

Star/galaxy separation at faint magnitudes: application to a simulated Dark Energy Survey

M. T. Soumagnac,^{1,2★} F. B. Abdalla,^{1,3} O. Lahav,¹ D. Kirk,¹ I. Sevilla,⁴ E. Bertin,⁵
B. T. P. Rowe,¹ J. Annis,⁶ M. T. Busha,^{7,8} L. N. Da Costa,^{9,10} J. A. Frieman,^{6,11,12}
E. Gaztanaga,¹³ M. Jarvis,¹⁴ H. Lin,⁶ W. J. Percival,¹⁵ B. X. Santiago,^{10,16}
C. G. Sabiu,^{17,18} R. H. Wechsler,^{19,20} L. Wolz^{1,21} and B. Yanny⁶

¹Department of Physics and Astronomy, University College London, Gower Street, London WC1E6BT, UK

²Ben-Zvi Center for Astrophysics, Weizmann Institute of Science, 76100 Rehovot, Israel

³Department of Physics and Electronics, Rhodes University, PO Box 94, Grahamstown 6140 South Africa

⁴Centro de Investigaciones Energéticas Medioambientales y Tecnológicas, Av. Complutense 40, E-28040 Madrid, Spain

⁵Institut d'Astrophysique de Paris, UMR 7095 CNRS, Université Pierre et Marie Curie, 98 bis boulevard Arago, F-75014 Paris, France

⁶Center for Particle Astrophysics, Fermi National Accelerator Laboratory, PO Box 500, Batavia, IL 60510, USA

⁷Institute for Theoretical Physics, University of Zurich, Wolfgang-Pauli-Str. 27 CH - 8093 Zurich, Switzerland

⁸Physics Division, Lawrence Berkeley National Laboratory, 1 Cyclotron Road Mailstop 50-4049 Berkeley, CA 94720-8153, USA

⁹Observatorio Nacional, Rua Gal. Jose Cristino 77, Rio de Janeiro, RJ – 20921-400, Brazil

¹⁰Laboratório Interinstitucional de e-Astronomia – LineA, Rua Gal. José Cristino 77, Rio de Janeiro, RJ – 20921-400, Brazil

¹¹Department of Astronomy and Astrophysics, The University of Chicago, 5640 South Ellis Avenue, Chicago, IL 60637, USA

¹²Kavli Institute for Cosmological Physics, The University of Chicago, 5640 South Ellis Avenue Chicago, IL 60637, USA

¹³Institut de Ciències de l'Espai (IEEC-CSIC), E-08193 Bellaterra (Barcelona), Spain

¹⁴Department of Physics and Astronomy, University of Pennsylvania, Philadelphia, PA 19104, USA

¹⁵Institute of Cosmology and Gravitation, University of Portsmouth, Portsmouth PO1 3FX, UK

¹⁶Departamento de Astronomia, Universidade Federal do Rio Grande do Sul Av. Bento Gonçalves 9500, Porto Alegre, RS 91501-970, Brazil

¹⁷Korea Institute for Advanced Study, Dongdaemun-gu, Seoul 130-722, Republic of Korea

¹⁸Korea Astronomy and Space Science Institute, 776, Daedeokdae-ro, Yuseong-gu, Daejeon 305-348, Korea

¹⁹Kavli Institute for Particle Astrophysics and Cosmology, SLAC National Accelerator Laboratory, 2575 Sand Hill Rd, Menlo Park, CA 94025, USA

²⁰Physics Department, Stanford University, Stanford, CA 94305, USA

²¹Sub-Department of Astrophysics, Department of Physics, University of Oxford, Keble Road, Oxford OX1 3RH, UK

Accepted 2014 July 11. Received 2014 July 10; in original form 2013 June 20

ABSTRACT

We address the problem of separating stars from galaxies in future large photometric surveys. We focus our analysis on simulations of the Dark Energy Survey (DES). In the first part of the paper, we derive the science requirements on star/galaxy separation, for measurement of the cosmological parameters with the gravitational weak lensing and large-scale structure probes. These requirements are dictated by the need to control both the statistical and systematic errors on the cosmological parameters, and by point spread function calibration. We formulate the requirements in terms of the *completeness* and *purity* provided by a given star/galaxy classifier. In order to achieve these requirements at faint magnitudes, we propose a new method for star/galaxy separation in the second part of the paper. We first use principal component analysis to outline the correlations between the objects parameters and extract from it the most relevant information. We then use the reduced set of parameters as input to an Artificial Neural Network. This multiparameter approach improves upon purely morphometric classifiers (such as the classifier implemented in SExtractor), especially at faint magnitudes: it increases the purity by up to 20 per cent for stars and by up to 12 per cent for galaxies, at *i*-magnitude fainter than 23.

Key words: gravitational lensing: weak – methods: data analysis – surveys – cosmology: observations – dark energy – large-scale structure of Universe.

1 INTRODUCTION

What makes a star look different from a galaxy in a deep image? This seemingly very simple question hides the much more

* E-mail: maayane.soumagnac@gmail.com

complicated issue of allocating a size and a scale to objects observed in the sky, which has concerned observers and theorists throughout the 20th century. The problem of classifying stars and galaxies in large scale surveys is a long-standing one. It has been encountered back in the early 1990s (e.g. the APM survey; Maddox et al. 1990) and poses a major challenge for all recent and large imaging cosmological surveys, including the Dark Energy Survey (DES) (<http://www.darkenergysurvey.org/>) and *Euclid* (<http://sci.esa.int/euclid>), which have been designed to uncover the nature of dark energy (DE). One common denominator of the wide variety of observational probes constraining DE is the necessity to select pure samples of galaxies. More specifically, all the surveys must differentiate galaxies at cosmological distances from local objects, to obtain pure, or at least well-understood, samples.

In the area of ‘precision cosmology’, any source of systematic error is likely to play a decisive role and needs to be taken into account in order to refine the standard inflationary big bang picture. An example of a scientific question for which star/galaxy separation is a potentially critical systematic is the precision measurement of primordial non-Gaussianities (PNG). These manifest themselves by making the bias of a given type of tracers of dark matter haloes strongly scale-dependent. This effect can easily be mimicked by any local systematic effect adding power at large scales and correlated with the galaxies. As the stellar distribution in the Milky Way is across large angular scales, star/galaxy separation is likely to introduce systematic errors in the measurement of PNG. Another example is the effect of occultation of galaxies by stars of comparable magnitudes. Ross et al. (2011) showed that this effect constitutes a source of systematic error in the measurement of angular and photometric distributions of luminous red galaxies. Photometric effects associated with faint stars could therefore partially account for the excess power seen in Thomas, Abdalla & Lahav (2011) for the MegaZ-Luminous Red Galaxy survey. This paper gives two other examples, in the case of weak lensing (WL) and large-scale structure (LSS) measurements, where star/galaxy separation is a key systematic, which needs to be taken into account in order to properly constrain DE.

The outline of this paper is as follows. In Section 2, we present the DES and the ‘DES-like’ simulations which we base our analysis on. In Section 3, we study the impact of star/galaxy misclassification on the measurement of the cosmological parameters, in the case of the WL and LSS probes, and show how the requirements on the statistical and systematic errors propagate into new requirements on the quality of star/galaxy separation. In Section 4, we summarize the current methods for star/galaxy classification and the motivations for our multiparameter approach. The details of the method are presented in Section 5. In Section 6, we compare our star/galaxy classification tool to the ones provided by other methods and confront these results to the science requirements derived in Section 3. Finally, we summarize our main conclusions in Section 7.

2 THE DARK ENERGY SURVEY

The Dark Energy Survey (DES)¹ is an imaging survey of 5000 deg² on southern sky, utilizing the four metre Blanco telescope in Chile. It will provide imaging of 300 million galaxies in five filters (g , r , i , z and Y). Photometric redshifts will be obtained from the colour information to produce a three dimensional survey. The main goal of DES is to determine the DE equation-of-state parameter, $w(z)$, and

other key cosmological parameters to high precision. DES will measure $w(z)$ using four complementary techniques in a single survey: counts of galaxy clusters (GC) (with synergy with clusters detected by the Sunyaev–Zel’dovich effect with the South Pole Telescope), weak gravitational lensing (WL), galaxy power spectra and Type Ia supernovae (SNe). It is expected that the uncertainty on $w(z)$ will be only a few per cent for each probe (see DES collaboration 2005, for detailed parameterisations and statistics). The science requirements of DES drove the construction of a new camera, the Dark Energy Survey camera (DECam), which had its first light in 2012 September and the survey has started in 2013 September.

As part of the process of testing and validation of the DES Data Management (DESDM) system (Mohr et al. 2012), a series of detailed simulations have been designed to serve as a test-bench for the development of the pipelines and for verifying the scientific reach of the experimental channels. Each of these iterations of the simulations are dubbed ‘Data Challenges’ (DC). The simulation starts with the creation of galaxy catalogues stemming from an N -body simulation (Busha et al. 2013) and detailed models of the Milky Way galaxy (Rossetto et al. 2011) for the star component. These are merged and fed to an image simulator which includes atmospheric and instrumental effects. The resulting images serve as inputs for DESDM and are processed as the data will be: the code SExtractor (Bertin & Arnouts 1996) produces a catalogue of more than 300 parameters encapsulating information about each detected object.

The most relevant features of these simulations for our study are as follows:

- (i) the seeing is introduced as a function of observing time;
- (ii) the galaxy shapes have been implemented using a Sérsic profile which matches the observed profile;
- (iii) the point spread function (PSF) takes into consideration the seeing for that time, the optics and the distortion as a function of separation from the optical axis.

The results shown in this paper are based on the latest release (internal to the DES Collaboration) of simulated data, DC6, which covers approximately 140 deg² to the full DES depth, corresponding to about 10 nights of observations. We select from it the objects with a model magnitude in the i band brighter than 24, as they are the ones most likely to be detected with DES.

3 SCIENCE REQUIREMENTS ON STAR/GALAXY SEPARATION

DES will be among the first surveys to combine in a single project the observation of the four preferred DE probes, as identified by the Dark Energy Task Force (DETF; Albrecht et al. 2006). SNe and Baryonic Acoustic Oscillation (BAO) constrain the expansion of the Universe as a whole and are therefore referred to as *purely geometric*. WL and GC constrain both the expansion on the Universe and the growth of LSS (see Weinberg et al. 2013 for a complete review).

In order to properly constrain DE, the broad variety of measures carried out within each probe must meet certain requirements defined by DES science teams. While there is no unique way to specify the constraints on DE experiments and probes, the Figure of Merit (FoM), defined by the DETF, provides a useful metric. If we parametrize the time evolution of DE by the equation of state $w(a) = w_0 + (1 - a)w_a$, where $a(t) = \frac{1}{1+z(t)}$ is the cosmic scale factor and $z(t)$ is the redshift of an object emitting at time t , the FoM is defined as the reciprocal of the area of the error ellipse

¹ <http://www.darkenergysurvey.org/>

enclosing 95 per cent confidence limit in the w_o – w_a plane. Larger FoM indicates smaller errors and therefore greater accuracy on the measurement of the parameters.

In other words, reaching the FoM goals requires to minimize the error on w_o and w_a . Since the total error is the sum of the *statistical* error and the *systematic* error, we can derive two types of science requirements. More concretely, the total mean square error (MSE) on a cosmological parameter p_α can be decomposed as

$$\text{MSE}[p_\alpha] = \sigma^2[p_\alpha] + \Delta^2[p_\alpha], \quad (1)$$

where $\sigma^2[p_\alpha]$ is the statistical error variance and $\Delta[p_\alpha]$ is the parameter shift due to the systematic signals. For each probe, both of these terms needs to be controlled in order to minimize the total error.

