Astronomy & Astrophysics

Point spread function calibration requirements for dark energy from cosmic shear

S. Paulin-Henriksson[1], A. Amara[1], L. Voigt[2], A. Refregier[1], and S. L. Bridle[2]

[1] Service d'Astrophysique, CEA Saclay, Batiment 709, 91191 Gif-sur-Yvette Cedex, France
[2] Department of Physics & Astronomy, University College London, London, WC1E 6BT, UK

ABSTRACT

Context. The control of systematic effects when measuring background galaxy shapes is one of the main challenges for cosmic shear analyses.

Aims. Study the fundamental limitations on shear accuracy due to the measurement of the point spread function (PSF) from the finite number of stars that are available. We translate the accuracy required for cosmological parameter estimation to the minimum number of stars over which the PSF must be calibrated.

Methods. We characterise the error made in the shear arising from errors on the PSF. We consider different PSF models, from a simple elliptical Gaussian to various shapelet parametrisations. First we derive our results analytically in the case of infinitely small pixels (i.e. infinitely high resolution), then image simulations are used to validate these results and investigate the effect of finite pixel size in the case of the elliptical Gaussian PSF.

Results. Our results are expressed in terms of the minimum number of stars required to calibrate the PSF in order to ensure that systematic errors are smaller than statistical errors when estimating the cosmological parameters. On scales smaller than the area containing this minimum number of stars, there is not enough information to model the PSF. This means that these small scales should not be used to constrain cosmology unless the instrument and the observing strategy are optimised to make this variability extremely small. The minimum number of stars varies with the square of the star Signal-to-Noise Ratio, with the complexity of the PSF and with the pixel size. In the case of an elliptical Gaussian PSF and in the absence of dithering, 2 pixels per PSF full width at half maximum (*FWHM*) implies a 20% increase of the minimum number of stars compared to the ideal case of infinitely small pixels; 0.9 pixels per PSF *FWHM* implies a factor 100 increase.

Conclusions. In the case of a good resolution and a typical Signal-to-Noise Ratio distribution of stars, we find that current surveys need the PSF to be calibrated over a few stars, which may explain residual systematics on scales smaller than a few arcmins. Future all-sky cosmic shear surveys require the PSF to be calibrated over a region containing about 50 stars. Due to the simplicity of our models these results should be interpreted as optimistic and therefore provide a measure of a systematic "floor" intrinsic to shape measurements.

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

1. Introduction

Gravitational lensing by large scale structure (or "cosmic shear") has grown rapidly as a research field over the last decade (Refregier 2003b; Hoekstra 2003; Munshi et al. 2006). In the coming years, we can expect that the field will continue to make important contributions to cosmology. For instance, Albrecht et al. (2006) and Peacock & Schneider (2006) have singled out cosmic shear as potentially the most powerful probe for constraining dark energy.

A great deal of work is currently under way to develop techniques that will allow the maximum possible potential to be reached and not limited by systematic measurement errors. This can be done by designing and building instruments with weak lensing as the central science driver. This approach leads to a top-down mission design process which begins with science requirements, which are translated into technical requirements and then into an instrument design. This method for tackling the problem of high accuracy weak lensing measurements is fundamentally different to the bottom-up approach that is currently used. In the latter approach an existing telescope is used and the science team is assigned the task of extracting the maximum possible information from their surveys (until they hit the systematic limit of the instrument).

The field of gravitational lensing lends itself most naturally to top-down design, since most (but not all) of the systematic errors are not astronomical but are associated with the instrument and the atmosphere. These can therefore be controlled through instrument design and an optimised survey strategy. For this reason, many of the ambitious future imaging surveys that are currently under development have placed weak lensing as primary science driver, including the Dark UNiverse Explorer[1] (DUNE), the SuperNovae Acceleration Probe[2] (SNAP), the Panoramic Survey Telescope & Rapid Response System[3] (Pan-STARRS), the Dark Energy Survey[4] (DES) and the Large Synoptic Survey Telescope[5] (LSST). The issue is then to establish the instrumental requirements needed to reach the full statistical potential of the survey.

In Amara & Refregier (2007a), a wide range of survey parameters, such as area and depth, were considered. Their impact on the statistical potential of a lensing survey was calculated and summarised in a scaling relation that can be used for survey

---

[1] http://www.dune-mission.net
[2] http://snap.lbl.gov/
[3] http://pan-starrs.ifa.hawaii.edu
[4] http://www.darkenergysurvey.org
[5] http://www.lsst.org

designs to trade-off one property of a survey against another. In brief, this work finds that, once the median redshift of a survey is sufficiently high ($z \gtrsim 0.7$), the optimal survey strategy is to make the lensing survey as wide as possible. Similar results have been found by Heavens et al. (2006). In the same spirit, Amara & Refregier (2007b) looked at the requirements an ultra-wide field survey places on the control of systematics and concluded with a set of scaling relations that show the tolerance on residual systematic errors as a function of survey parameters. In particular, two types of shape measurement systematics were considered: multiplicative and additive systematics which are, respectively, correlated and uncorrelated with the lensing signal. Huterer et al. (2006) have also studied the impact of multiplicative and additive errors.

In the present study, we link the above systematic requirements to errors associated with measurement of the point spread function (PSF) of the instrument. Indeed, since galaxy images need to be PSF-corrected before shapes can be measured, errors in the estimation of the PSF are propagated into errors on galaxy shapes and can mimic the shear signal. The PSF calibration is driven by 3 factors:

1. **The PSF model**: the model that is chosen to describe the PSF will never perfectly describe the true PSF hence the choice of PSF model introduces an "a priori" systematic.
2. **The interpolation scheme**: the PSF will vary over the field which means that the PSF at a galaxy position needs to be interpolated from the PSFs of the nearby stars.
3. **The finite information available from each star**: a star provides an image of the PSF that is noisy and pixelated. Thus, the PSF information that we are able to extract from each star is finite. As a result, to reach high precision, it is necessary to combine the information from several stars.

This paper is devoted to the last factor. We quantify the accuracy of the PSF calibration with analytical predictions in the case of infinitely small pixels and use image simulations to quantify the pixelation effects. Previous work has also studied the impact that PSF errors are likely to induce in cosmic shear measurements. For instance, Stabenau et al. (2007) looked specifically at the PSF from the current SNAP design and translated this into systematic errors. The study that we present here sets out to be more general and uses simplified PSFs to try and quantify the systematic floor of an instrument. We investigate the extent to which the pixel scale will degrade the amount of information relative to our analytic predictions but we do not explore the optimal methods for combining multiple exposures, such as the one proposed by Jain et al. (2006). High et al. (2007) have studied the impact of the pixel scale on the number of useful galaxies (assuming the PSF is known perfectly). While they find a weak dependence, we expect that the driving factor for pixel scale is shape measurement systematics rather than statistical errors.

For a survey with a given statistical potential we use the results of Amara & Refregier (2007b) to find the upper limit on the systematic errors in the shears such that the induced bias in cosmological parameters remain subdominant compared to the statistical (marginalised Fisher matrix) errors when estimating cosmological parameters. We then convert the requirements on shear systematics into the minimum number of stars (N_*) needed to measure the PSF to the level of accuracy required. On scales smaller than the area containing N_*, there is not enough information from the stars to calibrate the PSF. Therefore, the systematic errors in the cosmological parameters can be dominant over the statistical ones. This means that these small scales should not be used to constrain cosmology unless the PSF variability is known
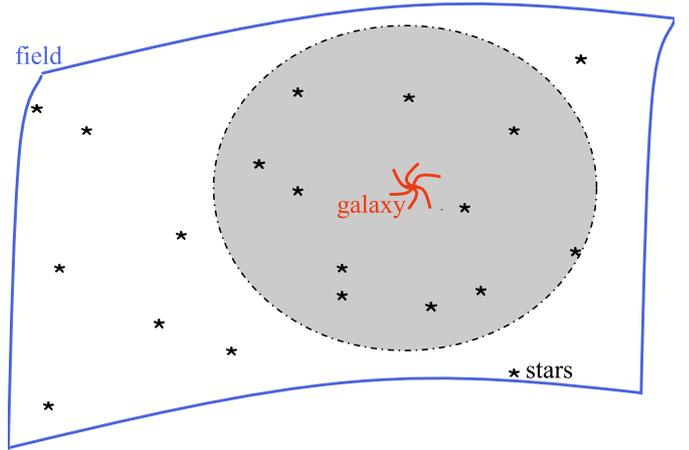
Fig. 1. Illustration of the required number of stars N_* characterised in this paper. When measuring the shape of a galaxy (marked by a red spiral), the PSF needs to be calibrated with at least N_* nearby stars (black asterisks) contained in the shaded region. In this example $N_* = 11$. On scales smaller than this there is not enough information coming from the stars to measure the PSF variations.

to be extremely small. N_* clearly depends on the Signal-to-Noise Ratio distribution of stars, since bright ones contain more information about the PSF than faint noisy ones. It also depends on the stability and the complexity of the PSF, characterised by the number of degrees of freedom that must be estimated or interpolated from the stars. Each degree of freedom increases N_*.

Figure 1 shows a graphical example to illustrate the calibration requirement we are considering. We see a central galaxy surrounded by several stars. The grey shaded region shows the area over which the PSF must be calibrated and in this example contains 11 stars. Stars do not provide enough information to model the PSF variations on scales smaller than this area.

