

# UNIONS-3500: a weak lensing catalogue for the northern sky

Fabian Hervas-Peters<sup>1</sup>, Sacha Guerrini<sup>2</sup>, Martin Kilbinger<sup>1</sup>, and others<sup>2,\*</sup>

<sup>1</sup> AIM, CEA, CNRS, Université Paris-Saclay, Université de Paris, F-91191 Gif-sur-Yvette, France  
e-mail: martin.kilbinger@cea.fr

<sup>2</sup> Other place

Received; accepted

## ABSTRACT

**Context.** Over the last two decades, weak lensing has become a widely used technique to characterise the dark matter distribution in the large-scale structure of the Universe. It is particularly used due to its unbiased nature as it purely relies on gravitational effects and bypasses any assumption connecting dark and luminous matter.

**Aims.** In this work, we present the first weak gravitational lensing catalogue for cosmological constraints with cosmic shear of the Ultraviolet Near Infrared Optical Northern Survey (UNIONS). We use 3,500 square degrees of sky area in the Northern Hemisphere, observed in the *r*-band by MegaCAM on the Canada-France Hawaïi Telescope, achieving a pristine median seeing of 0.7 arcsec.

**Methods.** Starting from images calibrated for astrometry and photometry, we describe the steps from image processing to catalogue creation. These steps include masking, source detection and selection, star selection, PSF modelling, shape measurement and calibration. We conduct extensive validation tests, particularly to assess and mitigate the leakage of PSF ellipticity into galaxy shapes. We demonstrate the robustness of the catalogue through the absence of correlation with other observational variables and structural elements, such as position in the image or proximity to bright stars.

**Results.** The final galaxy catalogue contains XX million galaxies, corresponding to a source density of  $x \text{ arcmin}^{-2}$ . The ellipticity dispersion, commonly referred to as shape noise is  $\sigma_\epsilon =$ . By covering the northern sky, it offers a large overlap with spectroscopy of the Sloan Digital Sky Survey and the Dark Energy Spectroscopic Instrument.

**Conclusions.** Initiating the first major cosmological analysis by the UNIONS collaboration, this is the first in a series of five papers which cover the various aspects of a robust cosmic shear analysis. Two companion papers discuss the robustness of the catalogue, one through the level of *B*-mode contamination and another by producing and analysing dedicated image simulations for shear calibration.

**Key words.** Cosmology – Large scale structure – Weak lensing – Imaging systematics – Shape Measurement

## 1. Introduction

Cosmic shear denotes the coherent distortion of galaxy images by the intervening large-scale structure through the effect of gravitational lensing. Since it is sensitive to the total matter content, this technique is of primary interest to map the matter distribution in the Universe, which is dominated by dark matter. Cosmic shear probes both the evolution of structure and the geometry of the Universe. Over the last two decades, it has become a main probe of cosmological parameters such as the matter density parameter  $\Omega_m$ , the matter density power-spectrum normalisation  $\sigma_8$ , or the dark-energy parameter of state  $w$ . For reviews, see (e.g., Kilbinger 2015; Mandelbaum 2018; Prat & Bacon 2026).

While the standard goal of a wide shear catalogue is a cosmological cosmic shear analysis, multiple measurements are enabled when probing the lensing field. These include the mass and profile measurement of different objects such as haloes (e.g., Li et al. 2025), galaxies (e.g., Huang et al. 2020), the detection of filaments (e.g., Xia et al. 2020), the study of voids distributions and profiles (e.g., Melchior et al. 2014), intrinsic alignments of galaxies (e.g., Singh et al. 2015) and mass maps (e.g., Jeffrey et al. 2021) to name some examples. This plethora of applications has made weak lensing catalogues a particularly rich and sought-after source of cosmological and astrophysical information.