Star/galaxy misclassification is an interesting effect because it contributes to both the statistical and systematic part of the total error, for the WL and LSS probes. This allows us to translate separately the requirement on the statistical term (Section 3.2) and the requirements on the systematic term (Section 3.3) into requirements on the quality of the star/galaxy separation. Additional requirements are specific to each probe, e.g. PSF calibration for WL (Section 3.4).

We outline below a formalism to derive these requirements.

3.1 Formalism

3.1.1 Completeness, contamination and purity

In the following, we define the parameters used to quantify the quality of a star/galaxy classifier. For a given class of objects, X (stars or galaxies), we distinguish the surface density of well-classified objects, N_X , and the density of misclassified objects, M_X , as specified in Table 1.

The galaxy *completeness* c_g^g is defined as the ratio of the number of true galaxies classified as galaxies to the total number of true galaxies. The stellar contamination f_s is defined as the ratio of stars classified as galaxies to the total amount of objects classified as galaxies.

$$c_g^g = \frac{N_G}{N_G + M_G}, \quad (2)$$

$$f_s = \frac{M_S}{N_G + M_S}. \quad (3)$$

The *purity* p^g is defined as $1 - f_s$:

$$p^g = \frac{N_G}{N_G + M_S} = 1 - f_s. \quad (4)$$

Similar parameters can be defined for a sample of stars: p^s , f_g and c^s .

We aim to formulate the requirements on the statistical and systematic errors in terms of constraints on these parameters. This will allow us to quickly compare the performance of the classifiers presented in Sections 4 and 5 and assess whether they allow us to achieve the goals of the DETF FoM.

Table 1. Notations for each category of objects (stars and galaxies), for the definitions of completeness and purity.

	True Galaxies	True stars
Objects classified as galaxies	N_G	M_S
Objects classified as stars	M_G	N_S

One should note that there are some inefficiencies in the image pipeline, which are studied in DC6 and which we do not deal with in this analysis. Instead, we define the latter parameters with respect to the mock galaxy samples used to produce the image simulations. With real DES data, our results could be tested e.g. on *Hubble Space Telescope* (HST) data in the same fields.

3.1.2 Fisher information matrix

The Fisher information matrix describes how the errors on the angular power spectrum $C(l)$ (of the cosmic shear in the case of WL, and the density fluctuations of galaxies in the case of LSS) propagate into the precision on the cosmological parameters p_α . We employ this formalism (see Tegmark, Taylor & Heavens 1997 for a review), to quantify the impact of star/galaxy misclassification on each of the terms in equation (1), i.e. on the statistical and systematic errors on the cosmological parameters.

The Fisher matrix can be expressed as

$$F_{\alpha\beta} = \sum_l \sum_{(i,j)(m,n)} \frac{\partial C_{ij}(l)}{\partial p_\alpha} \mathbf{Cov}^{-1}[C_{ij}(l), C_{mn}(l)] \frac{\partial C_{mn}(l)}{\partial p_\beta}, \quad (5)$$

where the sum is over multipole values and redshift bins (typically five for WL). $\mathbf{Cov}[X, Y]$ designates the covariance matrix of X and Y and is given by (Takada & Jain 2004)

$$\mathbf{Cov}[C_{ij}(l), C_{mn}(l)] = \frac{\{C_{im}(l)C_{jn}(l) + C_{in}(l)C_{jm}(l)\}}{f_{\text{sky}}(2l+1)\Delta l}, \quad (6)$$

where f_{sky} is the fraction of the sky covered by the survey ($f_{\text{sky}} = 0.1212$ for DES) and Δl is the width of the corresponding angular frequency bin.

3.2 Science requirements on the statistical errors

How does the need to control the statistical errors on the cosmological parameters propagate into a requirement on the quality of star/galaxy separation? In the following, we aim to answer this question in the case of the WL and LSS probes.

3.2.1 WL measurements

Gravitational lensing from distant intervening mass fluctuations causes the shapes of objects to be distorted such that they appear to be more or less elliptical. While no single object is intrinsically round, if the intrinsic shapes of galaxies are uncorrelated with one another, one can average the apparent shapes of many thousands of such objects to extract a distortion attributed to WL. The statistical properties of this observable pattern put a constraint on the power spectrum and therefore on the cosmological model and on DE. For some concise introductions to cosmic shear, see e.g. Mellier (1999), Bartelmann & Schneider (2001) and Refregier (2003).

How do star/galaxy misclassifications affect the WL shear measurement? The predicted shear angular power spectrum $C_{ij}(l)$ depends on N_{eff} , the effective density per unit area of galaxies with reliable shape measurements,

$$C_{ij}(l) = \int_0^{r_H} dr r^2 W_i(r) W_j(r) P(l/r; r) + \delta_{ij} \frac{\sigma_e^2}{N_{\text{eff}}}, \quad (7)$$

where $P(k = l/r)$ is the 3D matter power spectrum, $W_i(r)$ and $W_j(r)$ are the radial window functions of the redshift bins (i, j) , r is the comoving distance and r_H is the Universe horizon. The angular

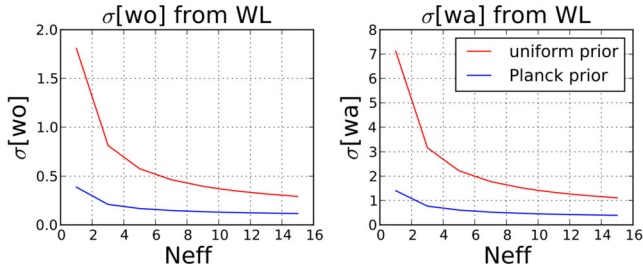


Figure 1. Marginalized statistical errors on the equation-of-state parameters w_o and w_a from the WL probe, for different values of the density of galaxies with reliable shape measurement N_{eff} . The errors are marginalized over $\{\Omega_m, H, \sigma_8, \Omega_b, n_s, b_g\}$ and computed using the assumptions and setup described in Section 3.2.1. The red curve shows the errors computed with a non-informative prior whereas the blue curve is obtained assuming a Planck prior.

power spectrum depends on N_{eff} through the last term, i.e. the ‘shot noise’ due to σ_e , the intrinsic ellipticity noise for the galaxy sample.

In order study the effect of N_{eff} on the statistical error $\sigma[p_\alpha]$, we compute the Fisher matrix for different values of N_{eff} . We estimate the $C_{ij}(l)$ and $\frac{\partial C_{mn}(l)}{\partial p_\alpha}$ terms (see equation 5) using the same code as in Laszlo et al. (2012) and Kirk et al. (2013). The setup is as follows: we use a model with eight free parameters: $\{w_o, w_a, \Omega_m, H, \sigma_8, \Omega_b, n_s, b_g\}$; we assume a Planck prior (Weller, personal communication); there are five tomographic bins of roughly equal number density between $z = 0$ and 3; the redshift distribution is a Smail-type distribution (e.g. equation 12 of Amara & Refregier 2008, with $\alpha = 2$, $\beta = 1.5$, $z_0 = \frac{0.8}{1.412}$); we compute the $C_{ij}(l)$ and $\frac{\partial C_{mn}(l)}{\partial p_\alpha}$ terms for $l \in [1, 1024]$, to avoid the strongly non-linear regime where baryon physics will start being important and following the l -cuts performed in most recent works by the WL community (Debono et al. 2010; Das et al. 2012; Audren et al. 2013); and the photometric redshift error is $\Delta z = 0.05^*(1 + z)$.

We then compute the marginalized statistical error on the cosmological parameters by approximating them with their Cramer–Rao lower bound

$$\sigma[p_\alpha] \approx \sqrt{(F^{-1})_{\alpha\alpha}}. \quad (8)$$

We show the results for w_o and w_a in Fig. 1 and for the other free parameters of our model in the appendix.

Fig. 1 shows that larger N_{eff} translates into smaller statistical errors on w_o and w_a , i.e. larger FoM, which puts a constraint on N_{eff} : it has to be higher than a threshold value N_{thresh} which can depend on the bandpass considered,

$$N_{\text{eff}} \geq N_{\text{thresh}}. \quad (9)$$

Fig. 1 also shows asymptotes above $N_{\text{thresh}} = 10$, i.e. the effect of any variation of N_{eff} on the statistical error decreases at high N_{eff} . In practice, we require the increase of the statistical error due to star/galaxy misclassification to be smaller than 2 per cent. If this reasonable but somewhat arbitrary goal is not achieved, it will only increase the statistical error and will not lead to a bias of the WL results. This translates into a decrease of N_{eff} smaller than 4 per cent, i.e.

$$c^g \geq 96.0 \text{ per cent}. \quad (10)$$

Star–galaxy misclassification is only one among many other sources of errors leading true galaxies to be rejected from the sample of galaxies with reliable shape measurements (e.g. shape measurement errors and photo- z errors). To insure that the statistical errors are

controlled, this condition on c^g should be completed by constraints on the survey parameters controlling all the other sources of errors.

3.2.2 LSS measurements

LSS measurements allow us to constrain DE in various ways. The position of the BAO feature provides a standard ruler to study the expansion history. The shape of the angular power spectrum of the galaxy density fluctuation encapsulates precious information about the clustering amplitude and the growth of structures.

Star/galaxy misclassification affects the power spectrum measurements and the statistical error on the cosmological parameters in a similar way as in the WL case. Indeed, we can write the same equation as equation (7) for the angular power spectrum of the galaxy density fluctuations. The shot noise term is then given by $\frac{1}{N_G}$, where N_G is simply the surface density of detected galaxies. In Fig. 2, we show the evolution of the statistical errors on w_o and w_a with the density of detected galaxies, computed using the same setup as in the WL case.

During the design phase of the project, it has been estimated that in order to achieve the goals of the LSS FoM, the 5000 deg² DES survey will need to provide reliable photo- z and position measurement for about 200 million galaxies, i.e. the number of galaxies correctly classified N_G should be higher than 11.1 arcmin^{−2} (when using combined measurements from the r , i and z bandpasses). This requirement is currently being revisited using data. When doing the latter calculation on the truth table of DC6, for which the surface density of galaxies is $N_{\text{tot}}^g \approx 12.5$, this threshold on N_G translates into the following requirement on the galaxy completeness provided by the star/galaxy classifier: $c^g > 88.9$ per cent.

Note that galaxies at different redshifts have different weights. This is somewhat related to the magnitude of galaxies, as the distribution of brighter galaxies will peak at a lower redshift than galaxies which are fainter. Even though this is not explicitly stated when mentioning that there is a requirement of 200 million galaxies, we are implying that we are effectively going to a given depth and therefore sampling galaxies out to a given redshift. These caveat are implicitly included within the Fisher matrix calculation.

Note also that the derived requirement is a necessary but not sufficient condition, as other sources of errors, apart from star/galaxy misclassification (e.g. photo- z errors), reduce the number of galaxies which can be used for LSS measurement.

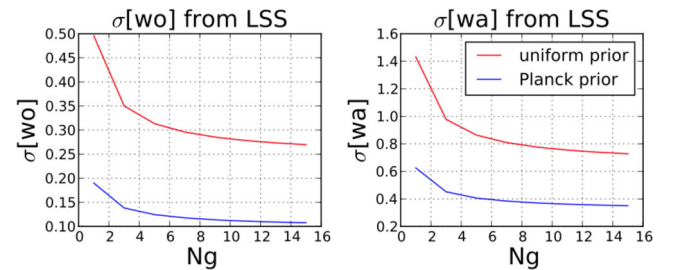


Figure 2. Marginalized statistical errors on the equation-of-state parameters w_o and w_a from the LSS probe, for different values of the density of detected galaxies N_g . The errors are marginalized over $\{\Omega_m, H, \sigma_8, \Omega_b, n_s, b_g\}$ and computed using the same assumptions and setup as in the WL case (see Section 3.2.1), with $l \in [10, 400]$, to avoid the non-linear regime and following most recent l -cuts work by the LSS community (Rassat et al. 2008; Audren et al. 2013). The red curve shows the errors computed with a non-informative prior whereas the blue curve is obtained assuming a Planck prior.

