4$.

This representation by a power-law function is particularly convenient because α can be identified with the sparsity: a high value of α means that the PSF model is efficient in representing the underlying PSF with a small number of free parameters. Conversely, a low value of α means the PSF model requires a large number of parameters to describe the underlying PSF without large residuals. In the following, α is called the "sparsity parameter". Together with this power-law representation (Eq. (24)), Eqs. (22) and (23) imply that:

$$\left(\sigma_{\rm sys}(\Psi)\right)^2 = C\left[B_0\left(\frac{\Psi_0}{\Psi}\right)^\alpha + \frac{\Psi^2}{n_* \mathcal{S}^2_{\rm eff}}\right], \tag{25}$$

$$\Psi_{\rm opt} = \left[\alpha\, B_0\, \Psi_0^\alpha\, n_*\, \mathcal{S}^2_{\rm eff}/2\right]^{1/(\alpha+2)}, \tag{26}$$

$$\left(\sigma^{\rm opt}_{\rm sys}\right)^2 = \left(\sigma_{\rm sys}\left(\Psi = \Psi_{\rm opt}\right)\right)^2 = C\left[B_0^2\left(\frac{2\,\Psi_0^2}{\alpha\, n_* \mathcal{S}^2_{\rm eff}}\right)^\alpha\right]^{1/(\alpha+2)} . \tag{27}$$

Note that Eq. (27) expresses $\left(\sigma^{\rm opt}_{\rm sys}\right)^2$ (the minimum variance in the systematic errors in shear measurements that can be achieved, see Eqs. (9), (10), and (23)) in terms of a set of parameters that can be divided into 2 families:

1. parameters that are properties of the data set, such as C, B_0, Ψ_0, and $\mathcal{S}_{\rm eff}$;

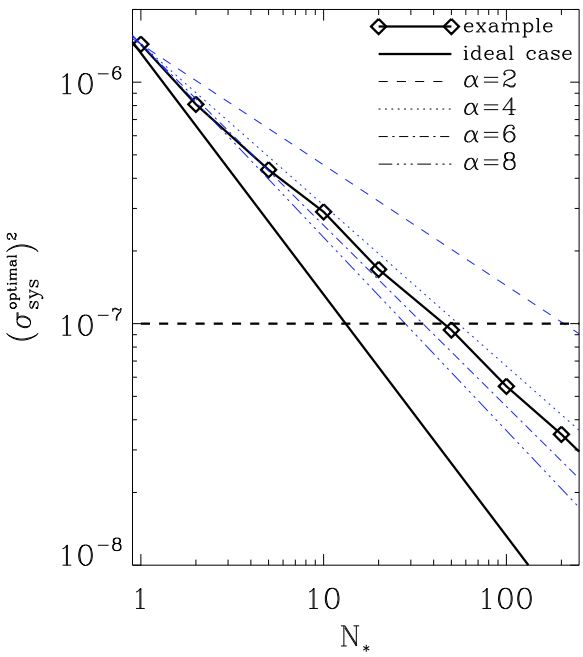

Fig. 4. Optimal variance of the shape measurement systematics $\left(\sigma_{\rm sys}^{\rm opt}\right)^2$ as a function of the number of stars n_* used to calibrate the PSF. The diamonds show the curve for our PSF example illustrated in all previous plots (presented in Fig. 1) and fitted with shapelets. The straight bold line shows the ideal case (addressed in P1) of a PSF model that exactly describes the PSF (i.e. with no residual). The horizontal line shows the values $\left(\sigma_{\rm sys}^{\rm opt}\right)^2 = 10^{-7}$ which is the requirement to be able to constraint w_0 and w_a at 0.02 and 0.1 respectively (Amara & Refregier 2007b). The blue lines (dashed, dotted and dotted-dashed) are discussed in Sect. 4.3. They are the curves expected when modeling the bias B with a power-law function of the complexity as stated by Eq. (24) and illustrated in Fig. 5.