This paper is organised as follows. Section 2 explains the weak lensing and cosmological context and summarises the issue of accurate PSF calibration. Section 3 shows the analytical predictions of the PSF calibration accuracy, expected in the case of infinitely small pixels. Section 4 describes the simulations we use to validate these predictions and extend them to finite pixels. In Sect. 5 we give the final accuracy of the PSF calibration and derive the number of stars it requires.

2. Weak lensing

2.1. Shear measurement

Weak gravitational shear is locally estimated using the shapes of background galaxies (for a review, see e.g. Bartelmann & Schneider 2001). Shear estimation methods can be divided into two families: those computing an estimator of the shear from a set of weighted sums over pixel values, for instance the common "KSB+" method (Kaiser et al. 1995; Luppino & Kaiser 1997; Hoekstra et al. 1998), and those fitting a model to the observed galaxy shape and deriving an estimator from the model. This includes the methods proposed by Kuijken (1999); Bernstein & Jarvis (2002); Refregier et al. (2002) and Voigt & Bridle (2008).

In this paper, we need a general formalism to propagate the error on the PSF into an error on the shear estimate. For this purpose, we consider as shape parameters the 2 component ellipticity ϵ and the squared radius R^2, both defined using the unweighted second order moments of the galaxy. For an object with

surface brightness $f(x_1, x_2)$, the total flux $F^{(0)}$ is the zeroth order moment of the surface brightness:

$$F^{(0)} = \int \mathrm{d}^2\boldsymbol{x}\, f(\boldsymbol{x}), \tag{1}$$

the centroid $\boldsymbol{x}^{\mathrm{cen}}$ is given by the first order moments divided by the total flux:

$$x_i^{\mathrm{cen}} = \frac{F_i^{(1)}}{F^{(0)}} = \frac{1}{F^{(0)}} \int \mathrm{d}^2\boldsymbol{x}\, x_i\, f(\boldsymbol{x}) \tag{2}$$

and the quadrupole moment matrix is given by the second order moments divided by the total flux:

$$Q_{ij} = \frac{F_{ij}^{(2)}}{F^{(0)}} = \frac{1}{F^{(0)}} \int \mathrm{d}^2\boldsymbol{x}\, (x_i - x_i^{\mathrm{cen}})(x_j - x_j^{\mathrm{cen}})\, f(\boldsymbol{x}). \tag{3}$$

The square rms radius R^2 and the 2 component ellipticity $\boldsymbol{\epsilon} = [\epsilon_1, \epsilon_2]$ are defined by:

$$R^2 \equiv Q_{11} + Q_{22}, \tag{4}$$

$$\epsilon_1 \equiv \frac{Q_{11} - Q_{22}}{Q_{11} + Q_{22}}, \quad \epsilon_2 \equiv \frac{2Q_{12}}{Q_{11} + Q_{22}}, \tag{5}$$

$$\epsilon \equiv \sqrt{\epsilon_1^2 + \epsilon_2^2}. \tag{6}$$

Consider now we have unbiased estimators of R^2 and $\boldsymbol{\epsilon}$ with variations δR^2 and $\boldsymbol{\delta\epsilon}$ around the true values. We adopt the following definitions:

$$\left\langle \left| \delta R^2 \right|^2 \right\rangle \equiv \sigma^2[R^2], \tag{7}$$

$$\left\langle \left| \boldsymbol{\delta\epsilon} \right|^2 \right\rangle \equiv 2\sigma^2[\epsilon]. \tag{8}$$

The factor 2 in Eq. (8) is due to the fact that ϵ has 2 components and we define $\sigma[\epsilon_{\mathrm{PSF}}]$ as the standard deviation of one of the components.

The weak gravitational shear $\boldsymbol{\gamma} = [\gamma_1, \gamma_2]$ can then be shown to be estimated using:

$$\widehat{\boldsymbol{\gamma}} = (P^\gamma)^{-1}\, \boldsymbol{\epsilon}_{\mathrm{gal}}, \tag{9}$$

where

$$P^\gamma = 2 - \left\langle \left| \boldsymbol{\epsilon}_{\mathrm{gal}} \right|^2 \right\rangle \approx 1.84. \tag{10}$$

The symbol "$\widehat{}$" indicates an estimator so that $\langle \widehat{\boldsymbol{\gamma}} \rangle = \boldsymbol{\gamma}$, P_γ is the shear susceptibility and the subscript "gal" corresponds to values measured on a galaxy. The value of $\sqrt{\langle | \boldsymbol{\epsilon}_{\mathrm{gal}} |^2 \rangle} \approx 0.4$ comes from the typical ellipticity distribution in current data sets.

In practice, the measurement of the galaxy ellipticity is uncertain because of the noise in the image. We write the induced error as $\boldsymbol{\delta\epsilon}^{\mathrm{noise}}$, which has a null average $\langle \boldsymbol{\delta\epsilon}^{\mathrm{noise}} \rangle = 0$. There may also be a systematic effect $\boldsymbol{\delta\epsilon}^{\mathrm{sys}}$ so that the estimated ellipticity is

$$\widehat{\boldsymbol{\epsilon}}_{\mathrm{gal}} = \boldsymbol{\epsilon}_{\mathrm{gal}} + \boldsymbol{\delta\epsilon}^{\mathrm{sys}} + \boldsymbol{\delta\epsilon}^{\mathrm{noise}}. \tag{11}$$

In practice $\boldsymbol{\delta\epsilon}^{\mathrm{sys}}$ may be due to several factors. For instance, the STEP collaboration (Heymans et al. 2006; Massey et al. 2007; Rhodes et al. 2008) investigate the contribution to $\boldsymbol{\delta\epsilon}^{\mathrm{sys}}$ from imperfections in the shear measurement methods. In a different spirit, this paper focusses on the contribution to $\boldsymbol{\delta\epsilon}^{\mathrm{sys}}$ from the limited information available on the PSF due to the photon noise

in the star images. We identify the variance $\langle |\boldsymbol{\delta\epsilon}^{\mathrm{sys}}|^2 \rangle$ with the quantity σ_{sys}^2 examined in Amara & Refregier (2007b), which indicates the variance of the systematic errors

$$\sigma_{\mathrm{sys}}^2 \equiv (P^\gamma)^{-2} \left\langle \left| \boldsymbol{\delta\epsilon}^{\mathrm{sys}} \right|^2 \right\rangle. \tag{12}$$

We translate an upper limit on σ_{sys}^2 into constraints on the number stars that are needed to calibrate the PSF.

2.2. Systematics from PSF calibration

Since a galaxy image needs to be PSF corrected before its shape can be measured, errors in the estimation of the PSF propagate into an error $\boldsymbol{\delta\epsilon}^{\mathrm{sys}}$ in the measured ellipticity. Here, we consider errors in the PSF radius parameter R_{PSF} and in the 2 component PSF ellipticity $\boldsymbol{\epsilon}_{\mathrm{PSF}}$. To first order and for the unweighted moments (definitions given by Eqs. (3) to (6)), we have (see Appendix A for details):

$$\boldsymbol{\delta\epsilon}^{\mathrm{sys}} \simeq \left(\boldsymbol{\epsilon}_{\mathrm{gal}} - \boldsymbol{\epsilon}_{\mathrm{PSF}} \right) \frac{\delta\left(R_{\mathrm{PSF}}^2\right)}{R_{\mathrm{gal}}^2} - \left(\frac{R_{\mathrm{PSF}}}{R_{\mathrm{gal}}} \right)^2 \boldsymbol{\delta\epsilon}_{\mathrm{PSF}} \tag{13}$$

where R_{gal} and R_{PSF} are the radius parameters of the galaxy and the PSF respectively. The first two terms show that systematics due to an error on the PSF size are proportional to the ellipticities of the galaxy and of the PSF. In the following, we show that the latter should be optimised to be as small as possible. The last term shows that systematics due to an error on the PSF ellipticity are proportional to the squared ratio between the PSF and galaxy sizes.

Combining Eqs. (12) and (13) gives the propagation of PSF errors into σ_{sys}^2. For this purpose we make the following simplifications:

1. the galaxy is not correlated with the PSF (i.e. the crossed-terms $\langle \boldsymbol{\epsilon}_{\mathrm{gal}}.\boldsymbol{\epsilon}_{\mathrm{PSF}} \rangle$ and $\langle \boldsymbol{\epsilon}_{\mathrm{gal}}.\boldsymbol{\delta\epsilon}_{\mathrm{PSF}} \rangle$ are equal to 0);
2. the error on the PSF ellipticity ($\boldsymbol{\delta\epsilon}_{\mathrm{PSF}}$) and the PSF ellipticity itself ($\boldsymbol{\epsilon}_{\mathrm{PSF}}$) are not correlated. This is warranted by the fact that, in the assumed small ellipticity regime, $\boldsymbol{\delta\epsilon}_{\mathrm{PSF}}$ does not have any preferred direction, implying $\langle \boldsymbol{\epsilon}_{\mathrm{PSF}}.\boldsymbol{\delta\epsilon}_{\mathrm{PSF}} \rangle = 0$;
3. we assume that the ellipticity and the inverse squared radius of the galaxy are also uncorrelated. More exactly we assume:

$$\left\langle \left| \frac{\boldsymbol{\epsilon}_{\mathrm{gal}}}{R_{\mathrm{gal}}^2} \right|^2 \right\rangle \simeq \left\langle \left| \boldsymbol{\epsilon}_{\mathrm{gal}} \right|^2 \right\rangle \left\langle \left| 1/R_{\mathrm{gal}}^2 \right|^2 \right\rangle. \tag{14}$$

This assumption is reasonable for our work on the PSF calibration presented in this paper.