3.3 Science requirements on the systematic errors

We now explore the contribution of star/galaxy misclassification as a source of *systematic* error, which need to be controlled in order for the FoM objectives to be achieved. Star/galaxy misclassifications generate a residual signal $\delta C^{\text{sys}}(l)$ in the angular power spectra (of the cosmic shear in the case of WL, and the density fluctuations of galaxies in the case of LSS), which propagates into a systematic shift $\Delta[p_\alpha]$ of the cosmological parameter p_α . We use the same formalism as in Amara & Refregier (2008) (see also Kirk et al. 2012 and Huterer et al. 2006), to derive $\Delta[p_\alpha]$,

$$\Delta[p_\alpha] = \sum_{\beta, l, (i, j), (m, n)} (\mathbf{F}^{-1})_{\alpha\beta} \delta C_{ij}^{\text{sys}}(l) \mathbf{Cov}^{-1}[C_{ij}^{\text{gal}}(l), C_{mn}^{\text{gal}}(l)] \frac{\partial C_{mn}^{\text{gal}}(l)}{\partial p_\beta}, \quad (11)$$

where \mathbf{F}^{-1} is the inverse Fisher matrix. A criterion, usually used to constrain the contribution of the systematic error to the total MSE, is to define a tolerance on the systematics such that they do not dominate over statistical error. This is verified when

$$|\Delta[p_\alpha]| \leq \sigma[p_\alpha]. \quad (12)$$

In the following sections, we derive the systematic parameter shift for seven cosmological parameters $p_\alpha = \{w_o, w_a, \Omega_m, H, \sigma_8, \Omega_b, n_s\}$ and the galaxy bias b_g , in the case of WL and LSS. This allows us to translate equation (12) into requirements on the quality of the star/galaxy separation.

3.3.1 Requirement from WL measurements

In the case of WL, the systematic error $\delta C^{\text{sys}}(l)$ comes from the fact that some stars are identified as galaxies, and therefore contribute to the measured cosmic shear. We decompose the measured shear γ_m into the contribution from the true galaxies and the contamination from the misclassified stars. The galaxy shear is measured by deconvolving the observed shear and a PSF model; therefore, the contamination from stars in a galaxy sample appears as a residual deconvolved shear:

$$\gamma_m = (1 - f_s)\gamma_g + f_s\gamma_{s,\text{res}}. \quad (13)$$

where $f_s = 1 - p^g$ is the stellar contamination rate (defined in equation 3) and $\gamma_{s,\text{res}}$ is the residual PSF shear, after deconvolution of the PSF model from the shape of misclassified stars. In the following analysis, we make a toy model where the residual deconvolved shear can be written as

$$\gamma_{s,\text{res}} = a\gamma_s, \quad (14)$$

where $a \in [0, 1]$ and γ_s is the stellar shear. The measured two-point shear correlation function is then

$$\langle \gamma_m \gamma_m \rangle = (1 - f_s)^2 \langle \gamma_g \gamma_g \rangle + f_s^2 a^2 \langle \gamma_s \gamma_s \rangle, \quad (15)$$

and in terms of measured angular power spectrum, the latter equation reads

$$C^{\text{obs}}(l) = (1 - f_s)^2 C^{\text{gal}}(l) + f_s^2 a^2 C^s(l), \quad (16)$$

where $a = \alpha^2$ and where we assumed that γ_g and γ_s are uncorrelated. In practice, this is not necessarily the case. Our toy model introduces into the same term, $\alpha C^s(l)$, the auto-correlation of the residual ‘deconvolved star shapes’ and possible cross-correlation between them and the galaxy shear γ_g . Setting α to zero comes to

neglecting both of these terms, and setting $\alpha = 1$ comes to over-estimating them both. We derive the requirement on the quality of star/galaxy separation in the two limiting cases $\alpha = 1$ and $\alpha = 0$ and leave the more general case for further analysis. Equation (16) gives the residual systematic signal

$$\delta C^{\text{sys}}(l) = f_s^2 (C^{\text{gal}}(l) + \alpha C^s(l)) - 2f_s C^{\text{gal}}(l). \quad (17)$$

The requirement stated in equation (12) can be reformulated as a requirement on the stellar contamination rate f_s ,

$$\mathcal{P}(f_s) \leq 0, \quad (18)$$

where \mathcal{P} is a second-order polynomial.

The assumptions made to solve equation (18) are detailed below. We use the setup detailed in Section 3.2.1 to compute the Fisher matrix and the marginalized statistical errors $\sigma[p_\alpha]$ on the cosmological parameters. To estimate $C^s(l)$ in equation (17), we assume that it is the sum of a ‘shot noise’ term and a term due to the correlation of stellar shapes across the field of view,

$$C^s(l) = C_{\text{noise}}^s + C_{\text{tile}}^s(l). \quad (19)$$

We measure $C_{\text{tile}}^s(l)$, the power spectrum of the shapes of the stars in DC6, using the same code as in Jarvis, Bernstein & Jain (2004). The ‘shot noise’ term is given by

$$C_{\text{noise}}^s = \frac{\sigma_s^2}{N_{\text{tot}}^s}, \quad (20)$$

where $N_{\text{tot}}^s = N_s + M_s$ (see Section 3.1.1) is the density of stars and σ_s is the ellipticity of stars. Various complex effects combine as different terms in the stellar ellipticity. First, the interpolated PSF can only constrain large-scale modes of the PSF variation. The power on small scales is not well-constrained at all, so it will show up more or less completely in the shapes of stars that are corrected by the PSF model. Secondly, the deconvolution process magnifies the errors in the shape measurement, especially for objects that are nearly the same size as the PSF, which is presumably the case for stars that are misidentified as galaxies. We approximate the contamination σ_s as being just the original PSF ellipticity. We believe this is a reasonably conservative treatment given the complexity of the effects that combine as different terms in this ellipticity.

To estimate σ_s , we use the *whisker length*. Given I_{xx} , I_{yy} and I_{xy} , the second moment of the light intensity from an object in x , y coordinates, a measure of the ellipticity of the light distribution is given by $e = (I_{xx} - I_{yy})(I_{xx} + I_{yy})$. The whisker length is then defined as $w \approx \sqrt{e(I_{xx} + I_{yy})} = \sqrt{e} \cdot r_{\text{psf}}$, where r_{psf}^2 is given by $(\text{FWHM})/2.35$. FWHM designates the full width at half-maximum and is given by $\text{FWHM} \approx 0.94$ in DES. In addition, the hardware has been designed with a requirement on the whisker length to be lower than a threshold value of 0.2 arcsec in the r , i and z band, which we take as an estimation of *whisk*. We get $C^s \approx 1.3187 \times 10^{-8}$ sr.

Here we consider the two limiting cases $\alpha = 0$ and $\alpha = 1$ and derive the lower bounds for f_s corresponding to each of these cases, referred to as $f_{s,\text{lim},\alpha=0}$ and $f_{s,\text{lim},\alpha=1}$. The true lower bound is in the interval corresponding to these limiting cases: $f_{s,\text{lim}} \in [f_{s,\text{lim},\alpha=1}, f_{s,\text{lim},\alpha=0}]$. In particular, Fig. 3 shows the limiting case $\alpha = 1$: we plot the two terms of the total error $\text{MSE}[p_\alpha]$ (see equation 1), i.e. the systematic parameter shift $\Delta[p_\alpha]$ due to star/galaxy misclassification, and the statistical error $\sigma[p_\alpha]$, for different values of the stellar contamination f_s and for each of the cosmological parameters of our model $p_\alpha = \{w_o, w_a, \Omega_m, H, \sigma_8, \Omega_b, n_s, b_g\}$. For the equation-of-state parameters w_o and w_a , we find that we require $f_s \leq f_{s,\text{lim}}$ with $f_{s,\text{lim},\alpha=0} = 0.122$ and $f_{s,\text{lim},\alpha=1} = 0.022$ (requirement driven by w_a). This translates into the following requirement

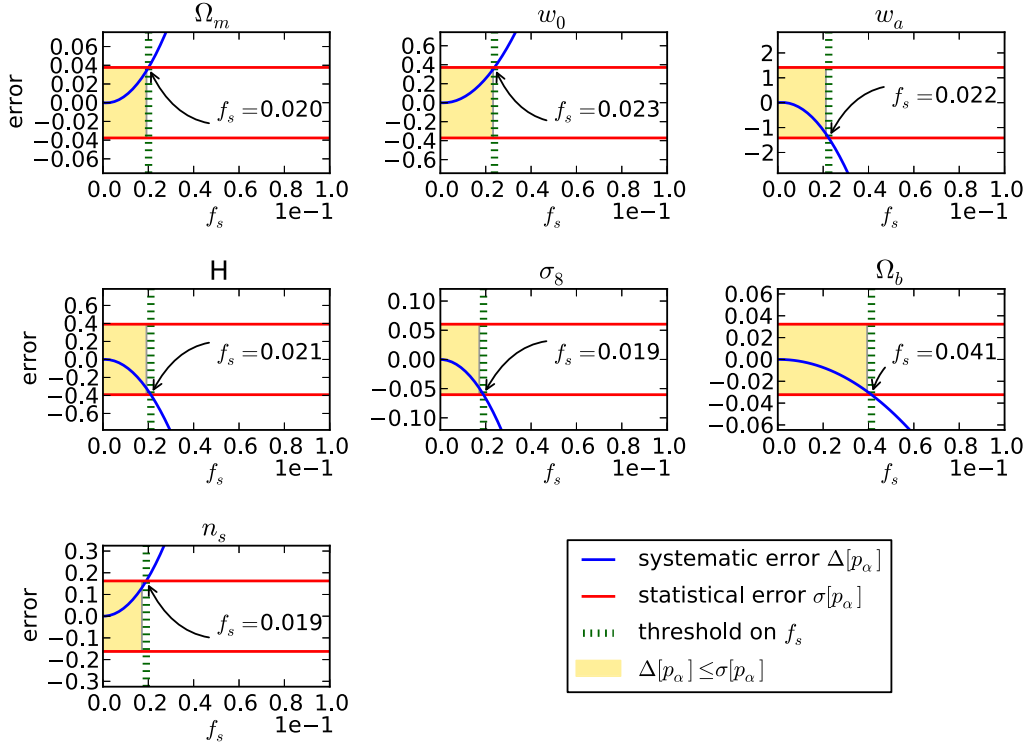


Figure 3. Marginalized statistical error σ (red line) and systematic parameter shift Δ (blue curve) from the WL probe, for different values of the stellar contamination f_s allowed by the star/galaxy classifier. Both σ and Δ are marginalized over $\{\Omega_m, H, \sigma_8, \Omega_b, n_s, b_g\}$ and are computed using the setup described in Section 3.2.1. The yellow area corresponds to the values of f_s for which the requirement on the systematic errors is achieved, i.e. it does not dominate over the statistical error. This requirement translates into a threshold on f_s , indicated by the green line. Unlike LSS measurements, WL measurements are not sensitive to the galaxy bias b_g , which is the reason why it does not appear above.

on $p^g = 1 - f_s$, the purity provided by the star/galaxy classifier: $p^g \geq p_{\text{lim}}^g$ with $p_{\text{lim}}^g \in [87.7 \text{ per cent}, 97.8 \text{ per cent}]$. To refine this requirement, we now allow α to vary. In Fig. 4, we show the evolution of p_{lim}^g when varying α and when considering the requirement on the parameters w_o and w_a . The threshold is driven by w_a (since the requirement to constrain the bias on w_a leads to a more stringent value of p_{lim}^g). The value of p_{lim}^g quickly grows with α . Above $\alpha = 0.4$, p_{lim}^g grows slower and stays above 96 per cent.