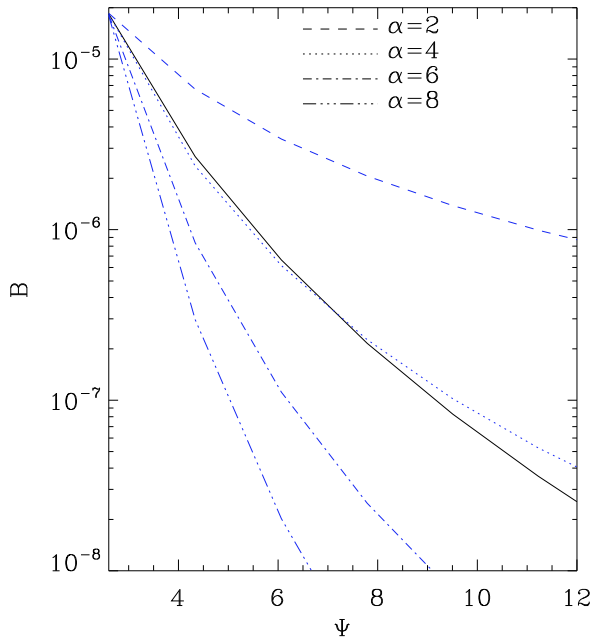

Fig. 5. Overall bias B versus Ψ for our PSF example (in black) and some power-law functions $B \propto 1/\Psi^{\alpha}$, $\alpha = 2, 4, 6, 8$ (in blue) normalized to intercept at $\Psi = 1.6$. We see that the case $\alpha = 4$ fits well with the example.

2. parameters that are properties of the analysis method, such as n_* (i.e. the number of stars used to calibrate the PSF) and α (i.e. the sparsity parameter of the PSF model).

When analysing a given data set, the parameters in the first family are kept fixed. The only parameters that can vary to achieve optimization during the analysis are those in the second family (i.e. n_* and α). For a given n_*, $\left(\sigma_{\rm sys}^{\rm opt}\right)^2$ is proportional to $\alpha^{-\alpha/(\alpha+2)}$.

5. Required number of stars

As discussed in the introduction, an important issue for cosmic shear surveys is to ensure that systematics are kept smaller than the statistical errors, by demanding an upper limit to $\sigma_{\rm sys}$. Part of the systematics have their origin in the PSF calibration, which is imperfect due to the limited number of stars available. In this section, we express N_*, the number of stars required to calibrate the PSF, in terms of the level of systematic errors $\sigma_{\rm sys}$ (note the capital "N", as opposed to "n_*" which is the number of stars involved in the PSF calibration process: we need $n_* \geq N_*$ to ensure that systematic effects are below $\sigma_{\rm sys}$. N_* is the lower limit of n_*).

In Sect. 5.1, we summarize the conclusions of P1 which apply when the underlying PSF and the PSF model have the same functional form (i.e. $B = 0$) and we extend these conclusions to the general case of PSF modeling performed with any model (i.e. B not necessarily equal to 0). In Sect. 5.2, we invert Eq. (27) (that holds when B is described by a power-law function of the complexity: $B \propto 1/\Psi^{\alpha}$) and express N_* as a function of α and of the minimum systematic level $\sigma_{\rm sys}^{\rm opt}$ achievable when the complexity of the PSF modeling is optimal.