Galaxy shape measurement algorithms have evolved significantly in the past two decades in response to the biases affecting weak-lensing shear estimation. Broadly, existing approaches fall into two categories: moments-based methods and model-fitting techniques. Moments-based estimators compute weighted integrals of the light distribution involving powers of the spatial coordinates and therefore avoid strong assumptions about galaxy morphology (Kaiser et al. 1995; Bernstein & Jarvis 2002; Hirata & Seljak 2003). In contrast, model-fitting approaches assume a parametric description of galaxy structure and optimize parameters against pixel data; the Sérsic profile (Sérsic 1963), either single-component or bulge–disk decomposed, remains the most widely adopted model, forming the basis of algorithms such as *lensfit* (Miller et al. 2013) used in the Kilo-Degree Survey (Giblin et al. 2021b) and *im3shape* (Zuntz et al. 2013) employed in DES Y1 analyses (Jarvis et al. 2016; Zuntz et al. 2018). More recently, *LensMC* uses an MCMC-based inference framework for *Euclid* (Congedo et al. 2024). While some works tried to reach minimally biased estimators, the field has progressively shifted toward calibration-driven strategies (Mandelbaum et al. 2014, 2015), either through realistic simulations or data-driven corrections. Modern surveys increasingly adopt self-calibrating techniques: in the Dark Energy Survey, Metacalibration (Huff & Mandelbaum 2017; Sheldon & Huff 2017; Gatti et al. 2021) and Metadetection (Sheldon et al. 2020; Yamamoto et al. 2025) estimate shear responses directly from the data by artificially shearing observed images, leveraging the simplicity of Gaussian

\* Just to show the usage of the elements in the author field

profile convolutions (Hogg & Lang 2013). The Kilo-Degree Survey recently applied Metacalibration with a moments-based approach (Yoon et al. 2025). Following the logic of the Dark Energy Survey Y3 (Gatti et al. 2021), we use a Gaussian mixture model `ngmix` (Sheldon 2015) with Metacalibration in this work.

This paper is the first in a series of five publications. This first paper presents the UNIONS galaxy catalogue and provides an overview of weak lensing image processing and overall calibration with ShapePipe (Farrens et al. 2022). The second paper in the series focuses on validation of the shear calibration with image simulations (Hervas-Peters et al. 2026). Paper III (Daley et al. 2026) validates the catalogue using B-mode tests and derives scale cuts. Papers IV and V present cosmological constraints in configuration space (Goh et al. 2026) and harmonic space (Guerrini et al. 2026), respectively.

UNIONS is a very unique dataset for cosmology in many respects. The Northern Hemisphere has been rather neglected by photometric surveys. In terms of the combination of depth, area, and image quality, UNIONS images are unprecedented: They are deeper than the Sloan Digital Sky Survey (York et al. 2000) and the Pan-STARRS  $3\pi$  survey Chambers et al. (2019), and cover significantly more area than Subaru Strategic Program (SSP) survey with the Hyper Suprime-Cam (HSC; Li et al. 2023b). Narrow-band optical images from UNIONS images will remain state-of-the-art for many years to come.

With observations in the  $r$ -band having started in 2017, UNIONS is the last of the stage-III surveys. Other weak-lensing Stage-III surveys, notably KiDS, DES, and HSC, have recently published results that build on previous iterations and data releases; in part, they correspond to joint analyses (DES and KiDS Collaboration et al. 2023; Jefferson et al. 2025). The UNIONS results presented here and in the companion papers are completely independent of those other surveys, and correspond to new constraints on cosmological parameters.

The wide-field imager MegaCAM Boulade et al. (2003) on the Canada-France Hawai’i Telescope (CFHT) remains, despite its age, one of the best optical cameras in the world. The CFHT has a long history of providing data for weak gravitational lensing measurements, including the first cosmic shear detection Van Waerbeke et al. (2000). Dedicated cosmic shear surveys were CFHTLS Hoekstra et al. (2006), and, using the same data, CFHTLenS Heymans et al. (2012). CFHTLS published cosmic shear results just after the quest of dark energy was placed into stage Albrecht et al. (2006), and can therefore be considered as one of the first Stage-II surveys. UNIONS, as coming out at the dawn of the next-generation surveys such as Euclid, LSST, and Roman, can be seen as the last of the Stage-III surveys.