Within an experiment designed to constrain DE such as DES, the constraints on the quality of star/galaxy separation comes from the need to control the errors on w_o and w_a . This being said, one should keep in mind that the contamination from stars affects the precision on the measurements of other cosmological parameters, as shown in Fig. 3.

3.3.2 Requirement from LSS measurement

Achieving the objectives of the LSS FoM requires the systematic error induced by star/galaxy misclassification to be smaller than the statistical error on w_o and w_a , and we can rewrite equation (12) in the case of LSS measurements. The shape of the residual systematic signal due to star/galaxy misclassification, δC^{sys} , is obtained following the same methodology as in the WL case, by decomposing the measured density fluctuation into the contribution from the true galaxies and the contamination from the stars identified as galaxies,

$$\delta_m = (1 - f_s)\delta_g + f_s\delta_s. \quad (21)$$

Replacing the shear angular power spectrum with the density fluctuation angular power spectrum in equation (17), we get the

same requirement on the stellar contamination rate f_s as in equation (18). To estimate $C^g(l)$, we use the same stellar catalogue as used for the DES simulated sky survey produced by Busha et al. (2013). We then calculate $C^g(l)$ using the approach from Thomas, Abdalla & Lahav (2010) and an adaptation of the HEALPIX code (Gorski et al. 2005). We estimate the $C_{ij}(l)$ and $\frac{\partial C_{ij}^g(l)}{\partial p_\alpha}$ terms using the same code and setup as for the WL case. Fig. 5 shows the systematic parameter shift induced by the stellar contamination, for each of the cosmological parameters of our model $p_\alpha = \{w_o, w_a, \Omega_m, H, \sigma_8, \Omega_b, n_s, b_g\}$. In particular, for the equation-of-state parameters w_o and w_a , we find that we require $f_s \leq 0.015$. This translates into the following requirement on $p^g = 1 - f_s$, the purity provided by the star/galaxy classifier: $p^g \geq 98.5 \text{ per cent}$. The requirement on star/galaxy separation in a DE experiment is dictated by the need to accurately measure w_o and w_a . This being said, Fig. 5 demonstrates that these two parameters are not the most sensitive to the contamination by stars, which we leave for further analysis.

3.4 Stellar PSF calibration for WL

In this section, we derive two additional requirements on the quality of the star/galaxy separation, from calibration constraints specific to the WL probe. The measured shapes of galaxies include a component due to the PSF of the combined telescope, atmosphere, and instrument which is correlated among galaxies. Removing this contribution requires careful measurement of the PSF, which is done using isolated stars. Therefore, additional requirements on star/galaxy separation come from PSF calibration for WL.

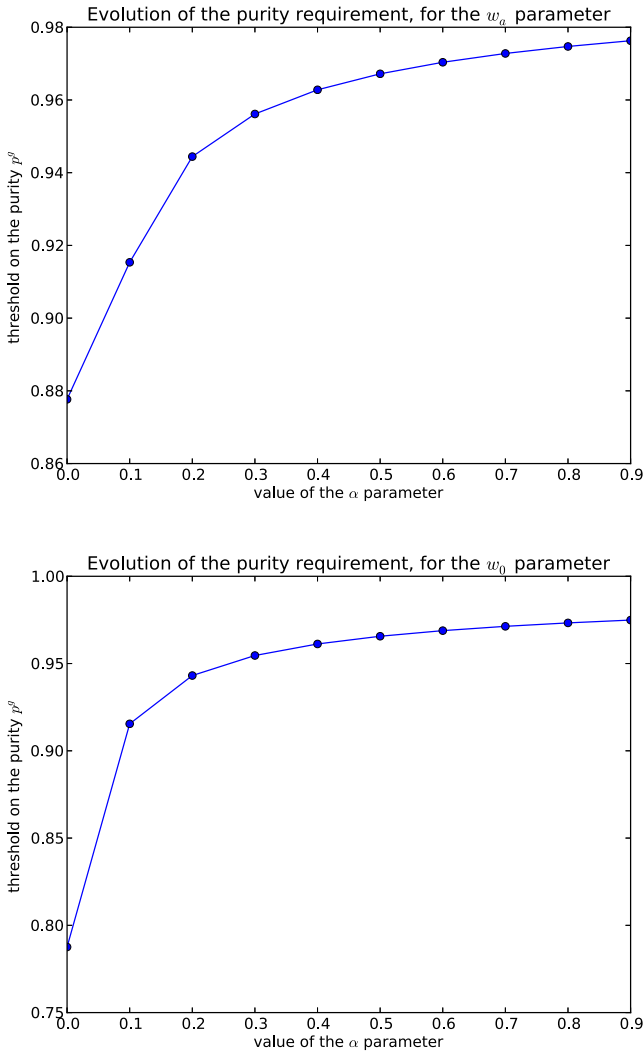


Figure 4. Evolution with the coefficient α of the value of p_{lim}^g , from the constraint on the bias of the equation of state parameter w_a (top) and w_o (bottom).

3.4.1 Requirement on c^s

In order to determine the interpolation pattern of the PSF, one needs to find enough stars to adequately cover the area of the CCD chip. Based on preliminary studies of the DES science verification data, we believe around 200 stars per DES CCD is enough to adequately cover the area of the CCD chip and determine the interpolation pattern of the PSF. From the truth tables, we know that the total number of stars per CCD is approximately 810 and therefore the technical constraint on the completeness of the stars samples is $c^s \geq 25$ per cent.

In this analysis, we assumed that all non-saturated stars can be used for PSF estimation. In practice, the latter lower limit on the completeness could be more stringent because of detector non-linearities. Indeed, the ‘blooming’ effect, caused by the voltages induced by the photons reaching the detector, leads brighter objects to appear larger than faint objects. This effect can lead to variations of the PSF between bright and faint stars, and therefore affect the PSF calibration. This reduces the number of stars available for PSF calibration.

3.4.2 Requirement on p^s

The upper limit on the contamination in a sample of stars comes from the fact that galaxies misclassified as stars will bias the inferred PSF, which in turn will bias the galaxy shapes. We use a toy model to estimate the bias on the shear estimate as a function of $f_g = M_g/(N_s + M_g)$, the galaxy contamination rate in the sample of stars.

Let us first consider the sample of objects classified as stars, used for the calibration of the PSF. Such a sample actually includes two types of objects: true stars and true galaxies which have been misclassified as stars. The PSF model derived from this sample can be approximated as the weighted average of both types of objects:

$$\chi_{\text{psf}}^{\text{biased}}(f_g) = f_g \chi^{\text{mis,gal}} + (1 - f_g) \chi_{\text{psf}}^{\text{true}}, \quad (22)$$

where χ is the polarization, and is related to the observed major and minor axis a and b of the image produced by a circular source via

$$|\chi| = \frac{a^2 - b^2}{a^2 + b^2}, \quad (23)$$

and to the shear and convergence fields via

$$\chi = \frac{2\gamma(1 - \kappa)}{(1 - \kappa)^2 + |\gamma|^2}, \quad (24)$$

so that $|\chi| \approx 2|\gamma|$.

Let us now consider a sample of galaxies of which we would like to measure the shear. The observed polarization χ^{obs} , i.e. the polarization after convolution with the PSF model, is linked to the true polarization of a galaxy through the following relation (Viola, Kitching & Joachimi 2014):

$$\chi_{\text{gal}}^{\text{obs}} = \frac{\chi_{\text{gal}}^{\text{true}}}{1 + 1/R} + \frac{\chi_{\text{psf}}^{\text{true}}}{1 + R}. \quad (25)$$

The *resolution* R in the above equation is the ratio of the galaxy to PSF size. In the absence of misclassified galaxy contaminating the sample used to measure χ_{psf} (and neglecting the other sources of errors in the PSF calibration), the measured polarization is

$$\chi_{\text{gal}}^{\text{true}} = (1 + 1/R) \left(\chi_{\text{gal}}^{\text{obs}} - \frac{\chi_{\text{psf}}^{\text{true}}}{1 + R} \right). \quad (26)$$

However, the contamination from galaxies biases the PSF model, and the measured galaxy polarization is rather

$$\chi_{\text{gal}}^{\text{measured}} = (1 + 1/R') \left(\chi_{\text{gal}}^{\text{obs}} - \frac{\chi_{\text{psf}}^{\text{biased}}}{1 + R'} \right), \quad (27)$$

where $\chi_{\text{psf}}^{\text{biased}}$ is given by equation (22).

As a result, the measured polarization can be written as

$$\chi_{\text{gal}}^{\text{measured}} = (1 + m) \chi_{\text{gal}}^{\text{true}} + c, \quad (28)$$

where

$$m = \frac{R/R' - 1}{R + 1}, \quad (29)$$

and

$$c = \left(1 + \frac{1}{R'} \right) \left(\frac{\chi_{\text{psf}}^{\text{true}}}{1 + R} - \frac{\chi_{\text{psf}}^{\text{biased}}}{1 + R'} \right). \quad (30)$$

The same relation can be written for the shear:

$$\gamma_{\text{gal}}^{\text{measured}} = (1 + m') \gamma_{\text{gal}}^{\text{true}} + c', \quad (31)$$

where $m'^q = m$ and $c' = 2c$.

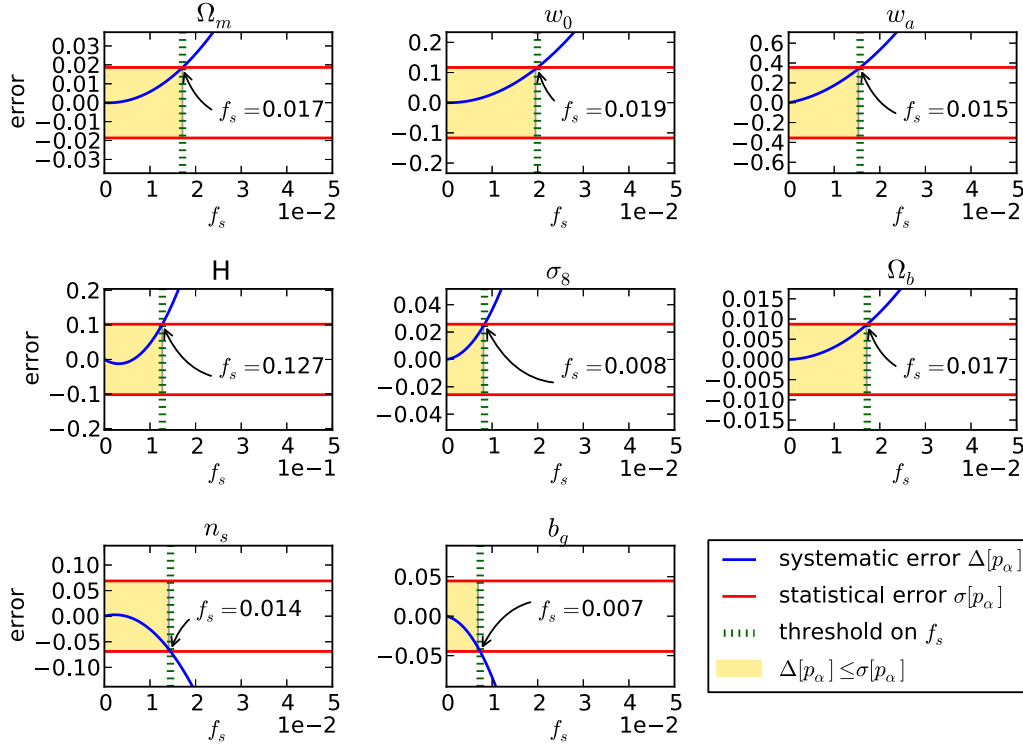


Figure 5. Marginalized statistical error σ (red line) and systematic parameter shift Δ (blue curve) from the LSS probe, for different values of the stellar contamination f_s allowed by the star/galaxy classifier. Both σ and Δ are marginalized over $\{\Omega_m, H, \sigma_8, \Omega_b, n_s, b_g\}$ and are computed using the setup described in Section 3.2.1, with $l \in [10, 400]$, to avoid the non-linear regime. The yellow area corresponds to the values of f_s for which the requirement on the systematic errors is achieved, i.e. it does not dominate over the statistical error. This requirement translates into a threshold on f_s , indicated by the green line. Unlike WL measurements, LSS measurements are sensitive to the galaxy bias b_g , as shown in the last panel.