5.1. Generalised scaling relation

In the optimistic case where the PSF calibration is the only significant source of systematic errors, a given value of N_* (i.e. a given number of stars involved in the PSF calibration) implies a value of $\sigma_{\rm sys}$. This is presented in P1 in the form of a scaling relation that links N_*, $\sigma_{\rm sys}$, $\mathcal{S}_{\rm eff}$ (the effective signal-to-noise ratio of stars), $\left(R_{\rm gal}/R_{\rm PSF}\right)_{\rm min}$ (the ratio between the smallest galaxy size and the PSF size), and Ψ (the complexity of the PSF):

$$N_* \simeq 50\left(\frac{\mathcal{S}_{\rm eff}}{500}\right)^{-2}\left(\frac{\left(R_{\rm gal}/R_{\rm PSF}\right)_{\rm min}}{1.5}\right)^{-4}\left(\frac{\sigma_{\rm sys}^2}{10^{-7}}\right)^{1}\left(\frac{\Psi}{3}\right)^{2}/2. \tag{28}$$

The factor 2 at the end comes from the fact that $\Psi^2 \simeq 2\psi_\epsilon^2$ (in P1, this scaling relation is written in terms of ψ_ϵ). This holds for the assumption that the PSF model is able to describe the PSF without any bias (i.e. $B = 0$). With non-zero B and adopting the same simplifications and the same typical values as in P1, Eq. (22) leads to the more general relation:

$$N_* \simeq 50\left(\frac{\mathcal{S}_{\rm eff}}{500}\right)^{-2}\left(\frac{\left(R_{\rm gal}/R_{\rm PSF}\right)_{\rm min}}{1.5}\right)^{-4}\left(\frac{\sigma_{\rm sys}^2}{10^{-7}}\right)^{1}\Phi^2\,, \tag{29}$$

where

$$\Phi^2 = \frac{(\Psi/3)^2}{2\left[1 - C\frac{B}{\sigma_{\rm sys}^2}\right]}\,. \tag{30}$$

Thus, taking B into account in the scaling relation translates into the new factor $1/\left[1 - C\frac{B}{\sigma_{\rm sys}^2}\right]$ in Eq. (30), which equals 1 when B

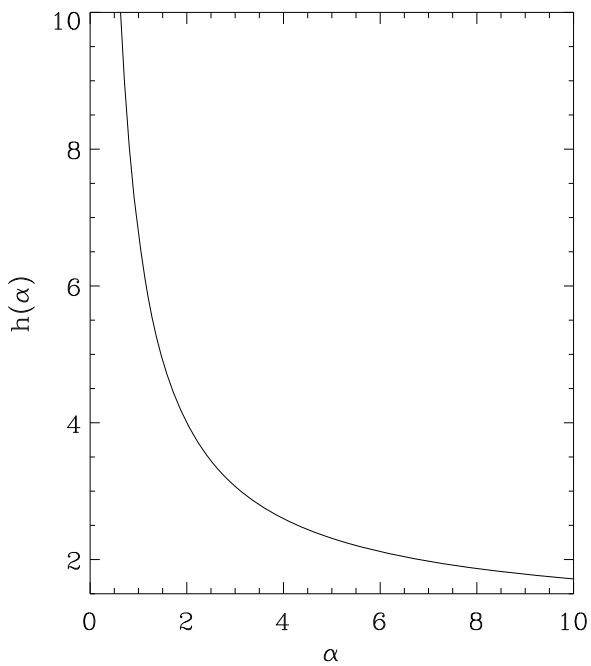

Fig. 6. $h(\alpha)$ as defined in Eq. (32).

is zero (then the relation (29) is equivalent to the scaling relation given in P1) and is related to the ratio $\frac{B}{\sigma^2_{\rm sys}}$, which is the relative weight of biases in the error budget.