The first weak-lensing analysis of UNIONS data was presented in Guinot et al. (2022), which analysed 1,600 square degrees of sky area in the  $r$ -band. Further, processing of a larger area of around 3,200 deg $^2$  was used in Li et al. (2024). The present paper corresponds to the analysis of 3,500 deg $^2$ .

This paper is organised as follows. Sect. 2 gives an overview of the UNIONS data and weak-lensing processing. Sect. 3 discusses shape measurement, shear estimation and calibration. In Sect. 5 the effect of the point-spread function (PSF) is quantified.

## 2. UNIONS data

The targeting of UNIONS, as described in Gwyn et al. (2025) can be initially traced back to the CFHT  $u$ -band survey driven by near field cosmology. When the need for complementary optical coverage for *Euclid* became more pressing, the targeting was

extended according to principles described in Euclid Collaboration et al. (2025). The target area was selected to be minimally contaminated by the galactic plane; objects from the solar system and zodiacal light. This means that at the end of the survey, UNIONS will have observed a large fraction of the darkest extragalactic sky available, least subject to galactic extinction. The survey was designed with a strong homogeneity target, which was particularly successful for the  $r$ -band when considering the spread of the point source depth distribution and both the median and width of the FWHM distribution of point sources.

### 2.1. Catalogue versions

Three weak-lensing processing campaigns of UNIONS  $r$ -band data have been undertaken and published to date. A first catalogue (v0.1) covering 1,600 deg $^2$  was created in 2020, published in (Guinot et al. 2022) and used subsequently for cosmological constraints from peak counts Ayçoberry et al. (2023), and measuring group halo shapes Robison et al. (2023).

A second campaign (v1.0 and sub-flavours v1.1, v1.3) using 3,200 deg $^2$  and thus nearly doubling the previous area was run in 2022. The PSF model was switched to M CCD Liaudat et al. (2021). The changes compared to v0.1 are discussed in Li et al. (2024). Version v1.0 was released internally to the UNIONS collaboration, but not used in publications. Additional cuts were applied to create a more robust catalogue. The subsequent flavours of that catalogue are, first, v1.1, used for a direct measurement of intrinsic alignment Hervas Peters et al. (2025). This catalogue includes objects that SExtractor flagged as blends (FLAGS=2). Second, with more conservative cuts on object size v1.3 was defined, used in a number of publications: Li et al. (2024) measured the black hole to halo mass relation from weak lensing; Guerrini et al. (2025) developed and measured PSF systematics for cosmic shear; Zhang et al. (2024) quantified additive biases from PSF systematics for lensing-density cross-correlations. Mpetha et al. (2025) derived cosmological constraints from lensing cluster profiles, whereas Cheng et al. (2025) measured the lensing to probe the dark-matter environment of merger galaxies. Martin et al. (2025) measured void properties from lensing by underdense regions. In Ahad et al. (2025) the lensing profile of clusters was used to classify their dynamical state.

A re-processing campaign was run in 2023-2024 to recover lower leakage. A small number of previously failed tiles were added, bringing the area to 3,500 deg $^2$ . The PSF model was switched back to the well established PSFEX Bertin (2011). This catalogue (v1.4) is used in the present work and in the companion papers. The difference between PSFEX and M CCD performances are presented in Appendix B. To demonstrate the footprint we show the conver