The **SEXTRACTOR** parameters A_{image} and B_{image} can be used to estimate the typical polarisations. In particular, $\chi_{\text{psf}}^{\text{biased}}$ can be computed as

$$\chi_{\text{psf}}^{\text{biased}}(f_g) = f_g \chi_{\text{psf}}^{\text{mis, gal}} + (1 - f_g) \chi_{\text{psf}}^{\text{true}}, \quad (32)$$

where

- (i) $\chi_{\text{psf}}^{\text{true}} \approx 2\gamma_{\text{psf}}$ and γ_{psf} is estimated as the ellipticity of the PSF,
- (ii) $\chi_{\text{psf}}^{\text{mis, gal}} = \frac{A_{\text{image}}^2 - B_{\text{image}}^2}{A_{\text{image}}^2 + B_{\text{image}}^2}$, for the misclassified galaxies.

R and R' can be computed using the **SEXTRACTOR** parameter **Flux_Radius**:

$$R = \frac{\overline{\text{Flux_Radius}}^{\text{gal}}}{\overline{\text{Flux_Radius}}^{\text{stars}}}, \quad (33)$$

where $\overline{\text{Flux_Radius}}^{\text{stars}}$ is the **Flux_Radius** parameters for the true stars, and

$$R' = \frac{\overline{\text{Flux_Radius}}^{\text{gal}}}{\overline{\text{Flux_Radius}}^{\text{stars+gal.mis}}}, \quad (34)$$

where $\overline{\text{Flux_Radius}}^{\text{stars+gal.mis}}$ is the **Flux_Radius** parameter for all the objects in the sample labelled as stars (i.e. true stars and misclassified galaxies).

Using equations (26) and (32), we can compute $\chi_{\text{gal}}^{\text{measured}}$ and $\chi_{\text{gal}}^{\text{true}}$, as well as the multiplicative and additive biases, m and c . Previous work by the DES collaboration led to the formulation of requirements on the value of m and c : $m < 0.004$ and $|c| < 8 \times 10^{-4}$. These requirements translate into requirements on the contamination from galaxies. In particular, within a toy model in which m and c depend linearly on f_g , i.e. $m = A_m f_g + B_m$ and $c = A_c f_g + B_c$, the

expected² values of the parameters are given by $A_m = 8.6 \times 10^{-2}$, $B_m = -1.6 \times 10^{-3}$, $A_c = -1.0 \times 10^{-1}$ and $B_c = 2.1 \times 10^{-3}$. Therefore, the requirement $m < 0.004$ translates into $f_g < f_{g, \text{lim}}$ with $f_{g, \text{lim}} = 0.07$, i.e. $p^s > p_{\text{lim}}^s$ with $p_{\text{lim}}^s = 93$ per cent. The requirement on the additive bias parameter $|c| < 8 \times 10^{-4}$ leads to a more stringent requirement on the contamination: $f_{g, \text{lim}} = 0.03$, i.e. $p_{\text{lim}}^s = 97$ per cent.

In practice, shear codes have the ability to sharpen the classification of stars and galaxies. Indeed, a shear measurement code convolves a model for the galaxy with the measured PSF function, and then adjusts the parameters of this model to best fit the observed data. If, for example, the best-fitting values for the parameters characterizing the size of the model are too small, it is likely that the observed object is a star (or a very small galaxy). This allows us to perform additional cuts of the sample of objects, using the output of the shear measurement code as an additional indication about the class of the object. For this reason, using the derived verbatim as a requirement on the star/galaxy separation is conservative.

3.5 Summary of the science requirements star/galaxy separation

The requirements on the quality of the star/galaxy separation derived in this section are summarized in Table 2.

A dedicated sample of stars is only needed when calibrating the PSF. Therefore, the two requirements on the samples of stars are

² Sevilla (personal communication).

Table 2. Summary of the science requirements on the quality of star/galaxy separation.

	LSS	WL
p^g	≥ 98.5 per cent (requirement)	$\geq p_{\text{lim}}^g$, with $p_{\text{lim}}^g \in [87.7 \text{ per cent}, 97.8 \text{ per cent}]$ (req. on the systematic error) > 97 per cent (req. on the PSF calibration)
c^g	≥ 88.9 per cent (requirement on the statistical error)	≥ 96.0 per cent (requirement on the statistical error)
c^s	-	≥ 25 per cent (requirement on the PSF calibration)

only required for WL science. As far as samples of galaxies are concerned, LSS science requires purer samples than WL science. This is due to star contamination affecting the corresponding measured ‘observable’ in different ways. The contribution of misclassified stars to the measured shear is dominated by the shot noise term (see equation 19), which is approximately scale independent, whereas they mimic an l -dependent density fluctuation of galaxies and therefore contribute to the LSS measurement in a more complicated way. On the other hand, WL requires a more complete samples of galaxies. This is because a ‘usable’ object means something different for LSS and WL. In order to be usable for LSS measurement, a galaxy needs to be detected with a reliable photometric redshift but WL also needs the shape of the galaxy to be measurable.

In the next sections, we will use these requirements to assess the performance of a new classifier, `multi_class`, and compare it to other classifiers currently used in galaxy surveys. In particular, we will use the most stringent requirement, in the cases of the purities p^g and p^s , i.e. $p_{\text{lim}}^g = 97.8$ per cent and $p_{\text{lim}}^s = 97$ per cent.

4 CURRENT TOOLS FOR STAR-GALAXY SEPARATION

Different strategies have been adopted to classify stars and galaxies in large sky surveys. The morphometric approach (e.g. Kron 1980; Yee 1991; Vasconcellos et al. 2011; Seaborn 1979; Valdes 1982) relies on the separation of point sources (the ones most likely to be stars) from resolved sources (presumably galaxies).

This approach is challenged at the faint magnitudes reached by the next generation of wide-field surveys, due to the vast number of unresolved galaxies.

Another strategy consists of using training algorithms. Machine learning distinguishes several types of learning strategies, Artificial Neural Network (ANN) being one successfully implemented example of *supervised learning*. ANN has previously been applied to the star/galaxy separation problem (e.g. Odewahn et al. 1992; Naim 1995; Bertin & Arnouts 1996; Oyaizu et al. 2008). Indeed, star/galaxy separation shares with many other classification problems the three criteria which usually make neural computing applications particularly successful:

- (i) The task is well defined in that we know precisely what we want, i.e. classify objects in two distinct classes.
- (ii) There is a sufficient amount of data available to train the network to acquire a useful function based on what it should have done in these past examples.
- (iii) No simple parametrization for the output (the class of the object) as a function of the input (the parameters derived from the

images) is known, and we would like to leave it to the algorithm to determine the optimal classification scheme.

Other supervised classifiers, such as support vectors machine, have been more recently used for the star/galaxy separation problem, as well as unsupervised tools such as Bayesian techniques (e.g. Henrion et al. 2011; Fadely, Hogg & Willman 2012).

Throughout this section, we will use the following notations to define:

- (i) *classification tools* – `class_star`; `spread_model` and `multi_class`
- (ii) *classification output* – $X_{\text{class_star}}$; $X_{\text{spread_model}}$ and $X_{\text{multi_class}}$.

As described below (Section 4.1), `class_star` and `spread_model` are two classifiers currently implemented in `SEXTRACTOR` (Bertin & Arnouts 1996) and in the next sections we present a new method for star/galaxy separation called ‘`multi_class`’, designed to achieve the science requirements derived in Section 3 at the faint magnitudes reached by DES.

4.1 Current approaches

Both the morphometric and the training approaches are implemented in `SEXTRACTOR` (Bertin & Arnouts 1996), with two classifiers, `class_star` and `spread_model`.

4.1.1 The training approach – `class_star`

The first classifier to be implemented in `SEXTRACTOR` was `class_star`. Its performance on our example sub-survey is shown in Fig. 6. It uses a set of features of the objects as the input space for a built-in previously trained ANN. These parameters are as follows.

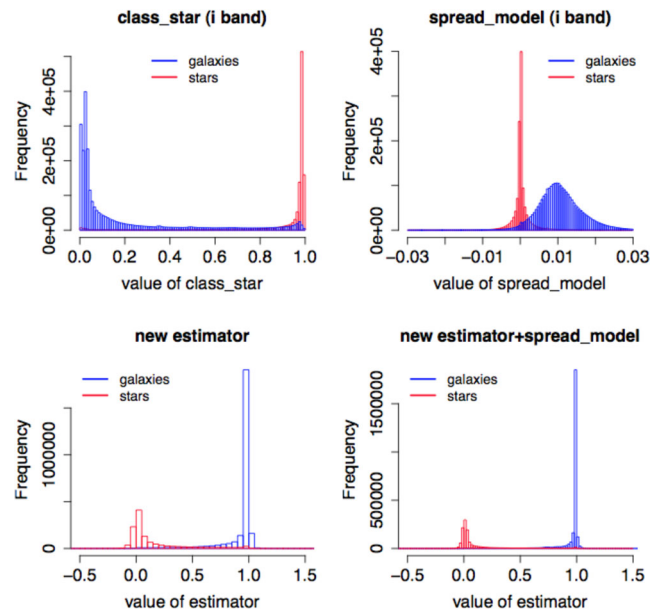


Figure 6. Distribution of the output of all the classifiers presented in the paper. The two upper histograms show the classification performed by `class_star` and `spread_model`. The lower histograms show the classification performed by our new estimator, `multi_class`. On the right one, we incorporate $X_{\text{spread_model}}$ in the input parameters of the ANN. The advantages of plugging $X_{\text{spread_model}}$ into our tool are explained in Section 5.3.2. This allows an increase of the purity for a given completeness, as shown in Fig. 10.

- (i) eight isophotal areas, at regular intervals spanning from the detection threshold to the intensity peak;
- (ii) the intensity peak;
- (iii) the local value of the seeing.

This specific pre-defined set of inputs, chosen mainly for historical reasons, is the main weakness of the `class_star` estimator. The choice of training the ANN on isophotal areas (normalized to the local PSF footprint area) makes it sensitive to close pairs of objects (star–star, star–galaxy, galaxy–galaxy) either blended or de-blended. Since star–star pairs are common on the bright end of the source population, the classifier has a tendency to miss bright, compact galaxies.

More generally speaking, given the large amount and diversity of information encapsulated in the parameters provided by `SEXTRACTOR`, this specific choice of inputs has become hard to justify as it is using a very small part of the available information. The photometry, the shape or the size of an object should also be useful indicators of whether it is a star or a galaxy.

`Class_star` has the advantage of making use of several parameters and combining the information they contain. In this sense it is a ‘multiparameter’ estimator. However, it does not use the most relevant parameters. A more flexible and sensible choice of the inputs is likely to give much better results. This is the main motivation for the new approach tested in this paper.