With these simplifications, we obtain:

$$\sigma_{\mathrm{sys}}^2 = (P^\gamma)^{-2} \left\langle \left(\frac{R_{\mathrm{PSF}}}{R_{\mathrm{gal}}} \right)^4 \right\rangle \times \left[2\sigma^2[\epsilon_{\mathrm{PSF}}] \right.$$
$$\left. + \left(\left\langle \left| \boldsymbol{\epsilon}_{\mathrm{gal}} \right|^2 \right\rangle + \left\langle \left| \boldsymbol{\epsilon}_{\mathrm{PSF}} \right|^2 \right\rangle \right) \left(\frac{\sigma[R_{\mathrm{PSF}}^2]}{R_{\mathrm{PSF}}^2} \right)^2 \right]. \tag{15}$$

This equation confirms the intuition that the PSF ellipticity of a cosmic shear survey should be small. We see that the last ellipticity term inside the brackets is $\langle | \boldsymbol{\epsilon}_{\mathrm{PSF}} |^2 \rangle$ and should be reduced when optimising the survey to have $\boldsymbol{\epsilon}_{\mathrm{PSF}}$ as small as possible.

For instance, to reach the requirements of $\sigma_{\mathrm{sys}}^2 \lesssim 10^{-7}$, with a typical well sampled cosmic shear survey with $\langle | \boldsymbol{\epsilon}_{\mathrm{gal}} |^2 \rangle = 0.16$,

$R_{gal} \geq 1.5 R_{PSF}$ and $\epsilon_{PSF} \lesssim 0.05$, would require:

$$\frac{\sigma[R_{PSF}^2]}{R_{PSF}^2} \lesssim 10^{-3}, \tag{16}$$

$$\sigma[\epsilon_{PSF}] \lesssim 10^{-3}. \tag{17}$$

In the following, we discuss how these upper limits translate into requirements on the PSF calibration.

3. Analytical model

In this section, we consider the general problem of fitting a 2D model of the PSF to a pixelated image of a star with additive uncorrelated Gaussian noise using χ^2 minimisation. This allows us to derive a number of analytic results in the limit of infinitely small pixels (i.e. infinitely high resolution), which will serve as a useful comparison base for the numerical simulation studies presented in Sect. 4.

3.1. General 2D fit

Consider the PSF surface brightness as a function of position on the image x to be described by a model $m(x; p)$ parameterised by parameters p. The observed surface brightness of a star is $f(x) = m(x; p(\text{PSF})) + n(x)$ where $p(\text{PSF})$ is the true values of the parameters, $n(x)$ is the noise, which is assumed to be uncorrelated (from pixel to pixel) and Gaussian with $\langle n(x) \rangle = 0$ and $\langle n(x)^2 \rangle = \sigma_n^2$ is assumed constant across the image. The usual χ^2-functional is given by

$$\chi^2(p) = \sum_k \sigma_n^{-2} [f(x_k) - m(x_k; p)]^2, \tag{18}$$

the sum being over all pixels k in the image. The usual estimator $\hat{p}$ is constructed by requiring that $d\chi^2/dp = 0$ when evaluated at $\hat{p}$. The covariance matrix of the estimated parameters is given by the inverse of the Fisher matrix:

$$\text{cov}[p_i, p_j] \simeq (\mathcal{F}^{-1})_{ij} \tag{19}$$

with

$$\mathcal{F}_{ij} = \sigma_n^{-2} \sum_k \frac{\partial m(x_k; p)}{\partial p_i} \frac{\partial m(x_k; p)}{\partial p_j} \tag{20}$$

and the variance of a function $P(p)$ of the fitted parameters is:

$$\sigma^2[P] \simeq \sum_{i,j} \frac{\partial P}{\partial p_i} \frac{\partial P}{\partial p_j} \text{cov}[p_i, p_j]. \tag{21}$$

In the following, we consider 2 such functions of the parameters p: the rms radius squared R^2 and the 2 component ellipticity ϵ (as defined in Eqs. (4) to (6)), in the case of a simple elliptical Gaussian PSF (Sect. 3.2) and in the case of more complex PSFs defined with shapelets (Sect. 3.3). We define the associated dimensionless complexity factors ψ_{R^2} and ψ_ϵ such as:

$$\sigma[R^2] \equiv \frac{R^2}{S} \psi_{R^2}, \tag{22}$$

$$\sigma[\epsilon] \equiv \frac{1}{S} \psi_\epsilon \tag{23}$$

where S is the Signal-to-Noise Ratio defined as:

$$S = \frac{F^{(0)}}{\sigma[F^{(0)}]}. \tag{24}$$

$\sigma[F^{(0)}]$ is the standard deviation of the total flux. Basically, ψ_{R^2} and ψ_ϵ characterise the numbers of degrees of freedom in the PSF model associated with R^2 and ϵ.

In the limit of infinitely small pixels, the sum over pixels in Eqs. (18) and (20) can be replaced by a continuous integral over the object. A number of analytical results can be derived. In Sect. 3.2 we derive these for an elliptical Gaussian PSF and in Sect. 3.3 we study more complex PSFs described with shapelets.

3.2. Elliptical Gaussian

We first consider a 2D elliptical Gaussian model parameterised as:

$$m(x; p) = \frac{A}{2\pi \sqrt{a_1 a_2}} \exp\left[-\frac{1}{2}(x - x^a)^T Q^{-1}(x - x^a)\right], \tag{25}$$

$$p = (x^a, A, a_1, a_2, \alpha) \tag{26}$$

where a_1 and a_2 are the rms major and minor axes of the Gaussian, respectively, A is a parameter which controls the amplitude, x^a is the (true) centroid, and ^{T} stands for the transpose operator. The total flux (as defined in Eq. (1)) is:

$$F^{(0)} = A \sqrt{a_1 a_2}, \tag{27}$$

the centroid (as defined in Eq. (2)) is:

$$x_i^{cen} = x_i^a \tag{28}$$

and the quadrupole moment (as defined in Eq. (3)) is:

$$Q = R(\alpha)^T \begin{pmatrix} a_1^2 & 0 \\ 0 & a_2^2 \end{pmatrix} R(\alpha) \tag{29}$$

where α is the position angle of the major axis counter-clockwise from the x-axis and:

$$R(\alpha) = \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix} \tag{30}$$

is the rotation matrix which aligns the coordinate system with the major axis. The ellipticity and the squared radius (as defined in Eqs. (4) to (6)) are:

$$R^2 = a_1^2 + a_2^2, \tag{31}$$

$$\epsilon = \frac{a_1^2 - a_2^2}{R^2}. \tag{32}$$

This parametrisation is particularly convenient because for infinitely small pixels the Fisher matrix of the parameters, given by Eq. (20), is diagonal with diagonal elements:

$$F_{ii} = S^2 \left(\frac{1}{a_1^2}, \frac{1}{a_2^2}, \frac{2}{A^2}, \frac{1}{a_1^2}, \frac{1}{a_2^2}, \frac{(a_1^2 - a_2^2)^2}{2a_1^2 a_2^2} \right). \tag{33}$$

Consequently, with Eq. (21), the errors on the major and minor axes are:

$$\sigma[a_i] = \frac{a_i}{S}. \tag{34}$$

From Eqs. (31), (32) and (34), it follows that:

$$\sigma[R^2] = \frac{2}{S} \sqrt{a_1^4 + a_2^4}, \tag{35}$$

$$\sigma[\epsilon] = \frac{a_1^2 a_2^2}{R^4} \frac{4\sqrt{2}}{S}. \tag{36}$$

We generalise this simple elliptical Gaussian model to more complex PSFs in the following section and test these equations for finite pixels using simulations in Sect. 4.

Under the approximation of small ellipticity ($\epsilon \lesssim 0.1$), the dimensionless complexity factors ψ_{R^2} and ψ_ϵ (defined by Eqs. 22 and 23) are constant (i.e. do not depend on the object). Indeed, $\epsilon \lesssim 0.1$ implies $a_1 \simeq a_2$ and therefore in Eqs. 35 and 36, we have $\sqrt{a_1^4 + a_2^4} \simeq R^2/\sqrt{2}$ and $4a_1^2 a_2^2 \simeq R^4$. This leads to:

$$\psi_{R^2}(\text{Gauss}) = \psi_\epsilon(\text{Gauss}) = \sqrt{2}. \tag{37}$$

3.3. Shapelet model

To explore a wide variety of possible PSFs we consider a basis set that allows for complexity. In this framework, the shapelet basis sets are particularly convenient because: (i) shapelets provide some orthonormal basis sets that have already been studied in number of publications (Refregier 2003a; Refregier & Bacon 2003; Massey & Refregier 2005), tested on the STEP2 and STEP3 simulated data (Massey et al. 2007; Rhodes et al. 2008) and used for the weak lensing analysis of the Canada-France-Hawai-Telescope-Legacy-Survey (CFHTLS) data set in the framework of a comparison with X-ray surveys (Bergé et al. 2007); (ii) as we show in the following, in the case of simple objects the standard deviations of the squared radius $\sigma[R^2]$ and ellipticity $\sigma[\epsilon]$ can be written as proportional to some complexity factors containing all the information about the basis.