### 2.2. SExtractor

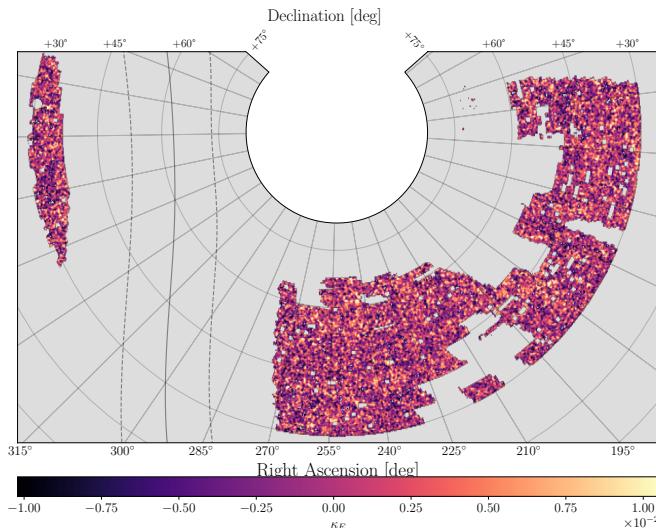
To identify sources, ShapePipe uses the widely adopted SExtractor library Bertin & Arnouts (1996). The detections are done at the tile level. The detected objects will all be subject to the shape measurement process. Due to cuts, only  $\sim 25\%$  objects will be part of the final catalog. The precise fraction of objects which are detected and which make it into the final catalogue can be inferred from Fig. 2. Our detection process is very conservative and exhaustive, which minimizes detection biases, i.e. any correlation between the shear of an object and the likelihood of it being detected. We rely on SExtractor for deblending and further masking. We then propagate the segmentation

Survey	Area [deg <sup>2</sup> ]	$n_{\text{eff}}$ [arcmin <sup>-2</sup> ]	$\sigma_e$	PSF FWHM [arcsec]
KiDS-Legacy	1,347	8.94	0.255	0.7 ( <i>r</i> )
DES-Y3	4,143	5.59	0.261	0.98 ( <i>r</i> ) – 0.85 ( <i>z</i> )
DES-Y6	4,422	8.22	0.29	0.95 ( <i>r</i> ) – 0.83 ( <i>z</i> )
HSC-Y3	433	22.9	0.236	0.59 ( <i>i</i> )
DECADE	5,412	4.59	0.254	
UNIONS-v0	1,700	6.8	0.35	0.65 ( <i>r</i> )
UNIONS-3500				0.65 ( <i>r</i> )

**Table 2.** ShapePipe catalogue versions and survey properties.  $A_{\text{obs}}$  is the total observed footprint;  $A_{\text{eff}}$  is the effective area after masking (v1.4 variants only).

Version	$A_{\text{obs}}$ (deg <sup>2</sup> )	$A_{\text{eff}}$ (deg <sup>2</sup> )	$n_{\text{gal}}$ (10 <sup>6</sup> )	$n_{\text{eff}}$ (arcmin <sup>-2</sup> )	$\sigma_e$	PSF	Notes	Reference
0.1	1,500	—	40	—	—	PSFex	P3 patch	Guinot et al. (2022); Ayçoberry et al. (2023); Robison et al. (2023)
1.1	3,200	—	128	—	—	MCCD	FLAGS=2	Hervas Peters et al. (2025)
1.3	3,200	—	84	—	—	MCCD	+ cuts	Li et al. (2023b); Guerrini et al. (2025); Zhang et al. (2024)
v1.4.5	3,200	2,894	85.0	6.48	0.28	PSFex	Initial	CosmoStat (2026) (this work); Hervas-Peters et al. (2026); Dally et al. (2026); Goh et al. (2026); Guerrini et al. (2026)
v1.4.6.3	3,200	2,894	61.4	4.96	0.27	PSFex	Size cut (fiducial)	
v1.4.8	3,200	2,517	53.9	4.98	0.27	PSFex	+ star-halo mask	
v1.4.11.3	3,200	2,894	78.2	6.26	0.27	PSFex	+ FLAGS ≤ 2	

**Notes.**  $n_{\text{eff}}$  follows Heymans et al. (2012);  $\sigma_e$  is the per-component ellipticity dispersion. Both use metacalibration-weighted galaxy samples. v1.4.11.3 relaxes the SExtractor flag criterion to  $\text{FLAGS} \leq 2$ , admitting objects with neighbour-biased photometry (1) or deblending (2). Versions v1.4.5, v1.4.6.3, and v1.4.11.3 share the same footprint (Section 2.3.1); v1.4.8 adds stellar-halo masking.