4.1.2 The morphometric approach – `spread_model`

The morphometric approach was used in several photometric surveys in the past. One possible implementations of this approach, adopted in the Sloan Digital Sky Survey (SDSS) pipeline and in early versions of the DES pipeline, consists of comparing a ‘model magnitude’, i.e. the optimal measure of the magnitude obtained by fitting a galaxy model to the object, to the ‘PSF magnitude’, i.e. the optimal measure of the magnitude determined by fitting a PSF model to the object. A similar strategy was adopted in the Canada-France-Hawaii Telescope Legacy Survey (CFHTLS) pipeline, where classes are assigned to objects according to their half-light radius, i.e. the circular radius which encloses half the light of an object.

The classifier implemented in recent development versions of `SEXTRACTOR` is called `spread_model` (Desai et al. 2012; Bertin et al., in preparation). It carries out diverse operations directly on the image pixels with no use of the object’s parameters generated by `SEXTRACTOR`. The newest version of `spread_model` acts as a linear discriminant between the best-fitting local PSF model ϕ and a slightly ‘fuzzier’ version made from the same PSF model, convolved with a circular exponential model with scalelength given by $\text{FWHM}/16$ (FWHM being the full width at half-maximum of the local PSF model). `spread_model` is normalized to allow for comparison of sources with different PSFs throughout the field. It is defined as

$$X_{\text{spread_model}} = \frac{\phi^T \mathbf{W}x}{\phi^T \mathbf{W}\phi} - \frac{G^T \mathbf{W}x}{G^T \mathbf{W}G}, \quad (35)$$

where x is the image centred on the source, \mathbf{W} is the inverse of the covariance matrix of the pixel noise, which is assumed to be diagonal, ϕ is the PSF and G is the circular exponential model convolved with the PSF. By construction, `spread_model` is close to zero for point sources (most likely to be stars), positive for extended sources (most likely to be galaxies) and negative for detections smaller than the PSF, such as cosmic ray hits.

The performance of this late version of `spread_model` on our example sub-survey is shown in Fig. 6. Although this morphometric approach is quite efficient, it is not entirely satisfying as it does not make use of any of the 300 `SEXTRACTOR` parameters, which are likely to encapsulate a lot of relevant information for star/galaxy separation.

5 THE MULTI_CLASS METHOD

5.1 Motivation and principle

Our goal is to combine the assets of both the morphometric approach and the training approach. We adopt the multiparameter approach allowed by the training method and focus on making the optimal choice of input parameters. The steps of the method are as follows:

- (1) optimal choice of input parameters using a PCA;
- (2) training and running an ANN.

5.2 Step 1 – optimal choice of input parameters using principal component analysis

We make a broad pre-selection of all the parameters likely to be relevant for star/galaxy classification. These parameters are listed in Table 3. They include:

- (i) photometry in five bands (g, r, i, z and y);
- (ii) the size of objects;
- (iii) the shape of objects;
- (iv) the surface brightness of objects;
- (v) qualifiers of the fitting procedure;
- (vi) the output of the `class_star` classifier, $X_{\text{class_star}}$;
- (vii) additional analysis-dependent information.

Ideally, we could run an ANN with this full set of relevant inputs. In practice, training the ANN is a non-linear iterative process, which becomes more time consuming and less robust as the number of input parameters increases. In fact, defining an optimal set of input parameters consists of minimizing its size while maximizing the amount of relevant information it contains.

Our initial set of parameter is redundant, as many of the parameters within each sub-group are dependent variables. For example, we show in Fig. 7 the dependences between four types of magnitudes parameters measured in a given band. In order to reveal the redundancies within the data and compress it, we use a principal component analysis (PCA). This statistical method, which comes down to diagonalizing the covariance matrix of the data, allows us to re-express the pre-selected parameters detailed above in a more meaningful basis of orthogonal, i.e. uncorrelated variables called *principal components*. The first principal component is chosen to account for most of the data variability and thus to have the highest possible variance. Then each succeeding principal component has the highest possible variance under the constraint of being orthogonal – that is uncorrelated – to the preceding one.

We run several ‘well-informed’ PCAs on sub-ensembles of parameters, rather than a ‘blind’ PCA on the full set of initial parameters. We choose to group in these sub-ensembles parameters which have the same units (or measure) and which are linearly dependent on each other (such as the magnitudes in a given band, as shown in Fig. 7). Indeed, when the parameters are linearly dependent, PCA is successful at finding a new basis of meaningful independent variables.

Table 3. DC6 pre-selected parameters, grouped as defined in Section 5.2, by type of information they provide: (i) photometry; (ii) size; (iii) shape; (iv) surface brightness; (v) qualifiers of the fitting procedure; (vi) output of the *class_star* classifier; (vii) additional analysis-dependent information. It should be noted that all of these parameters are distance dependent. The need for *K*-correction to the magnitudes is therefore dealt with by including the photometric redshift in this pre-selected parameters space.

Parameters	Description
(i) <i>mag_aper_</i> in five bands	Fixed aperture magnitude with six different apertures
<i>mag_auto</i> in five bands	Kron-like elliptical aperture magnitude
<i>mag_iso</i> in five bands	Isophotal magnitude
<i>mag_model</i> in five bands	Magnitude from model-fitting
<i>mag_petro</i> in five bands	Petrosian-like elliptical aperture magnitude
<i>mag_psf</i> in five bands	Magnitude from PSF-fitting
<i>mag_spheroid</i> in five bands	Spheroid total magnitude from fitting
(ii) <i>kron_radius</i> (from the detection image)	Kron apertures
(iii) <i>ellipticity</i> (from the detection image)	$1 - B_{\text{image}}/A_{\text{image}}$
(iv) <i>isoarea_world</i> in five bands	Isophotal area above analysis threshold
<i>FWHM_world</i> in five bands	FWHM assuming a Gaussian core
(v) <i>chi2_model</i> in five bands	Reduced chi-square of the fit
<i>chi2_psf</i> in five bands	Reduced chi-square from PSF-fitting
<i>niter_model</i> in five bands	Number of iterations for model-fitting
(vi) <i>X_{class_star}</i> in five bands	Output from <i>class_star</i>
(vii) <i>nlowweight_iso</i>	Number of pixels with low detection weight over the isophotal profile
<i>photoZ</i>	photometric redshift

Our new set of parameters includes uni-band parameters from the initial set (such as the photometric redshift or the ellipticity), as well as the principal components from the PCAs listed below:

- (i) PCA on the five bands of each multiband parameter;
- (ii) PCA on the six fixed-aperture magnitudes in each band;
- (iii) PCA on the six other types of magnitudes in each band (i.e. *mag_auto*, *mag_iso*, *mag_model*, *mag_petro*, *mag_spheroid* and *mag_psf*).

Fig. 8 shows the variances of the principal components of these six types of magnitudes in each band as a function of their index. Each of these PCAs shows that most of the variance of the data is encapsulated in a reduced number of principal components. In many cases, using PCA for data reduction consists of selecting only the principal components with the highest variance and approximating the data by its projection on this smaller set of variables.

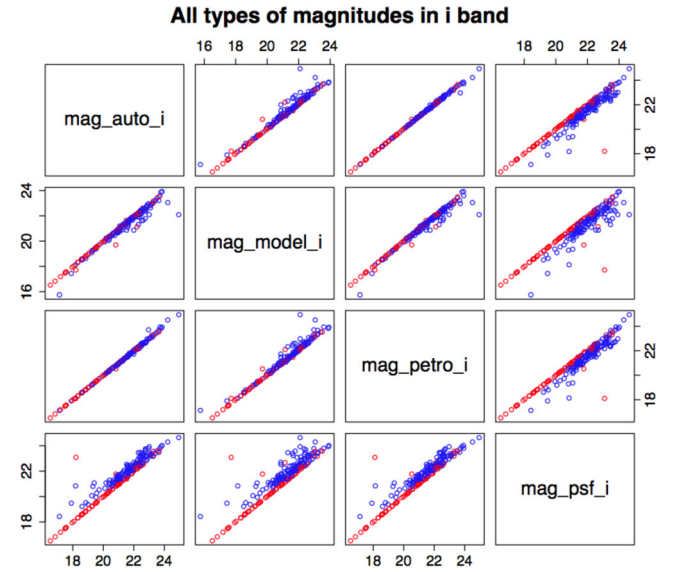


Figure 7. Scatter plots for stars (red markers) and galaxies (blue markers) for four different types of magnitudes in the *i* band. The magnitudes are strongly correlated and PCA is therefore well adapted to re-express them in a new basis of independent variables.

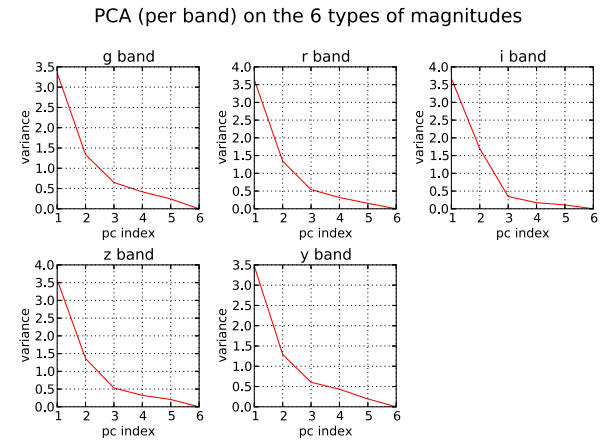


Figure 8. Value of the variance of the principal components as a function of their index for the five (per-band) PCAs performed on the six types of magnitudes: *mag_auto*, *mag_iso*, *mag_model*, *mag_petro*, *mag_spheroid* and *mag_psf*.

This encompasses the assumption that the important information is represented by the components with the highest variances. In the case of star/galaxy separation, this assumption is too simplistic. Indeed, the class of an object is only one possible source of variance and high variance could also be due to differences between objects in a given class. Therefore, when looking for the most relevant components for star/galaxy separation, we need another criterion to quantify their aptitude to separate between the classes. We calculate the *Fisher discriminant* (Fisher 1936) for each of the new parameters, defined as the inter-class variance over the intra-class variance,

$$\mathcal{F}_i = \frac{(\overline{X}_{G,i} - \overline{X}_{S,i})^2}{\sigma_{G,i}^2 + \sigma_{S,i}^2}, \quad (36)$$

where $\overline{X}_{A,i}$ is the empirical mean value of the *i*th parameter for class A and $\sigma_{A,i}^2$ is its empirical variance. Fig. 9 shows the classification

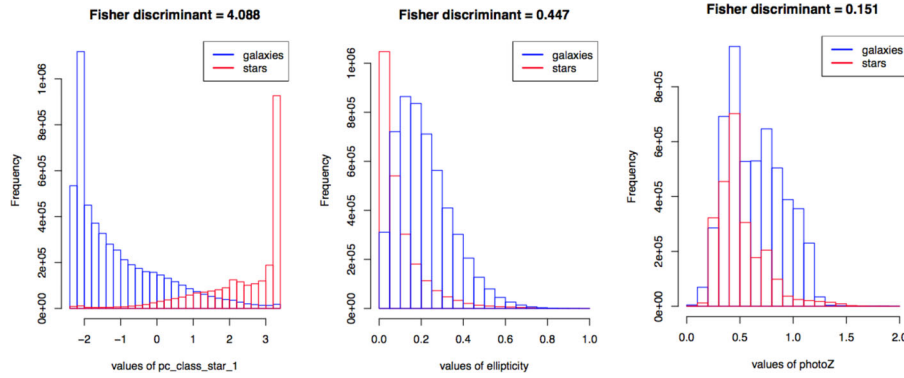


Figure 9. Distribution of the three parameters with the highest Fisher discriminant, for stars and galaxies as indicated in the figure. *pc_class_star_1* (left) is the first principal component from a PCA performed on the five bands of $X_{\text{class_star}}$ (see Section 5.3.2). The two other parameters shown, *ellipticity* (centre) and *photoZ* (left) have not gone through any PCA.

performed by the three parameters with the highest Fisher discriminant. The 15 parameters with the highest Fisher discriminant form our final set of input parameters for the ANN (as discussed in Section 5.3.1, more than 20 input parameters make the ANN less robust, so we limit the basic set to 15 parameters, in anticipation of the other five that will be added in Section 5.3.2).