A shapelet basis set is characterised by 2 parameters: $n_{\max}$, the maximum order of the functions in the basis and the scale parameter β. The surface brightness $f(\boldsymbol{x})$ of an object is described by:

$$f(\boldsymbol{x}) = \sum_{n=0}^{n_{\max}} \sum_{m=-n}^{n} f_{n,m} \chi_{n,m}(\boldsymbol{x}, \beta) \tag{38}$$

with $\chi_{n,m}$ the polar shapelet functions and $f_{n,m}$ the (complex) polar shapelet coefficients of the object. The scale of the oscillations described by the function $\chi_{n,m}$ is proportional to $1/(n+|m|)$. It is often convenient to impose a lower limit to the scales which are described. This corresponds to setting the coefficients with $|m| > n_{\max} - n$ to 0. This configuration is called "diamond" and is only defined for an even $n_{\max}$. For simple objects with reasonable substructures and tails, the number of coefficients required is small when $R^2 \sim 2\beta^2$. This relation is exact for a circular Gaussian represented with $n_{\max} = 0$.

We show in Appendix B that, under the approximation of small ellipticity ($\epsilon \lesssim 0.1$) and for stars described with a judicious scale parameter where $R^2 = 2\beta^2$, the dimensionless complexity factors defined by Eqs. (22) and (23) ($\sigma[R^2] \equiv R^2 \psi_{R^2}/S$ and $\sigma[\epsilon] \equiv \psi_\epsilon/S$) depend only on the basis (i.e. not on the object itself) and are given by:

$$\psi_{R^2} \simeq \sqrt{\frac{N}{3}(N+1)}, \tag{39}$$

$$\psi_\epsilon \simeq \sqrt{\frac{N}{3}(N+4)} \quad \text{without diamond}, \tag{40}$$

$$\psi_\epsilon \simeq \sqrt{\frac{N}{3}(N-2)} \quad \text{with diamond} \tag{41}$$

where N is the largest even integer lower than or equal to $n_{\max}$. Figure 3 illustrates Eqs. (22), (23) and (39) to (41) by showing $\sigma[\epsilon]$ and $\sigma[R^2]$ as a function of $n_{\max}$. ψ_ϵ and ψ_{R^2} are given in Table 2 for 16 bases with $n_{\max}$ between 4 and 16 and for the elliptical Gaussian model presented in the previous section.

4. Simulations

In this section, we use image simulations to test and validate the analytic predictions of the previous section. These simulations allow us to investigate the effects of pixelation in the case of an elliptical Gaussian PSF. In this simple case we show that pixelation degrades the accuracy of the shape measurements and thus, if the pixel scale is increased, the number of stars needed to calibrate the PSF also increases.

Star images are simulated for: (i) a simple elliptical Gaussian PSF and (ii) complex PSFs derived from the STEP III simulations (Rhodes et al. 2008). In all cases the pixelated images are produced by distributing the star centroids randomly and uniformly over the central pixel. Additive Gaussian noise is then added to each image and the model fitting is done using χ^2 minimisation (for the work presented here, we have not added Poisson noise, which arises from the object itself). Gaussian stars are simulated at high resolution (each pixel is considered as the sum of 7×7 sub-pixels), while shapelet stars are simulated using the analytic pixel integration scheme proposed in the online IDL pipeline[6] by Massey & Refregier (2005).

4.1. Elliptical Gaussian PSF

Figure 2 shows the measured standard deviations on R^2 and ϵ as a function of S, for a range of pixel scales (from 0.9 to 3.9 pixels per PSF *FWHM*). The black diamond symbols correpond to the highest resolution simulations (with 3.9 pixels per PSF *FWHM*). These points are close to the predictions (black curve) which are calculated analytically for the idealised case with infinitly small pixels (i.e. inifinitly high resolution, see Sect. 3.2). The upper and bottom panels show different representations of the same curves: on upper panels the y-axis shows σ, while on the bottom panels, it shows the relative differences with the infinite resolution case. As the pixel size increases (i.e. when the resolution decreases), the standard deviation increases monotonically and moves away from the analytical predictions. This shows that the analytic model can describe a high resolution case while simulations are required to take into account pixelation effects.

Our results in this paper are based on analytical predictions and assume the optimistic case where pixelation is good enough not to degrade the results. From Fig. 2 we see that, for the smallest pixel scale we consider (3.9 pixels per PSF *FWHM*), the results from the analytic calculations and the simulations agree to better than 10%. As the size of the pixels is increased to 0.9 pixels per PSF *FWHM* we see a dramatic increase of the standard deviation of a factor $\sim$10. Some of this pixel scale effect will be mitigated through the use of dithering, but more simulations are needed to fully quantify the extent to which this will help.

4.2. Shapelet PSF

We now turn our attention to the expected standard deviations of R^2 and ϵ (predicted by Eqs. (22) and (23)) for more complicated PSFs. We still work under the setup that the model fitted to the data can exactly describe the truth. With the notation used in Sect. 3.1, this means that when we fit the parameters $\boldsymbol{p}$ of a model $m(\boldsymbol{x}, \boldsymbol{p})$ there exists a solution $\boldsymbol{p}(\text{PSF})$ that exactly describes the PSF.

We model the PSF by decomposing PSF-D of STEP III (Rhodes et al. 2008) into shapelets. We create 16 PSF models with increasing levels of complexity by decomposing the

---

[6] http://www.astro.caltech.edu/~rjm/shapelets
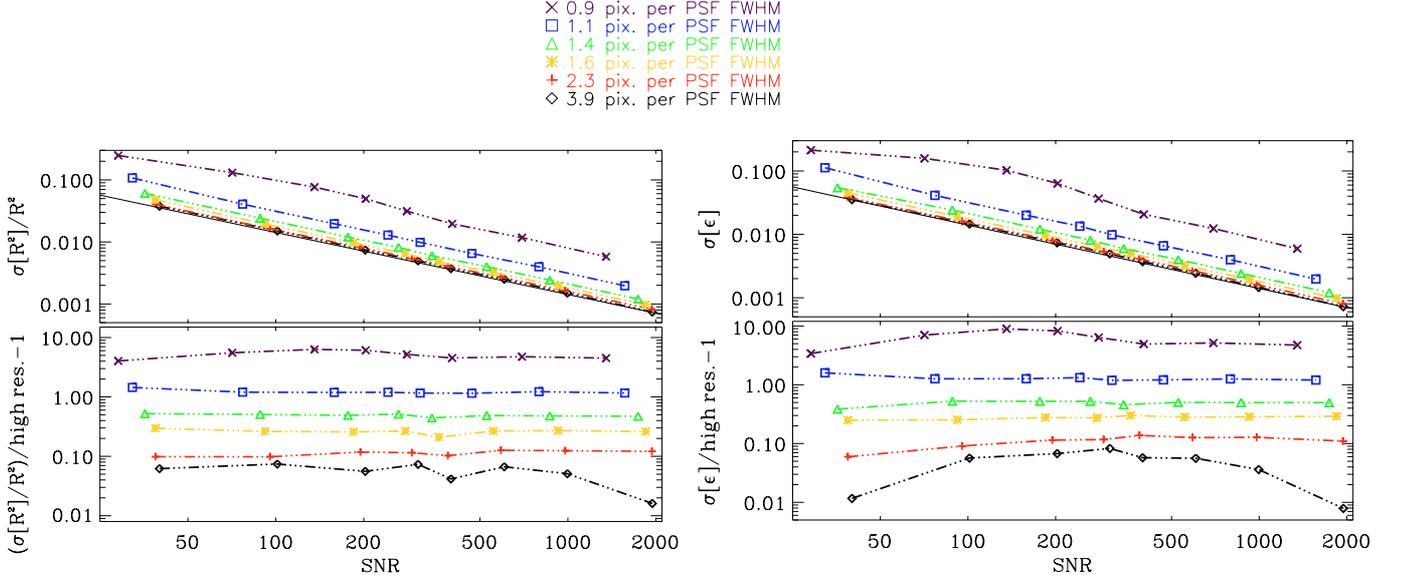
Fig. 2. Solid lines: analytical predictions in the case of infinitely small pixels for $\sigma[\epsilon]$ and $\sigma[R^2]$ (see Eqs. (35) and (36)). Data points joined by straight lines: simulation with finite pixel scales. All these results are for an elliptical Gaussian PSF with an ellipticity $\epsilon \simeq 0.1$. Each data point corresponds to about 10^4 realisations. Black diamonds, red horizontal-vertical crosses, orange stars, green triangles, blue squares and purple diagonal crosses correspond to sizes a_1 = 1.7, 1, 0.7, 0.6, 0.5 and 0.4 pixels respectively. This corresponds to about 3.9, 2.3, 1.6, 1.4, 1.2 and 0.9 pixels respectively per PSF *FWHM*. Each data point corresponds to about 10^4 realisations. On the upper panels the y-axis shows $\sigma[R^2]/R^2$ and σ_ϵ, while on the bottom panels it shows the relative differences as compared to the infinite resolution case. The Signal-to-Noise Ratio is noted "SNR" rather than $\mathcal{S}$ in the text.