**Fig. 1.** Map of the convergence  $\kappa$  from the 3500 deg<sup>2</sup> of data presented in this release.

map of objects detected by SExtractor to remove neighbouring objects during the shape measurement process. The parameters for deblending are shown in Table 3; they have been lowered compared to Guinot et al. (2022).

**Table 1.** Recent weak-lensing surveys and their key properties. Information is taken from Wright et al. (2024, 2025) for KiDS legacy , Sevilla-Noarbe et al. (2021); Abbott et al. (2021); Yamamoto et al. (2025) for DES-Y3 & Y6, Li et al. (2022) for HSC, Anbjagane et al. (2025) for DECADE, and Guinot et al. (2022) for UNIONS v0.

Parameter	Value
THRESH_TYPE	RELATIVE
DETECT_THRESH	1.5
DETECT_MINAREA	5
FILTER	Y
FILTER_NAME	default.conv
DEBLEND_NTHRESH	32
DEBLEND_MINCONT	0.0005

**Table 3.** SExtractor parametrisation. All other parameters are kept to their default values, following ?.

## 2.3. Data processing

### 2.3.1. Masking

We apply different types of masking that defines the final shear catalogue. Mask information is created before, during, and after processing of ShapePipe. The first, pre-processing masks are generated during MegaCam processing (Marmo & Bertin 2008), and mask chip defects and cosmic rays.

The second masks are generated by ShapePipe. This includes flags generated by SExtractor and ngmix. These masks are defined by the following criteria. The SExtractor flags are

- $\text{FLAGS} = 0$ , removing blended objects;
- $\text{IMA_FLAGS} = 0$ , removing objects masked by the `mask_runner` module of ShapePipe. This includes bright stars, Messier and NGC objects, diffraction spikes, and reflection halos. The parameters of these masks are based on THELI pipeline; to mask bright stars we use the Guide

Star Catalogue (GSC, version 2.2; ?). We mask stellar halos (diffraction spikes) for stars brighter than magnitude 13 (18), where we use the mean magnitude over all bands available in the GSC.

The general ShapePipe selection criteria are

- $n_{\text{ep}} \geq 2$ , keeping objects observed during at least two epochs;

Next, we apply flags generated by `ngmix` shape measurement. These indicate a failed moment calculation to set the initial guess for the fit, and an invalid PSF model at galaxy position.

Additionally, we remove duplicate objects in the border regions of tiles. These duplicate objects are due to the small overlapping area between neighbouring tiles.

The third kind of masks are applied in post-processing. These are both pixel-based and area-based masks. The pixel-based masks are generated by the neural-network based software `maximask` (Paillassa et al. 2020), and contain masks for various contaminants such as cosmic rays, residual fringe pattern, diffraction spikes and saturated pixels among others.

After processing, we noticed that the ShapePipe masks were not large enough to mask some NGC objects entirely, resulting in many fake detections in the outskirts of those objects. We therefore applied additional pixel-based masks from the THELI pipeline to conservatively mask bright stars, Messier, and NGC objects. We also identify regions around stars, sorted in two magnitude bins, that show a faint halo-like emission. These regions can however safely be used for most purposes: Shapes of galaxies falling in these regions can be well measured, and we do not apply this very conservative mask for our shear catalogue, following Erben et al. (2013). **Fabian:is this final?**

The area-based masks indicate the core region of each tile, thus removing overlapping regions between neighbours. An additional criterium concerns the overall exposure coverage. A rough exposure count map is created and only areas with  $n_{\text{pointing}} \geq 3$  are kept. This selection is complementary to the above-defined  $n_{\text{ep}} \geq 2$  criterium. It is useful to homogenize the coverage map. Finally, a tile-based flag indicates whether the tile has been observed and processed in the  $r$  band. This flag is currently redundant, but will be used with corresponding flags for the  $u$ ,  $i$ , and  $z$  band for upcoming multi-band shear tomography catalogues.

The pixel-based masks are transformed into `healsparse` area maps, and then the `healsparse` pixel values are attributed to each catalogue object.