The set of parameters with the highest Fisher discriminant is data specific. In our analysis of DC6, the most discriminative combinations of parameters are those of the class_star different bands, (group (vi) of Table 3), morphometric parameters such as the ellipticity (group (ii) of Table 3), followed by photometric parameters such as the photometric redshift and the magnitudes (group (i) of Table 3). This being said, when generalizing the method to other data sets, the Fisher discriminant should be recalculated.

5.3 Step 2 – running an ANN on the optimal inputs space

Once a set of optimal parameters is defined, the next step consists of mapping these parameters to the class of the objects. This mapping is performed by training an ANN.

5.3.1 ANN: principle and advantages

In essence, an ANN is a highly-flexible, fully non-linear fitting algorithm. During the training phase, it receives a set of input patterns and a given property (in our case the class of the object), which needs to be fitted to them. The training consists of several iterations during which a number of free parameters known as *weights* are adjusted so as to minimize the difference between the outputs of the neural network for each pattern and the desired property. The algorithm then learns how to link the inputs to the desired property. After the training phase, the ANN can be used to infer this property from a set of input objects for which it is unknown. For our analysis, we train an ANN to map the set of optimal input parameters selected in Section 5.2 to the class of the object (star or galaxy) on a sample of objects for which the answer is known (the training is made on the DC6 simulations for which we know the true class of each object). The ANN is then used to deduce the class of a distinct set of objects.

An ANN is made of computing units called *neurons*, arranged in several layers and connected by synapses in which the information flows in a single direction. The complexity of the network depends on the number of layers and neurons in each layer. We chose to use the ANNZ photometric redshift code (Collister & Lahav 2004),

which was originally designed for photometric redshift measurements, but can be effectively and straightforwardly applied to our classification problem. The trade-off between the complexity of the network and its performance has been investigated by Firth, Lahav & Somerville (2003). For the same number of parameters, adding extra hidden layers is found to give greater gains than widening existing layers. As the network complexity is increased, the accuracy eventually converges so that no further improvement is gained by adding additional nodes. We chose a network architecture with an input layer of 15 parameters (or twenty, as explained in the next section) and two hidden layers of twenty nodes, which turns out to be sufficiently complex for such convergence to be achieved.

Training on real data, as opposed to simulations, is preferable, yet more challenging. One option would be to use data from space-based surveys, as in space the PSF is not affected by the seeing. Data from the *HST* could be used to train our tool for the real DES survey data.

5.3.2 Plugging other classifiers in the method

Using an ANN brings flexibility to the training approach. It allows us not only to choose which inputs to use, but also in what number. In particular, we can take the output of other classifiers as inputs to our method.

We run a PCA on the five $X_{\text{class_stars}}$ (in the five bands). Not surprisingly, the first principal component has a high Fisher discriminant (as shown in Fig. 9) and is therefore included in the 15 input parameters selected in Section 5.2. As the five bands of $X_{\text{spread_model}}$ are less clearly linearly dependent, we choose not to run a PCA on them and add the five $X_{\text{spread_model}}$ to the set of 15 input parameters, which amounts to 20 input parameters.

Fig. 10 presents the purity level at a given completeness for these two different configurations of our method. The performance of our method with fifteen input parameters (orange curve) can be compared to the performance when plugging in $X_{\text{spread_model}}$ (pink curve). Including $X_{\text{spread_model}}$ in the inputs allows an increase in the level of the purity by 2 per cent at faint magnitudes. Running the ANN on the fifteen preselected parameters (orange curve) already gives better results than spread_model (blue curve) for most of the magnitude range (except for the very faint magnitudes, in the galaxies case). However, the best results are obtained by combining the two, i.e. by running the ANN on a hybrid input space combining the 15 selected parameters and $X_{\text{spread_model}}$.

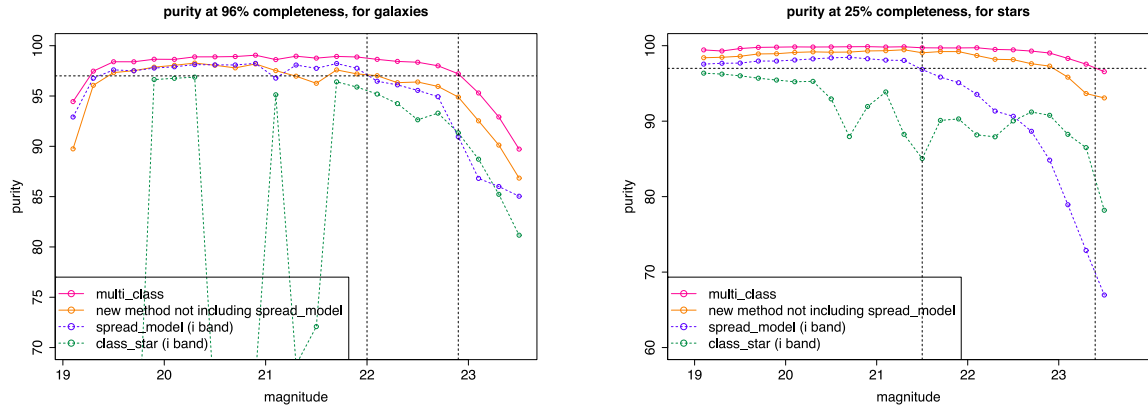


Figure 10. Purity level at the required completeness, for the WL probe, as a function of magnitude in the i band. The orange and pink curves correspond to different versions of our method, whereas the blue and green ones show the performance of the classifiers `class_star` and `spread_model`. The orange curve is obtained when running the ANN on the 15 parameters selected in Section 5.2 and the pink curve, the final version of `multi_class`, is obtained when adding `spread_model` in five bands to this set of inputs. The dashed horizontal line shows the science requirement from WL science on p^g (97.8 per cent, Section 3.3.1) and p^s (97.0 per cent, Section 3.4). The requirement on p^g is achieved by `multi_class` up to magnitudes of 22.9, whereas `spread_model` only allows us to reach 22.0. The requirement on p^s is achieved up to magnitudes of 23.4 with `multi_class`, versus 21.5 with `spread_model`.

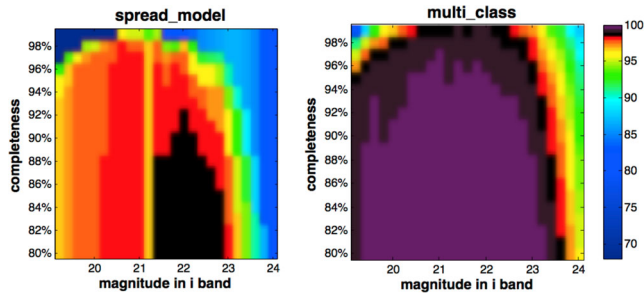


Figure 11. Level of purity for a sample of galaxies p^g , for different magnitudes and values of the completeness. The 98.5 per cent level requirement from LSS (Section 3.3.2) is shown in purple, and the 97.8 per cent limit required for WL (Section 3.3.1) is shown in black. `spread_model` does not allow us to achieve the LSS requirement, which `multi_class` can reach. `multi_class` also allows us to achieve the requirement from WL at fainter magnitudes than `spread_model`.

6 CLASSIFICATION RESULTS

We showed that we can optimise our classifier performance by using a ‘well-informed’ PCA strategy (Section 5.2), and by incorporating $X_{\text{spread_model}}$ into the method (Section 5.3.2). We now compare our classifier performance to the one of the other classifiers. We will focus on comparing `multi_class` to `spread_model`, as the performance of `class_star` is widely surpassed by both `spread_model` and `multi_class` for most of the magnitude range (as shown in Fig. 10).

For LSS, our new classifier allows us to achieve requirements which cannot be fulfilled by `spread_model`. Fig. 11 shows that the 98.5 per cent limit on p^g (derived in Section 3.3.2 and shown in purple on the figure) cannot be reached by `spread_model`, whereas `multi_class` allows us to reach it up to magnitudes of 22.9 (at the required 88.9 per cent completeness level, derived in Section 3.2.2).

For WL, `multi_class` allows us to increase the magnitude limit below which the science requirements are achieved. Fig. 10 shows that this magnitude limit increases from 21.5 to 23.4 for the requirement on the stars purity p^s , and from 22.0 to 22.9 for the requirement on the galaxy purity p^g . Fig. 11 and Fig. 12 generalize this to a broad range of completenesses. In Fig. 13, we consider the improvement in the purity of a sample of stars and a sample of galaxies, as a

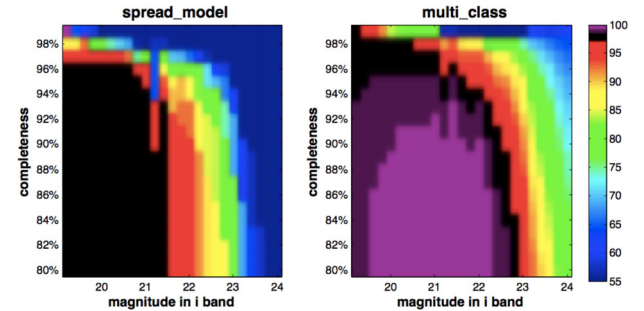


Figure 12. Level of purity for a sample of stars p^s , for different magnitudes and values of the completeness. The 97 per cent science requirement (from WL, derived in Section 3.4) is shown in black. Higher purity levels are shown in purple and light purple. Our new estimator, `multi_class`, allows us to widen the range of both magnitude and completeness where this requirement is achieved.

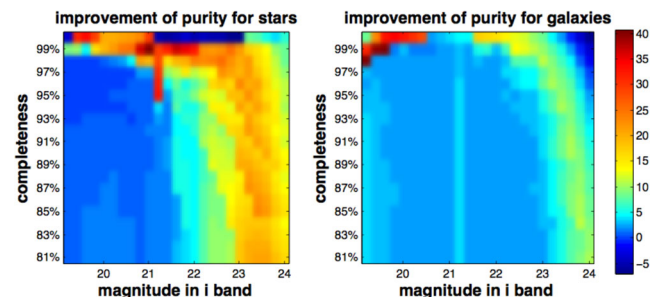


Figure 13. Difference of the purity level achieved by `multi_class` and `spread_model`, $p_{\text{multi_class}} - p_{\text{spread_model}}$ for stars (left) and galaxies. At faint magnitudes (ranging from 23 to 24), `multi_class` allows us to increase the level of p^s achieved by `spread_model` by up to 20 per cent, and p^g by up to 12 per cent.

function of magnitude, for a large range of completenesses. At faint magnitudes – typically fainter than 23 – `multi_class` improves the purity achieved by `spread_model` by up to 12 per cent for galaxies and by up to 20 per cent for stars.

7 CONCLUSIONS

We showed that star/galaxy misclassification contributes to both the statistical and systematic error on the measurement of cosmological parameters. In particular, it affects the measurement of the DE equation-of-state parameters, w_o and w_a , which future large photometric surveys such as DES aim to measure accurately. In the case of WL and LSS measurements, we translated the DETF FoM requirements on the statistical and systematic errors and the constraints from PSF calibration into the corresponding science requirements on the quality of star/galaxy separation. We formulated these requirements using two parameters: the purity and completeness of classified samples of stars and galaxy.