STEP PSF into basis sets with increasing n_{max} values. The fitted models are listed in Table 2. This gives 16 slightly different PSFs close to PSF-D (ϵ_i vary between –0.02 and 0.02) which are shown in Fig. 4.

The effect of the level of complexity of the PSF model on the standard deviation of a measurement of the size and ellipticity of the PSF is found by generating $\sim 10^5$ realisations in the interval $10 \le \mathcal{S} \le 1000$ for each of the 16 PSF shapelet basis sets and by fitting with the corresponding shapelet model.

We compare the analytical predictions given by Eqs. (22), (23) and (39) to (41) with simulations in Fig. 3. We see they provide a good description, although they sometimes underestimate $\sigma[R^2]$ and $\sigma[\epsilon]$ by a few percent. The results are similar whether or not the diamond configuration is used.

5. Requirements

In this section, we estimate the number of stars N_* required to estimate the PSF to sufficient accuracy for a given survey. First, each star is an independent realisation of the PSF. Thus the PSF size and ellipticity can be estimated by:

$$R_{PSF}^2 = \frac{\sum_k R_k^2/\sigma^2[R_k^2]}{\sum_k 1/\sigma^2[R_k^2]} = \frac{\sum_k R_k^2 \mathcal{S}_k^2}{\sum_k \mathcal{S}_k^2}, \qquad (42)$$

$$\epsilon_{PSF} = \frac{\sum_k \epsilon_k/\sigma^2[\epsilon_k]}{\sum_k 1/\sigma^2[\epsilon_k]} = \frac{\sum_k \epsilon_k \mathcal{S}_k^2}{\sum_k \mathcal{S}_k^2} \qquad (43)$$

where the sums are made on all the stars. These estimators are the minimum variance weighted averages if we assume the probability distribution in each of R^2 and the two components of ϵ is a Gaussian for each individual star. Second, variances are given

by Eqs. (22) and (23) and can be written:

$$\sigma^2[R_{PSF}^2] = \frac{R_{PSF}^4 \psi_{R^2}^2}{\sum_* \mathcal{S}_{*i}^2} \equiv \frac{R_{PSF}^4 \psi_{R^2}^2}{N_* \mathcal{S}_{eff}^2}, \qquad (44)$$

$$\sigma^2[\epsilon_{PSF}] = \frac{\psi_\epsilon^2}{\sum_* \mathcal{S}_{*i}^2} \equiv \frac{\psi_\epsilon^2}{N_* \mathcal{S}_{eff}^2}. \qquad (45)$$

We have defined an effective signal-to-noise of the stars $\mathcal{S}_{eff}$ as the expected rms signal-to-noise given the maximum and minimum limiting values of $\mathcal{S}$ used for PSF calibration ($\mathcal{S}_{max}$ and $\mathcal{S}_{min}$) and the number density of stars per unit of $\mathcal{S}$: $dn/d\mathcal{S}$, in the data set:

$$\mathcal{S}_{eff}^2 = \left(\int_{\mathcal{S}_{min}}^{\mathcal{S}_{max}} \frac{dn}{d\mathcal{S}} d\mathcal{S} \right)^{-1} \int_{\mathcal{S}_{min}}^{\mathcal{S}_{max}} \frac{dn}{d\mathcal{S}} \mathcal{S}^2 d\mathcal{S}. \qquad (46)$$

We then define the effective number of stars N_* by

$$N_* \mathcal{S}_{eff}^2 = \sum_k \mathcal{S}_k^2 \qquad (47)$$

where the sum is over the stars available around the galaxy. Therefore, the effective number of stars N_* is equal to the actual number of stars if all the stars have the same $\mathcal{S}$.

Substituting into Eq. (15) gives the requirement on the number of stars needed to calibrate the PSF. If the PSF ellipticity is at most a few percent then we can neglect the $\langle |\epsilon_{PSF}|^2 \rangle$ term leaving:

$$N_* \gtrsim \frac{\left(P^\gamma \sigma_{sys}^{lim} \right)^{-2}}{\mathcal{S}_{eff}^2} \left\langle \left(\frac{R_{PSF}}{R_{gal}} \right)^4 \right\rangle \left[\psi_{R^2}^2 \left\langle |\epsilon_{gal}|^2 \right\rangle + 2\psi_\epsilon^2 \right] \qquad (48)$$

where σ_{sys}^{lim} is the upper limit acceptable for σ_{sys}. The number of stars N_* is a central requirement of cosmic shear surveys: if
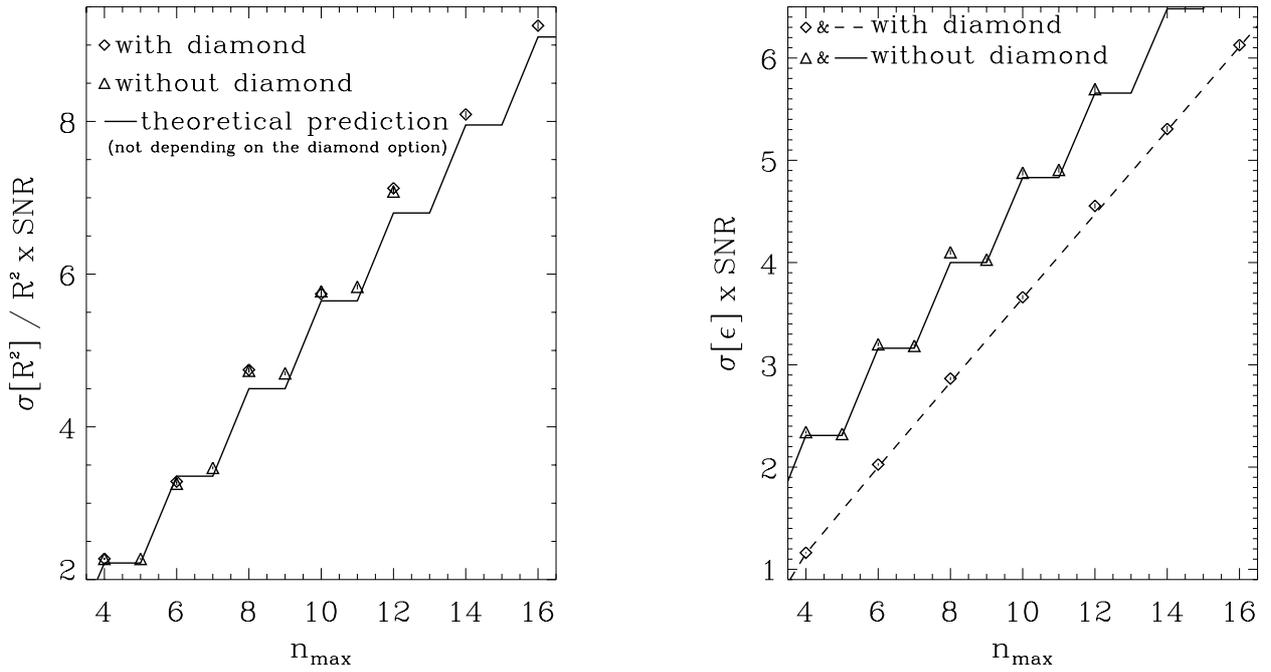
Fig. 3. Lines: analytical predictions in the case of infinitely small pixels for $\sigma[R^2]$ and $\sigma[\epsilon]$ (see Eqs. (22) and (23)). Data points: simulation with small pixels (i.e. 10 pixels per PSF *FWHM*). All these results are for stars imaged through the complex PSFs shown in Fig. 4 and with $10 < \mathcal{S} < 1000$. Each data point corresponds to about 1.5×10^4 realisations. Triangles and diamonds correspond to basis sets with and without the diamond configuration respectively. For $\sigma[R^2]$ the solid line corresponds to the theoretical prediction (which does depend on the diamond option). For $\sigma[\epsilon]$, solid and dashed lines correspond to theoretical predictions with and without the diamond option respectively. Without the diamond option, the complexity factors depend on the largest even integer lower than or equal to $n_{\max}$ (see Eqs. (39) and (40)), this is why theoretical predictions look like a "staircase". On the other hand this behaviour does not appear with the diamond configuration because the latter is defined only for even values of $n_{\max}$. We see that analytical predictions provide a good description, although they sometimes underestimate $\sigma[R^2]$ and $\sigma[\epsilon]$ by a few percent. The Signal-to-Noise Ratio is noted "SNR" rather than $\mathcal{S}$ in the text.
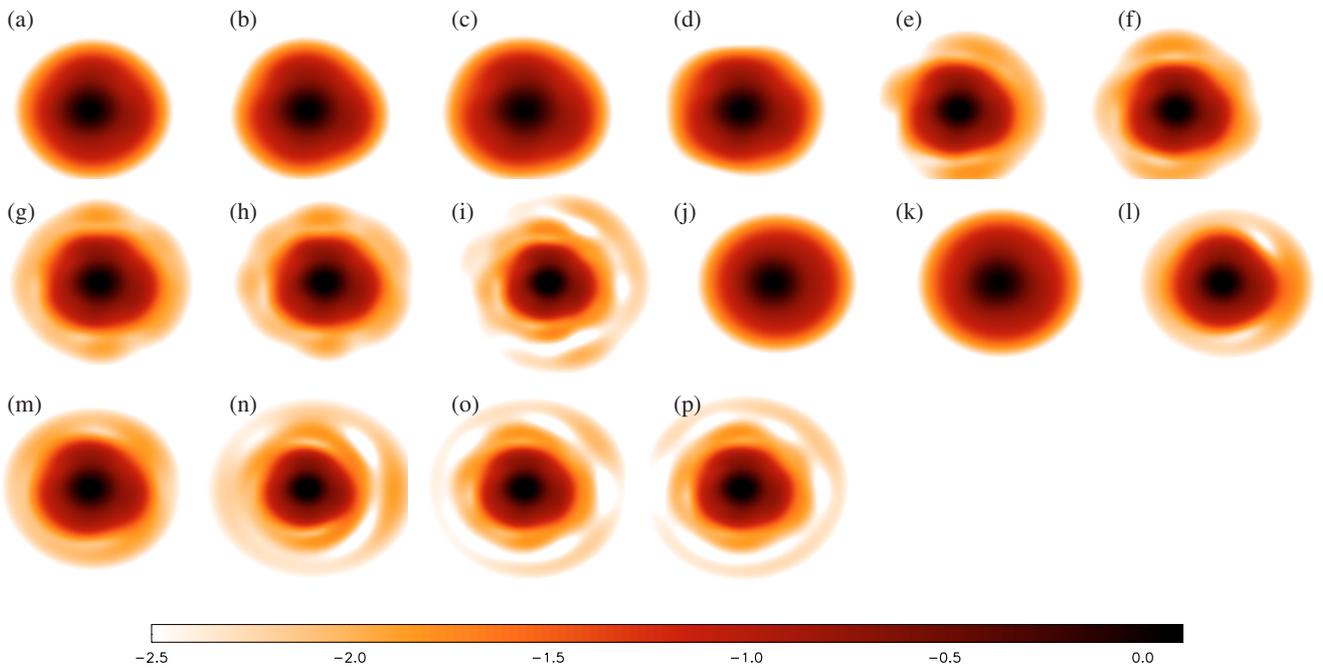
Fig. 4. Non-Gaussian PSFs used to study the effect of complexity in Sect. 4.2 (10 pixels per PSF *FWHM*, logarithmic color scale, the normalisation is such as the total flux is 10^3). Each of them is described with a different shapelet model, described in Table 2.