5.2. Application to the power-law model

Equation (27) can be inverted to provide N_* (the number of stars required to calibrate the PSF) as a function of $\sigma^{\rm opt}_{\rm sys}$ (the minimum level of systematics achievable when optimizing the complexity of the PSF modeling), α (the sparsity parameter), $\mathcal{S}_{\rm eff}$ (the effective SNR of stars defined in Eq. (3)), and C (a dimensionless factor defined in Eq. (12)):

$$N_* = \frac{\Psi_0^2 B_0^{2/\alpha}}{\mathcal{S}^2_{\rm eff}} h(\alpha) \left[\frac{C}{\left(\sigma^{\rm opt}_{\rm sys}\right)^2}\right]^{(1+2/\alpha)} \tag{31}$$

with the dimensionless function:

$$h(\alpha) = \left(\frac{\alpha}{2}\right)^{2/\alpha} \left(1 + \frac{2}{\alpha}\right)^{1+2/\alpha} \tag{32}$$

shown in Fig. 6. With the notation and scaling of Eq. (29), Eq. (31) is equivalent to:

$$N_* = 16 \left(\frac{\Psi_0}{2.6}\right)^2 \left(12\,\frac{B_0}{2\times 10^{-5}}\right)^{2/\alpha} h(\alpha) \left(\frac{\mathcal{S}_{\rm eff}}{500}\right)^{-2} \times \left(\frac{\left(R_{\rm gal}/R_{\rm PSF}\right)_{\rm min}}{1.5}\right)^{-4(1+2/\alpha)} \left(\frac{\left(\sigma^{\rm opt}_{\rm sys}\right)^2}{10^{-7}}\right)^{-(1+2/\alpha)} . \tag{33}$$

This equation allows one to estimate the number of stars required to calibrate the PSF and thus, with respect to the stellar density, the minimum scale on which the PSF calibration is possible. On smaller scales, stars provide insufficient information for calibrating PSF variations. This implies that these smaller scales may be contaminated by systematics due to a poor correction of the PSF and should not be used to estimate cosmological parameters, unless the variabilities on small scales are known to be extremely small. As shown by Amara & Refregier (2007b) and discussed in P1, future all-sky cosmic shear surveys will need to achieve $\left(\sigma^{\rm opt}_{\rm sys}\right)^2 \leq 10^{-7}$ to be able to estimate w_0 and w_a with uncertainties of about 0.02 and 0.1, respectively. Fig. 4 shows that, for our PSF example, it is possible to achieve this accuracy when calibrating the PSF with 50 stars, if $\alpha \geq 4$. Although this is not a general statement (this assumes that B can be described by a power-law function, see Eq. (24), and depends on the normalization parameter B_0), this is a representative example of the sparsity requirement for future cosmic shear surveys. On the other hand, for current cosmic shear surveys of areas ~50 deg^2, we have the requirement $\left(\sigma^{\rm opt}_{\rm sys}\right)^2 \leq 4\times 10^{-6}$ (see Amara & Refregier (2007b) and P1). In this case, it is possible to calibrate the PSF with a few stars when $\alpha \geq 2$. This sparsity requirement is reached with the current PSF correction methods (for instance that based on shapelets as in Bergé et al. 2008), and, assuming a star density of about 1 per arcmin^2, this is consistent with the presence of significant B modes usually found on scales smaller than a few arcmins.

6. Conclusions

We explore the systematics induced in cosmic shear by the PSF calibration/correction process and study how to optimize the PSF model to minimize the systematic errors in cosmological parameter estimations. In this framework, we revisit the concept of the complexity of the PSF, defined in our previous paper (P1), and introduce the concept of the "sparsity" of the PSF model. The complexity Ψ characterizes the number of degrees of freedom in the model. A small number of degrees of freedom corresponds to a low Ψ and relates to a simple PSF model, which can be fitted to the stellar observations of low signal-to-noise ratio, but is likely to be highly biased. On the other hand, a large number of degrees of freedom corresponds to a high Ψ and relates to a complex PSF model, which is expected to have a low bias but requires stellar observations of high signal-to-noise ratio to avoid large statistical scatters in the fitted parameters. In P1, we related the complexity Ψ of the PSF model to the systematic errors in cosmological parameter estimations. In this paper, we show how the complexity can be optimized depending on the stars available, using the concept of sparsity. The sparsity characterizes the decrease of residuals between the best fit of the PSF model on the underlying PSF, when adding degrees of freedom to the model.