In order to meet these new requirements, we built an efficient method for star/galaxy classification, called `multi_class`, which combines a PCA with a learning algorithm. Our multiparameter approach allows us to make use of the huge amount of information provided by `SEXTRACTOR`. In particular, the use of PCA allows us to better understand the correlations in the data, and to implement this physical knowledge in the classifier.

In ground-based surveys such as DES, the image quality is not constant with sky position and therefore any purely morphometric method gives limited performance, especially at faint magnitudes. The flexibility of using an ANN allows us to consider the morphometry as one input parameters among many others and to integrate the performance of other classifiers to our new tool. Our new classifier, `multi_class`, significantly improves the performance of the morphometric classifier implemented in `SEXTRACTOR` (`spread_model`), which cannot achieve the LSS science requirements on star/galaxy separation. For both the LSS and WL probes, it allows us to widen the range of both magnitude and completeness where the derived science requirements are achieved. For magnitudes fainter than 23, `multi_class` improves the purity achieved by `spread_model` by up to 12 per cent for galaxies and by up to 20 per cent for stars.

DES began survey operations in 2013 September, and will be running for five years. Therefore, we should be able to test the results shown in this paper on real data in the near future. The faint magnitudes reached by this new classifier constitute an important asset, which should allow us to achieve the science requirements on star/galaxy separation in the next generation of wide-field photometric surveys.

ACKNOWLEDGEMENTS

MTS would like to thank Gary Bernstein, Benjamin Joachimi and Alan Heavens for very useful comments and advice; Alexandre Refregier for a very useful discussion; and Ashley Ross, Adam Hawken, Manda Banerji, Alex Merson, Foteini Oikonomou, Boris Leistedt, Sreekumar Balan and Iftach Sadeh for their input to the project. MTS is grateful for the support from the University College London Perren and Impact studentships. FBA acknowledges the support of the Royal Society via a University Research Fellowship. OL acknowledges a Royal Society Wolfson Research Merit Award, a Leverhulme Senior Research Fellowship and an Advanced Grant from the European Research Council. BR acknowledges support from the European Research Council in the form of a Starting Grant with number 240672. We acknowledge UK's STFC for supporting DES optics and science.

Funding for the DES Projects has been provided by the US Department of Energy, the US National Science Foundation, the Ministry of Science and Education of Spain, the Science and Technology Facilities Council of the United Kingdom, the Higher

Education Funding Council for England, the National Center for Supercomputing Applications at the University of Illinois at Urbana-Champaign, the Kavli Institute of Cosmological Physics at the University of Chicago, Financiadora de Estudos e Projetos, Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro, Conselho Nacional de Desenvolvimento Científico e Tecnológico and the Ministério da Ciência e Tecnologia, the Deutsche Forschungsgemeinschaft and the Collaborating Institutions in the DES.

The Collaborating Institutions are Argonne National Laboratories, the University of California at Santa Cruz, the University of Cambridge, Centro de Investigaciones Energeticas, Medioambientales y Tecnologicas-Madrid, the University of Chicago, University College London, DES-Brazil, Fermilab, the University of Edinburgh, the University of Illinois at Urbana-Champaign, the Institut de Ciències de l'Espai (IEEC/CSIC), the Institut de Física d'Altes Energies, the Lawrence Berkeley National Laboratory, the Ludwig-Maximilians Universität and the associated Excellence Cluster Universe, the University of Michigan, the National Optical Astronomy Observatory, the University of Nottingham, the Ohio State University, the University of Pennsylvania, the University of Portsmouth, SLAC, Stanford University, the University of Sussex, Texas A&M University, and the Institute of Astronomy at ETH-Zurich.

REFERENCES

- Albrecht A. et al., 2006, preprint ([arXiv:astro-ph/0609591](https://arxiv.org/abs/astro-ph/0609591))
- Amara A., Refregier A., 2008, MNRAS, 391, 228
- Audren B., Lesgourgues J., Bird S., Haehnelt M. G., Viel M., 2013, J. Cosmol. Astropart. Phys., 1, 26
- Bartelmann M., Schneider P., 2001, Phys. Rep., 340, 291
- Bertin E., Arnouts S., 1996, A&A, 393, 404
- Busha M. T., Wechsler R. H., Becker M. R., Erickson B., Evrard A. E., 2013, Am. Astron. Soc. AAS Meeting #221, 341.07
- Collister A., Lahav O., 2004, PASP, 345, 351
- Das S., de Putter R., Linder E. V., Nakajima R., 2012, J. Cosmol. Astropart. Phys., 11, 011
- Debono I., Rassat A., Refregier A., Amara A., Kitching T., 2010, MNRAS, 404, 110
- DES Collaboration, 2005, preprint ([arXiv:astro-ph/0510346](https://arxiv.org/abs/astro-ph/0510346))
- Desai S. et al., 2012, ApJ, 757, 83
- Fadely R., Hogg D. W., Willman B., 2012, ApJ, 760, 15
- Firth A. E., Lahav O., Somerville R. S., 2003, MNRAS, 339, 1195
- Fisher R. A., 1936, Ann. Eugenics, 179, 1936
- Gorski K. M., Hivon E., Banday A. J., Wandelt B. D., Hansen F. K., Reinecke M., Bartelmann M., 2005, AJ, 622, 759
- Henion M., Mortlock D. J., Hand D. J., Gandy A., 2011, MNRAS, 412, 2286
- Huterer D., Takada M., Bernstein G., Jain B., 2006, MNRAS, 366, 101
- Jarvis M., Bernstein G., Jain N., 2004, MNRAS, 352, 338
- Kirk D., Laszlo I., Bridle S., Bean R., 2013, MNRAS, 430, 197
- Kirk D., Rassat A., Host O., Bridle S., 2012, MNRAS, 424, 1647
- Kron R. G., 1980, ApJS, 43, 305
- Laszlo I., Bean R., Kirk D., Bridle S., 2012, MNRAS, 423, 1750
- Maddox S. J., Efstathiou G., Sutherland W. J., 1990, MNRAS, 246, 433
- Mellier Y., 1999, ARA&A, 37, 127
- Mohr J. J. et al., 2012, Proc. SPIE, 8451, 84510D
- Naim A., 1995, PhD thesis, Univ. Cambridge
- Odewahn S. C., Stockwell E. B., Pennington R. L., Humphreys R. M., Zumach W. A., 1992, AJ, 103, 318
- Oyaizu H., Lima M., Cunha C. E., Lin H., Frieman J., Sheldon E. S., 2008, ApJ, 674, 768
- Rassat A. et al., 2008, preprint ([arXiv:0810.0003](https://arxiv.org/abs/0810.0003))
- Refregier A., 2003, ARA&A, 41, 645
- Ross A. et al., 2011, MNRAS, 417, 1350

- Rossetto B. M. et al., 2011, AJ, 141, 185
 Sebok W. L., 1979, AJ, 84, 1526
 Takad M., Jain B., 2004, MNRAS, 348, 897
 Tegmark M., Taylor A. N., Heavens A. F., 1997, ApJ, 480, 22
 Thomas S., Abdalla F. B., Lahav O., 2010, MNRAS, 1669, 1685
 Thomas S., Abdalla F. B., Lahav O., 2011, Phys. Rev. Lett., 106, 241301
 Valdes F., 1982, Proc. SPIE, 465, 472
 Vasconcellos E. C., de Carvalho R. R., Gal R. R., LaBarbera F. L., Capelato H. V., Frago Campos Velho H., Trevisan M., Ruiz R. S. R., 2011, AJ, 141, 189
 Viola M., Kitching T. D., Joachimi B., 2014, MNRAS, 439, 1909
 Weinberg D. H., Mortonson M. J., Eisenstein D. J., Hirata C., Riess A. G., Rozo E., 2013, Phys. Rep., 530, 87
 Yee H. K. C., 1991, PASP, 103, 396

APPENDIX A: STATISTICAL ERRORS ON THE COSMOLOGICAL PARAMETERS $\{\Omega_m, H, \sigma_8, \Omega_b, n_s, b_g\}$ FROM WL AND LSS MEASUREMENTS

In the following, we show the marginalized statistical errors on $\{\Omega_m, H, \sigma_8, \Omega_b, n_s, b_g\}$ from the WL probe (Fig. A1) and the LSS probe (Fig. A2), for different values of the density of galaxies with reliable shape measurement N_{eff} and of the density of detected galaxies N_G , respectively. The errors are marginalized and computed using the assumptions and setup described in Section 3.2.1, with $l \in [1, 1024]$ in the WL case and with $l \in [10, 400]$ in the LSS case. The red curve shows the errors computed with a non-informative prior whereas the blue curve is obtained assuming a Planck prior.

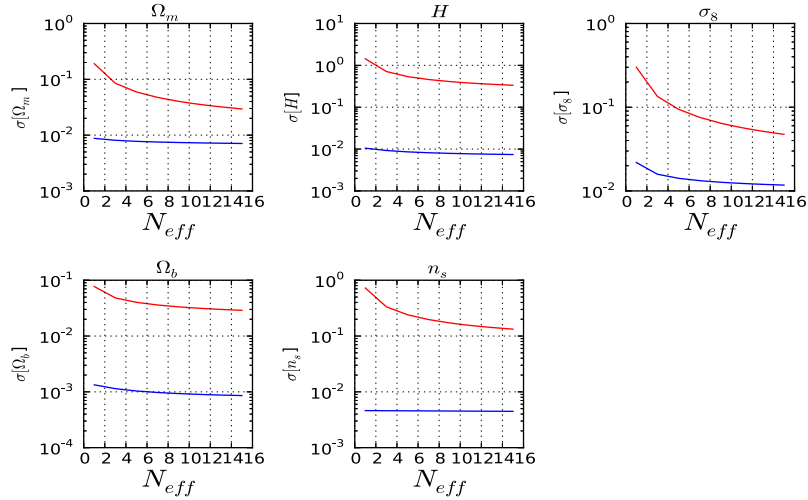


Figure A1. Marginalized statistical errors on $\{\Omega_m, H, \sigma_8, \Omega_b, n_s, b_g\}$ from the WL probe.

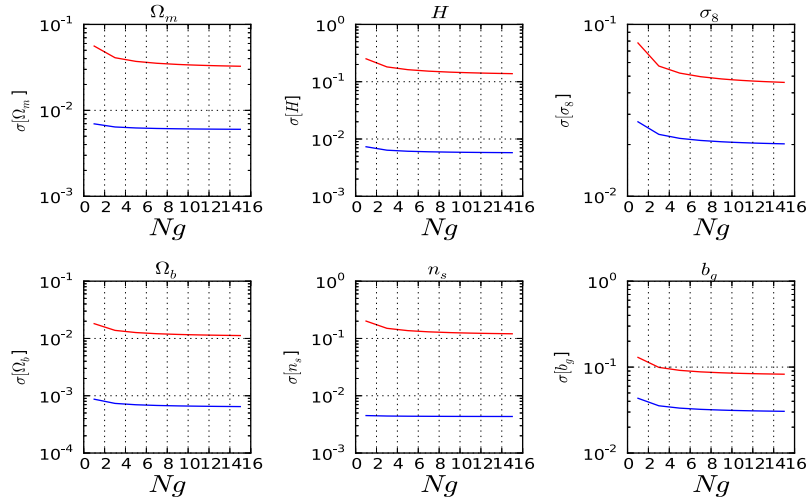


Figure A2. Marginalized statistical errors on $\{\Omega_m, H, \sigma_8, \Omega_b, n_s, b_g\}$ from the LSS probe.

This paper has been typeset from a \LaTeX file prepared by the author.