Table 1. The required number of stars N_* for PSF calibration of typical surveys, as predicted by our scaling relation (Eq. (50)) in function of the effective Signal-to-Noise Ratio of stars $S_{\rm eff}$ (defined in Eq. (46)) and of the scientific requirement given in terms of $(\sigma_{\rm sys}^{\rm lim.})^2$ as defined by Amara & Refregier (2007b) (i.e. the upper limit on the variance of systematics to ensure that systematics are sub-dominant compared to statistical errors when fitting cosmological parameters). The estimations of $\sigma[w_0]$ and $\sigma[w_a]$ are made through the scaling relations proposed by Amara & Refregier (2007a). The mention "nr" means that the constraint is "not relevant". Corresponding surveys (similar to the current generation of cosmic shear surveys) do not constraint the dark energy but rather constraint σ_8. However, $(\sigma_{\rm sys}^{\rm lim.})^2$ does not depend on the considered cosmological parameter and thus these poor constraints on w_0 do contain all the information to compute $(\sigma_{\rm sys}^{\rm lim.})^2$. We consider that stars have an average Signal-to-Noise Ratio of 500 in space and 300 from the ground. We also consider a fixed dilution factor $(R_{\rm gal}/R_{\rm PSF})_{\rm min}$ of 1.5 for every survey. The bold line corresponds to the set-up of a DUNE-like survey (i.e. full-sky from space with medium depth and galaxy density) as taken as the reference in the scaling relation. Note that the surface of 20 000 deg^2 (i.e. half of the sky) is the maximum surface of the sky relevant for cosmic shear. That is why it is often called "full-sky". In the same spirit, 15 000 deg^2 is the maximum surface of the sky (relevant for cosmic shear) that can be covered from Mauna Kea with a reasonable airmass.

	Surface (deg^2)	$z_{\rm m}$	$n_{\rm g}$ (.arcmin^{-2})	$\sigma[w_0]$ ($\sigma[w_a]$)	$(\sigma_{\rm sys}^{\rm lim.})^2$	N_* /$(\psi_\epsilon/3)^2$
Ground surveys	50	0.8	10	0.9 (nr)	4×10^{-6}	3
	170	0.8	10	0.5 (nr)	2×10^{-6}	6
($S_{\rm eff} = 300$)	1500	0.8	20	0.1 (0.6)	6×10^{-7}	25
	15 000	0.9	30	0.03 (0.1)	10^{-7}	104
Space surveys	1000	1.2	50	0.07 (0.3)	3×10^{-7}	15
($S_{\rm eff} = 500$)	**20 000**	**0.9**	**40**	**0.02 (0.1)**	**10^{-7}**	**50**

we need to measure the PSF so that the systematic effects of the shear power spectrum stay below $\sigma_{\rm sys}^{\rm lim.}$, we need to combine the information from at least N_* stars. Thus, for a given survey, one needs to estimate N_* in order to optimise the instrument and mission design. This is a requirement on the PSF stability. Indeed the star density must be taken into account to give the minimum area containing N_* stars, over which the PSF must be stable.

Low S stars do not contribute a great deal of information to the PSF calibration. For instance if the density of stars scales roughly as $dn/dS(S) \propto 1/S$, which is consistent with the simple star count model of Bahcall & Soneira (1980), Eq. (46) shows that the bright stars strongly dominate. In fact, not only do low S stars bring little information, they may also induce a bias in the PSF calibration. This bias will be studied in forthcoming work. For the moment, we avoid this regime by adopting an arbitrary lower limit of $S_{\rm min} = 100$.

On the other hand, the high S cut off $S_{\rm max}$ and the star density distribution dn/dS, depend on the properties of the surveys. They depend on a number of factors including: the line of sight in the Milky Way, the instrumental configuration and the observing strategy. Equation (48), therefore, can not be computed for a general case, but it can be simplified and scaled to typical values, to give easily readable requirements on the star population in a data set. First, since the galaxy size distribution is steep, it is pessimistic but reasonable to approximate:

$$\left\langle \left(\frac{R_{\rm gal}}{R_{\rm PSF}}\right)^4 \right\rangle \approx \left[\left(\frac{R_{\rm gal}}{R_{\rm PSF}}\right)_{\rm min}\right]^4 \tag{49}$$

where $(R_{\rm gal}/R_{\rm PSF})_{\rm min}$ is the minimum value that this ratio can reach, typically about 1.5. Second, we note that the variance $\left\langle |\epsilon_{\rm gal}|^2 \right\rangle$ is around 0.16, as stated in Sect. 2, thus for usual PSF models where the complexity factor of the squared size ψ_{R^2} is of the same order as the complexity factor of the ellipticity ψ_ϵ the expression inside the brackets in Eq. (48) is driven by ψ_ϵ^2 and we can neglect ψ_{R^2}. For example, for non-diamond shapelets with $n_{\rm max} = 4$ or for diamond shapelets with $n_{\rm max} = 6$, the ellipticity complexity factor is $\psi_\epsilon \sim 3$ (see Table 2). Third, as mentioned in Sect. 2, Amara & Refregier (2007b) find for a DUNE-like cosmic shear experiment $(\sigma_{\rm sys}^{\rm lim.})^2 = 10^{-7}$. Fourth, we take for $S_{\rm eff}$ a typical value of 500, which is roughly S of an AB magnitude

Table 2. Properties of the PSF models investigated in Sect. 4.2. The first one is the elliptical Gaussian (labelled "Gauss"). For each of the 16 other basis sets, we study one PSF (labelled a to p) shown in Fig. 4. From left to right, the columns are: $n_{\rm max}$ and diamond defined in Sect. 3.3 and by Massey & Refregier (2005); the number $N_{\rm coef}$ of coefficients in the basis; the numbers $N_{\rm coef}^{m=2(0)}$ of coefficients with $m = 2$ (and $m = 0$) in the basis and the complexity factors defined by Eqs. (39) to (41).

label	$n_{\rm max}$	diamond (yes or no)	$N_{\rm coef}$	$N_{\rm coef}^{m=2(0)}$	complexity factor $\Psi_\epsilon(\Psi_{R^2})$
Gauss	–	–	–	–	$\sqrt{2}$ ($\sqrt{2}$)
a	4	no	15	2(3)	3.27(2.58)
b	5	no	21	2(3)	3.27(2.58)
c	6	no	28	3(4)	4.47(3.74)
d	7	no	36	3(4)	4.47(3.74)
e	8	no	45	4(5)	5.66(4.90)
f	9	no	55	4(5)	5.66(4.90)
g	10	no	66	5(6)	6.83(6.06)
h	11	no	78	5(6)	6.83(6.06)
i	12	no	91	6(7)	8.00(7.21)
j	4	yes	9	1(3)	1.63(2.58)
k	6	yes	16	2(4)	2.83(3.74)
l	8	yes	25	3(5)	4.00(4.90)
m	10	yes	36	4(6)	5.16(6.06)
n	12	yes	49	5(7)	6.32(7.21)
o	14	yes	64	6(8)	7.48(8.37)
p	16	yes	81	7(9)	8.64(9.52)

20 star when imaged with a DUNE-like telescope (1500 s in a broad RIZ band). With these central values, we obtain:

$$N_* \gtrsim 50 \left(\frac{S_{\rm eff}}{500}\right)^{-2} \left(\frac{(R_{\rm gal}/R_{\rm PSF})_{\rm min}}{1.5}\right)^{-4} \left(\frac{\psi_\epsilon}{3}\right)^2 \left(\frac{(\sigma_{\rm sys}^{\rm lim})^2}{10^{-7}}\right)^{-1}. \tag{50}$$

Therefore the PSF of a DUNE-like survey will need to be stable over a region containing about 50 stars.