In the general case, we also extend the scaling relation, proposed in P1, between the number of stars used to calibrate the PSF and the systematic errors in the cosmological parameter estimations.

As discussed in P1, this relation, with the constraint of maintaining the systematics below the statistical uncertainties when estimating cosmological parameters, infers the number of stars N_* required for the PSF calibration. N_* corresponds to the minimum scale on which the PSF modeling is accurate: on scales smaller than this minimum, there is insufficent information in the data to calibrate the PSF variations. This implies that these smaller scales may be contaminated by systematics related to a poor PSF correction and should not be used when estimating cosmological parameters (unless the variabilities are known to be small, due, for instance, to the quality of the hardware).

We consider a realistic PSF example and model the amount of bias B between the PSF fit and the underlying PSF by a power-law function of the fitted complexity: $B \propto 1/\Psi^{\alpha}$ where α is the sparsity parameter. We find that, for this PSF, current cosmic shear analyses that cover 50 deg^2 or less, need α to be higher than 2, which is achievable by current analysis methods. Thus, current cosmic shear analyses do not require a rigorous optimization of the PSF model. On the other hand, future cosmic shear surveys that aim to measure w_0 and w_a to an accuracy of 0.02 and 0.1, respectively, will require $\alpha \geq 4$ to calibrate the PSF with 50 stars. This relation between the required number of stars N_* and the accuracy of the calibration depends on the underlying PSF. This explains why these values, although corresponding to realistic orders of magnitude, cannot be assumed to represent a general result. Two parameters drive this relation: the amount of biases B_0 when fitting the underlying PSF with a PSF model of low complexity (N_* being proportional to $B_0^{2/\alpha}$; in our example, $B_0 = 2 \times 10^{-5}$), and the sparsity parameter α of the PSF modeling during the analysis.

It is thus possible to optimize cosmic shear surveys at two levels: when optimizing the observational conditions, the PSF must be as simple and stable as possible in order to make possible its description by a low complexity model (this minimizes B_0); when analysing the data, the PSF modeling must be optimized to have as high a value of the sparsity α as possible.

The approach suggested in this paper is a first step toward introducing the concept of sparsity to weak lensing shape measurements. We do not address issues related to the pixelation. Moreover, although we only address the PSF calibration, this approach is also applicable to other topics such as description of galaxy shapes.

Acknowledgements. We thank Jean-Luc Starck and Jérome Bobin for useful discussions and insight on sparsity. We also thank Sarah Bridle and Lisa Voigt for an ongoing collaboration on weak lensing shape measurements. SPH is supported by the P2I program, contract number 102759. AA is supported by the Swiss Institute of Technology through a Zwicky Prize.

References

Amara, A., & Refregier, A. 2007a, MNRAS, 381, 1018
Amara, A., & Refregier, A. 2007b, [arXiv:0710.5171]
Benjamin, J., Heymans, C., Semboloni, E., et al. 2007, MNRAS, 381, 702
Bergé, J., Pacaud, F., Réfrégier, A., et al. 2008, MNRAS, 385, 695
Bridle, S., Shawe-Taylor, J., Amara, A., et al. 2008, [arXiv:0802.1214]
Fu, L., Semboloni, E., Hoekstra, H., et al. 2008, A&A, 479, 9
Heymans, C., VanWaerbeke, L., Bacon, D., et al. 2006, MNRAS, 139, 313
Massey, R., & Refregier, A. 2005, MNRAS, 363, 197
Massey, R., Heymans, C., Bergé, J., et al. 2007, MNRAS, 376, 13
Paulin-Henriksson, S., Amara, A., Voigt, L., Refregier, A., & Bridle, S. L. 2008, A&A, 484, 67
Rhodes, J., Refregier, A., & Groth, E. J. 2000, ApJ, 536, 79
Rhodes, J., Refregier, A., & Groth, E. J. 2008, in preparation