**Galaxy-property cuts** (size, magnitude, shape measurement quality) do not enter the footprint definition; only spatially defined criteria do (tile coverage, image quality flags, and star and artifact masks), reducing the full 3,500 deg<sup>2</sup> imaging area to an effective footprint. We compute this footprint by counting HEALPix pixels at  $N_{\text{side}} = 4096$  occupied by at least one source in the detection catalog before galaxy-property cuts, giving  $A_{\text{eff}} = N_{\text{pix}} \times 0.74$  arcmin<sup>2</sup>. The standard footprint covers 2,894 deg<sup>2</sup> (versions v1.4.5, v1.4.6.3, v1.4.11.3); the stellar-halo-masked variant (v1.4.8) covers 2,517 deg<sup>2</sup>.

### 2.3.2. Galaxy selection

The following criteria define the galaxy sample of the shear catalogue.

- $15 \leq r \leq 30$ , where the  $r$ -band magnitude is measured via `SExtractor`;

- $10 \leq v \leq 500$ , where  $v$  is the signal-to-noise ratio, defined as the ratio of flux and flux error as measured during the `ngmix` modelling.
- $0.707 \leq r_h/r_{h,\text{psf}} \leq 3$ , with  $r_h$  being the galaxy half-light radius of the original (deconvolved) galaxy image, and  $r_{h,\text{psf}}$  the PSF half-light radius;

The latter two selections are applied during `Metacalibration`, and thus determine the selection response matrix. The cut on the relative size of small objects is identical to the one used by DES Gatti et al. (2021); our numerical value is larger by a factor of  $\sqrt{2}$  because we use the half-light radius  $r_h$ , whereas DES defines size as the area  $T \propto r_h^2$ .

### 2.3.3. Weights

A first estimate of galaxy weights are inverse-variance weights, which depend on the properties of individual galaxies. The weight for the  $i^{\text{th}}$  galaxy is given by

$$w_i = \frac{1}{\sigma_e^2 + \sigma_{i,1}^2 + \sigma_{i,2}^2}, \quad (1)$$

where  $\sigma_e = 0.34$  is the estimated sample ellipticity dispersion, and  $\sigma_{i,j}$  is the shape measurement error associated with component  $j = 1, 2$ . The total variance is thus the sum of intrinsic and measurement variance. The latter is the trace of the measurement error covariance that `ngmix` computes.

A shear estimator as the mean over the shapes of galaxies (or any linear combination thereof) has a minimum variance by using the weights given by Eq. (1).

We further compute weights that are less sensitive to noise in the estimated measurement errors. Following Gatti et al. (2021), we compute the weights as piecewise constant values in bins of relative size and signal-to-noise ratio. These two quantities are good predictors of galaxy weight.

$$\begin{aligned} w(r_{\text{gal}}/r_{\text{PSF}}, v) &= \sigma_{\gamma}^{-2}(r_{\text{gal}}/r_{\text{PSF}}, v) \\ &= [\sigma_e \langle R_{\gamma} \rangle]^{-2}(r_{\text{gal}}/r_{\text{PSF}}, v). \end{aligned} \quad (2)$$