The scaling relation given by Eq. (50) allows one to study different survey configurations. This is illustrated by Table 1 that

gives N_* for 6 typical configurations comparable to: the Canada-France-Hawai-Telescope Legacy Survey (CFHTLS), the Kilo-Degree Survey[7] (KIDS), the Large Synoptic Survey Telescope (LSST), the SuperNovae Acceleration Probe (SNAP) and the Dark UNiverse Explorer (DUNE). To build this table, we have estimated the accuracy $\sigma[w_0]$ and $\sigma[w_a]$ that can be achieved according to the surface covered, the median redshift z_m and the galactic surface density n_g in the case where systematics are lower than statistical errors and when fitting the 7 cosmological parameters: Ω_m, Ω_b, σ_8, h, w_0, w_a and n. This estimation is made through the scaling relations proposed by Amara & Refregier (2007a). We have then estimated the value of $(\sigma_{sys}^{lim.})^2$ required to achieve this assumption of sub-dominant systematics, according to Amara & Refregier (2007b). Finally we have computed N_* according to our scaling relation (Eq. (50)).

1. The two first lines show surveys comparable to the current largest data sets optimised for cosmic shear. For instance the CFHTLS which currently covers $50 \, \text{deg}^2$ (Fu et al. 2007) and will eventually cover $170 \, \text{deg}^2$. The relatively poor constraints got on w_0 with such surveys (and the absence of constraint on w_a) illustrate the fact that cosmic shear surveys of the current generation do not constraint significantly the dark energy. They rather aim to constrain the combination $\sigma_8 \times \Omega_M^\alpha$. However this does not change the requirement on $\sigma_{sys}^{lim.}$ since Amara & Refregier (2007b) show that this requirement does not depend on the considered cosmological parameter. One can see that these surveys need to calibrate their PSF over few stars. Assuming a typical star density of $\sim 1.\text{arcmin}^{-2}$ at these magnitudes, this shows that at scales smaller than few arcmin the PSF correction may introduce significant systematics. This is consistent with the results of Benjamin et al. (2007) who describe a joint analysis of $100 \, \text{deg}^2$ from several surveys and find significant B modes on these scales;

2. the third line corresponds to a cosmic shear survey of the next generation, such as KIDS/VIKING;

3. the fourth line shows an ambitious deep survey from the ground covering $15\,000 \, \text{deg}^2$, which is the largest cosmic shear survey that can be perfomed from Mauna Kea with a small airmass[8]. This illustrates one of the main limitation of doing cosmic shear from the ground: although the survey is rather deep, the galaxy surface density n_g and the median redshift z_m are limited. This is due to the fact that faint galaxies are too small: they have sizes comparable to the PSF and can not be included in a cosmic shear analysis. This is included in the scaling relation 50 through the minimum dilution factor $(R_{gal}/R_{PSF})_{min}$. Note that it is not possible to change significantly the reference value of $(R_{gal}/R_{PSF})_{min} = 1.5$ adopted in Eq. (50) because it is at the power 4. For instance, a change from 1.5 to 1.2 to increase n_g would also increase N_* by a factor 2.5. Such a survey is similar to the LSST (Kahn & LSST Collaboration 2006) except that the LSST collaboration is developing a specific observing strategy that aims to drastically

reduce ψ_ϵ. This strategy is based on stacking many very short exposures taken with different orientations of the camera in order to circularise the PSF. Another effect of this strategy is that the effective Signal-to-Noise Ration $\mathcal{S}_{eff}$ of stars may be much larger than 300, as assumed in Table 1. But this strategy also has a fundamental limit: the more ψ_ϵ is reduced, the more ψ_{R^2} increases (for instance because of the astrometric errors while stacking the low Signal-to-Noise Ratio exposures). For such a LSST-like strategy, estimating N_* through Eq. (50) is not relevant because in this equation we assume that ψ_ϵ and ψ_{R^2} are of the same order. N_* must be estimated through Eq. (48);

4. the fifth line corresponds to a deep space survey covering $1000 \, \text{deg}^2$. This is similar to SNAP as described by Massey et al. (2004) and Refregier et al. (2004);

5. the last line corresponds to a DUNE-like survey (full sky from space with a medium depth and a reasonable galaxy density) adopted as the reference in the scaling relation (Eq. (50)).

6. Conclusions

We have studied the PSF calibration requirements for cosmic shear to measure cosmological parameters. We connect the finite information that we are able to extract from stars to the statistical error of the PSF calibration, then to the error on the galaxy ellipticity estimation. We express our results in terms of the minimum number of stars (N_*) that are needed to calibrate the PSF in order to keep the systematic errors below statistical errors for cosmological parameter estimation. On scales smaller than the area containing N_* stars there is not enough information coming from the stars to calibrate the PSF. Therefore, the systematic errors in the cosmological parameters may dominate over the statistical errors. This means that these small scales should not be used to constrain cosmology unless the variability is known to be extremely small. Our results show that for current cosmic shear surveys this scale is about an arcminute, which may explain residual systematics found on these scales in current analyses. In future all sky cosmic shear surveys, the data set will be increased by several orders of magnitude and a tight control of the PSF behaviour will be required to prevent this scale from being increased and to reach smaller scales. This places strong requirements on the hardware and observing strategy. For instance, for ground observations where the atmosphere prevents the PSF from being constant over several stars in a single exposure, Jarvis & Jain (2004); Jain et al. (2006) suggest the shear correlation functions should be calculated by cross-correlating galaxies in different exposures. The PSF calibration error contribution would then be uncorrelated and averaged down.

We also demonstrate how N_* can depend on the complexity of the PSF. We define a "complexity factor" for two different shapelet parametrisations and an elliptical Gaussian parametrisation. This complexity factor is a number lower than the number of parameters used to describe the PSF and increases with the PSF complexity. There is a different complexity factor associated with each PSF shape parameter. For accurate galaxy shear measurements, we find that the complexity factor for the PSF ellipticity is more important than that for the size. We summarise the dependence of N_* as a function of complexity, star Signal-to-Noise Ratio, galaxy size and cosmological requirements in a convenient scaling relation (see Eq. (50)). Note that we work under the setup that the PSF can be perfectly described by the model, thus we do not include in our error budget the systematics due to a poor choice of the PSF model.

---

[7] http://www.eso.org/sci/observing/policies/cSurveys/sciencecSurveys.html

[8] On one hand, half of the sky is available for cosmic shear, corresponding to $\sim 20\,000 \, \text{deg}^2$. The other half is masked by the Milky Way. This is why one uses to call "full-sky" the surveys covering $20\,000 \, \text{deg}^2$. On the other hand, in order to maintain the air mass in a reasonable range, we make the optimistic assumption that the minimum elevation of ground observations is 50 deg in local horizon coordinates. With this constraint, a telescope located at Mauna Kea (i.e. Hawaii) can cover about $15\,000 \, \text{deg}^2$ during the year.

We find that pixelation degrades the PSF calibration accuracy. Our calculation of N_* holds for the optimistic case of infinitely small pixels (i.e. infinitely high resolution). A finite pixel size increases N_*. We have performed simulations to predict the increase in the case of an elliptical Gaussian PSF. We show that for 2 pixels (or more) per *FWHM*, the rms scatter agrees with the standard deviation predicted for infinitely small pixels to within 10%, which in turn can be translated into a 20% increase in the required number of calibration stars. For large pixels we see that the pixelation effect can be dramatic, for instance having 0.9 pixels per PSF *FWHM* would lead to a factor of 100 increase in the number of stars needed. This effect may be mitigated to some extent by using dithering, which is beyond the scope of this paper. The main results of this paper are for the high resolution case that provides a strict lower limit on N_*.

Appendix A: Propagation of the PSF error

In this appendix we detail how we derive Eqs. (13) and (15) by propagating the error on each measurement of the PSF ellipticity and size to the error on the galaxy ellipticity and then to σ_{sys}.