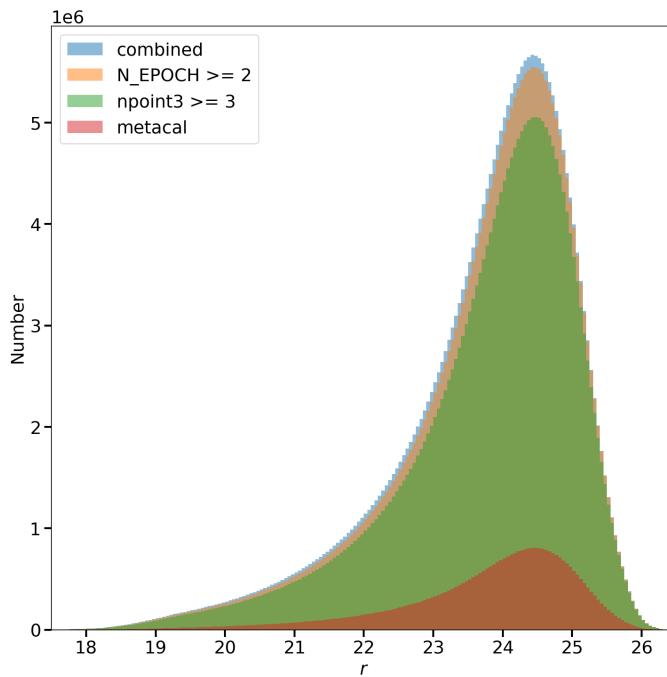
The shear dispersion  $\sigma_{\gamma}$  is rewritten as the product of the ellipticity dispersion  $\sigma_e$  and the shear response. The motivation to introduce this weight was to decouple weight from ellipticity, since the shape measurement error in Eq. 1 can correlate with shear. The distribution of these weights as a function of SNR and size-ratio is shown in Fig. 3. **Most galaxies lie in the small-size, low-SNR regime, where weights are small. The weight distribution shows a complex pattern.** In particular the top left corner could be affected by blended galaxies, i.e. large diffuse objects, which seem to be strongly down-weighted.

### 2.3.4. v1.3

Tile-by-tile processing. **Fabian: How far do we want to bring this comparison and how much do we want to talk about v1.3? Messaging is also important**

### 2.3.5. v1.4

Of the 14,596 tiles, 9 (0.06%) failed to finish processing, for the following reasons. 2 tiles had zero weight everywhere, due to some quirk in the CFHT pipeline. Three tiles had exceptionally large number ( $\sim 10$  times the mean) of `SExtractor` detections



**Fig. 2.** FHP: This plot feels to deep to me, metacal galaxies don't go down to 26, but are lost by 25 I believe. Histogram of the  $r$ -band magnitude of detected objects with various additive selection criteria: all unmasked objects with valid shapes (blue); with  $n_{ep} \geq 2$  (orange); with  $n_{pointing} \geq 3$  (green); with size and SNR cuts performed during metacalibration (brown).

due to its vicinity to a nearby Messier galaxy. Since all those detections fall into the masked regions, we removed those tiles from further processing. Finally, processing of 4 tiles failed after multiple re-tries for unknown reasons.

### 3. Shapes, shears, and calibration

#### 3.1. Number density

We compute the number density using the Heymans 2012 definition

$$N_{\text{eff}} = \frac{1}{A} \frac{(\sum_i w_i)^2}{\sum_i w_i^2} \quad (3)$$

where  $w_i$  is the shear weight per galaxy and  $A$  the area of the survey. In Fig. 4 the distribution of the effective number density is shown across the survey footprint. The mask has been subtracted from the area in each pixel. This is key since the raw number density correlates strongly with right ascension due to the stellar density increasing towards the galactic plane as shown in Fig. A.1. The spread of  $N_{\text{eff}}$  seems to be well constrained. This can be attributed to the homogeneous data in the  $r$ -band shown in Gwyn et al. (2025). A small decrease towards the galactic plane can hypothetically be attributed to dust extinction.

#### 3.2. Additive bias

Table 4 shows the mean and R.M.S. [MK: todo] of the additive bias  $c_1$  and  $c_2$ . We remove these values from each ellipticity component to empirically account for the additive bias. These values are relatively small and well aligned with other

stage-III surveys. In DES Y3 (DES Y6) the reported values are  $c_1 = 3.5 \times 10^{-4}$  ( $1.9 \times 10^{-4}$ ) and  $c_2 = 0.6 \times 10^{-4}$  ( $-0.3 \times 10^{-4}$ ). This places our average shear at the same magnitudes. During our work on PSF contamination mitigation subtraction, no indication appeared that this global calibration was in conflict with the leakage correction.

Note that the *Euclid* requirements Cropper et al. (2013) are not obtained through the raw values, but ones removed the desirable requirement is achieved through empirical calibration. **Fabian:** writing this I realised that the Euclid requirements are on the errors on  $c$  at least in the Cropper and Kitching so it would be nice to have them for comparison. They used to be computed automatically i believe?