The 2 component galaxy ellipticity ϵ_{gal} can be written in terms of radii R_{obs}^2 and R_{PSF}^2 and ellipticities ϵ_{obs} and ϵ_{PSF} of the observed image and the PSF:

$$\epsilon = \frac{\left[F_{11}^{\text{obs}} - F_{11}^{\text{PSF}} - F_{22}^{\text{obs}} + F_{22}^{\text{PSF}}; \; 2(F_{12}^{\text{obs}} - F_{12}^{\text{PSF}})\right]}{F_{11}^{\text{obs}} - F_{11}^{\text{PSF}} + F_{22}^{\text{obs}} - F_{22}^{\text{PSF}}}$$

$$= \frac{\epsilon_{\text{obs}} R_{\text{obs}}^2 - \epsilon_{\text{PSF}} R_{\text{PSF}}^2}{R_{\text{obs}}^2 - R_{\text{PSF}}^2}. \tag{A.1}$$

Consider now that we have estimators of ϵ_{PSF} and R_{PSF}^2 and the estimated values have small errors $\delta\epsilon_{\text{PSF}}$ and $\delta\left(R_{\text{PSF}}^2\right)$ with respect to the true values. We can find the deviation from the truth of a single measurement of one ellipticity component $\delta\epsilon_{\text{gal},i}$ by differentiation:

$$\delta\epsilon_{\text{gal},i} \approx \frac{\partial\epsilon_{\text{gal},i}}{\partial R_{\text{PSF}}^2}\delta R_{\text{PSF}}^2 + \frac{\partial\epsilon_{\text{gal},i}}{\partial\epsilon_{\text{PSF},i}}\delta\epsilon_{\text{PSF},i} \tag{A.2}$$

which can be expanded to give Eq. (13) and when combined with Eq. (12):

$$\left\langle\left|\delta\epsilon^{\text{sys}}\right|^2\right\rangle = \left(\left\langle\left|\frac{\epsilon_{\text{gal}}}{R_{\text{gal}}^2}\right|^2\right\rangle + \left\langle\left|\frac{\epsilon_{\text{PSF}}}{R_{\text{gal}}^2}\right|^2\right\rangle\right)\left\langle\left|\delta\left(R_{\text{PSF}}^2\right)\right|^2\right\rangle$$
$$+ \left\langle\left(\frac{R_{\text{PSF}}}{R_{\text{gal}}}\right)^4\right\rangle\left\langle\left|\delta\epsilon_{\text{PSF}}\right|^2\right\rangle \tag{A.3}$$

which leads to Eq. (15) when assuming that the ellipticity and size of the galaxy are uncorrelated (Eq. (14)) and using Eqs. (7) and (8).

Appendix B: Complexity factors for shapelet models

As stated in Sect. 3.3, any object can be described by the sum of polar shapelet functions $\chi_{n,m}$ weighted by a set of complex coefficients $f_{n,m}$ (see Eq. (38)). The scale parameter β and the center of the basis can be tuned to get a sparse description of the object. Most of the time, a good choice is to take $\beta^2 = R^2/2$ and the center of the basis corresponding to the centroid of the surface brightness. Then, for simple objects like stars with reasonable substructures and tails, the required number of coefficients is typically less than 20. The "diamond" configuration, defined for $n_{\max}$ even, consists in setting $f_{n,m} = 0$ when $|m| > n_{\max} - n$.

For given values of β and the center of the basis, Eqs. (50), (54) and (55) of Massey & Refregier (2005) can be summarised by:

$$F^{(0)} = \sqrt{4\pi}\beta \times S_0$$

$$R^2 = \frac{\sqrt{16\pi}\beta^3}{F^{(0)}}S_1$$

$$= 2\beta^2 \times S_1/S_0 \tag{B.1}$$

$$\epsilon = \frac{\sqrt{16\pi}\beta^3}{F^{(0)}R^2}S_2$$

$$= S_2/S_1 \tag{B.2}$$

with:

$$S_0 = \sum_{n\,\text{even}=0}^{n_{\max}} f_{n0} \tag{B.3}$$

$$S_1 = \sum_{n\,\text{even}=0}^{n_{\max}} (n+1)f_{n0} \tag{B.4}$$

$$S_2 = \sum_{n\,\text{even}=2}^{n_{\max}} \sqrt{n(n+2)}f_{n2}. \tag{B.5}$$

From Eqs. (B.1), (B.2) and (21), it follows that:

$$\sigma[R^2] = \frac{1}{S} \times 2\beta^2 \times \psi_{R^2} \tag{B.6}$$

$$\sigma[\epsilon] = \frac{1}{S} \times \frac{2\beta^2}{R^2} \times \psi_\epsilon \tag{B.7}$$

with:

$$\psi_{R^2} = \frac{1}{N/2+1} \sum_{n\,\text{even}=0}^{n_{\max}} \left[(n+1-\frac{R^2}{2\beta^2})^2\right] \tag{B.8}$$

and, for the non-diamond configuration:

$$\psi_\epsilon^2 = \frac{1}{N/2+1} \sum_{n\,\text{even}=0}^{n_{\max}} \left[n(n+2) + \epsilon^2(n+1)^2\right]$$

$$= \frac{N(N+4)}{3} + \epsilon^2\left(\frac{N}{3}(N+4)+1\right) \tag{B.9}$$

or alternatively, if the diamond configuration is used:

$$\psi_\epsilon^2 = \frac{1}{N/2+1} \sum_{n\,\text{even}=0}^{n_{\max}-2} [n(n+2)] + \epsilon^2 \sum_{n\,\text{even}=0}^{n_{\max}} \left[(n+1)^2\right]$$

$$= \frac{N(N-2)}{3} + \epsilon^2\left(\frac{N}{3}(N+4)+1\right) \tag{B.10}$$

where N is the largest even integer lower than or equal to $n_{\max}$. Note that $\sigma[\epsilon]$ is the standard deviation of one component of ϵ, according to our definition given in Eq. (8).

To obtain Eqs. (39) to (41), we make two simplifications:

1. We neglect the term proportional to ϵ^2 in Eqs. (B.9) and (B.10). This is justified by the "small ellipticity" assumption adopted all along this paper.
2. We choose $R^2 = 2\beta^2$ in Eqs. (B.6), (B.7) and (B.8).

References

Albrecht, A., Bernstein, G., Cahn, R., et al. 2006, ArXiv Astrophysics e-prints
Amara, A., & Refregier, A. 2007a, MNRAS, 381, 1018
Amara, A., & Refregier, A. 2007b, ArXiv e-prints, 710
Bahcall, J. N., & Soneira, R. M. 1980, ApJ, 238, L17
Bartelmann, M., & Schneider, P. 2001, Phys. Rep., 340, 291
Benjamin, J., Heymans, C., Semboloni, E., et al. 2007, MNRAS, 381, 702
Bergé, J., Pacaud, F., Réfrégier, A., et al. 2007, ArXiv e-prints, 712
Bernstein, G. M., & Jarvis, M. 2002, AJ, 123, 583
Fu, L., Semboloni, E., Hoekstra, H., et al. 2007, ArXiv e-prints, 712
Heavens, A. F., Kitching, T. D., & Taylor, A. N. 2006, MNRAS, 373, 105
Heymans, C., VanWaerbeke, L., Bacon, D., et al. 2006, MNRAS, 139, 313
High, F. W., Rhodes, J., Massey, R., & Ellis, R. 2007, ArXiv Astrophysics e-prints
Hoekstra, H. 2003, ArXiv Astrophysics e-prints
Hoekstra, H., Franx, M., Kuijken, K., & Squires, G. 1998, ApJ, 504, 636
Huterer, D., Takada, M., Bernstein, G., & Jain, B. 2006, MNRAS, 366, 101
Jain, B., Jarvis, M., & Bernstein, G. 2006, J. Cosmol. Astro-Part. Phys., 2, 1
Jarvis, M., & Jain, B. 2004, ArXiv Astrophysics e-prints
Kahn, S., & LSST Collaboration 2006, BAAS, 38, 1020
Kaiser, N., Squires, G., & Broadhurst, T. 1995, ApJ, 449, 460
Kuijken, K. 1999, A&A, 352, 355
Luppino, G. A., & Kaiser, N. 1997, ApJ, 475, 20
Massey, R., & Refregier, A. 2005, MNRAS, 363, 197
Massey, R., Rhodes, J., Refregier, A., et al. 2004, AJ, 127, 3089
Massey, R., Heymans, C., Bergé, J., et al. 2007, MNRAS, 376, 13
Munshi, D., Valageas, P., Van Waerbeke, L., & Heavens, A. 2006, ArXiv Astrophysics e-prints
Peacock, J., & Schneider, P. 2006, The Messenger, 125, 48
Refregier, A. 2003a, MNRAS, 338, 35
Refregier, A. 2003b, ARA&A, 41, 645
Refregier, A., & Bacon, D. 2003, MNRAS, 338, 48
Refregier, A., Chang, T.-C., & Bacon, D. J. 2002, in The shapes of galaxies and their dark halos, Proceedings of the Yale Cosmology Workshop "The Shapes of Galaxies and Their Dark Matter Halos", New Haven, Connecticut, USA, 28–30 May 2001, ed. P. Natarajan (Singapore: World Scientific), ISBN 9810248482, 29
Refregier, A., Massey, R., Rhodes, J., et al. 2004, AJ, 127, 3102
Rhodes, J., et al. 2008, in prep.
Stabenau, H. F., Jain, B., Bernstein, G., & Lampton, M. 2007, ArXiv e-prints, 710
Voigt, L., & Bridle, S. L. 2008, in prep.