**Table 4.** Additive bias components  $c_1$  and  $c_2$ , for different weightings.

version	$c_1$	$c_2$
DES weights	$-1.566 \times 10^{-4}$	$2.259 \times 10^{-4}$
inverse-variance	$-1.363 \times 10^{-4}$	$2.325 \times 10^{-4}$
No weights	$-1.038 \times 10^{-4}$	$2.302 \times 10^{-4}$
version	$c_1$	$c_2$
v1.4.5	$-1.094 \times 10^{-4}$	$2.117 \times 10^{-4}$
v1.4.6	$-1.566 \times 10^{-4}$	$2.259 \times 10^{-4}$
v1.4.7	$-1.332 \times 10^{-4}$	$2.213 \times 10^{-4}$
v1.4.8	$-1.341 \times 10^{-4}$	$2.215 \times 10^{-4}$

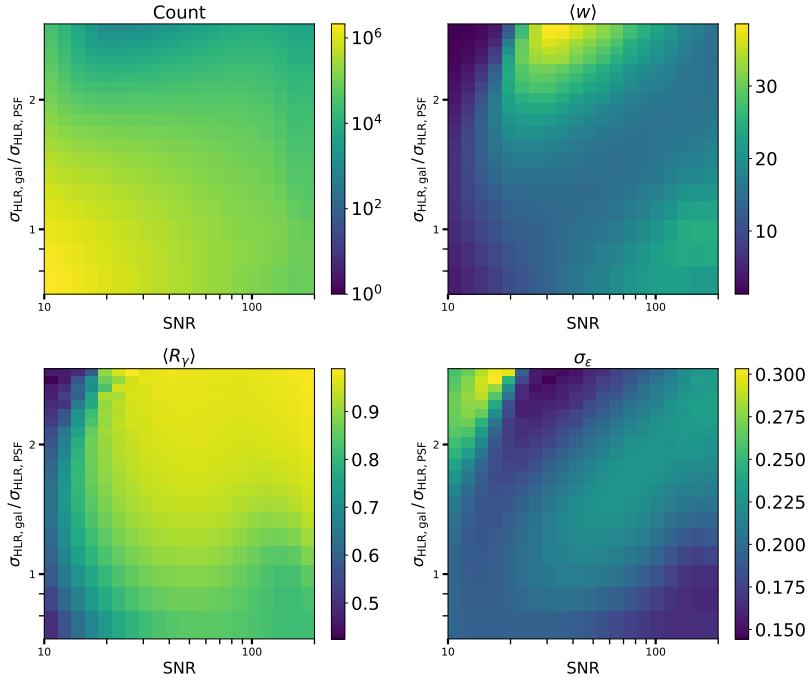
#### 3.2.1. Metacalibration shear and selection response

As mentioned in the introduction, multiple shape measurements methods have been explored by the weak lensing community in the past decade. Some rely on estimating moments of the surface brightness profiles, others try fitting a parametric profile to the galaxies, such as a mixture of Gaussians or a bulge+disk model using Sérsic profiles. While each of this methods has their own advantages and biases, a method that has been widely employed in recent years to estimate model, noise and selection biases is Metacalibration. This technique quantifies capacity of a given shape measurement method to recover a known shear  $\Delta\gamma$  that is applied to the data directly (Huff & Mandelbaum 2017; Sheldon & Huff 2017). This is done by first deconvolving the image from the PSF model, shearing the deconvolved image and then reconvolving it by a slightly dilated PSF to suppress high frequency modes. During the shearing process, correlated noise can appear, causing in the recovered shear. To cancel out the effect of correlated noise, artificial noise is introduced, referred to as *fixnoise*. This artificial noise causes an approximate 20% loss in signal-to-noise ratio. Some techniques leveraging deep fields Zhang et al. (2022) have been investigated to reduce this loss by relying on the higher signal to noise of galaxies in these regions. From a theoretical point of view the shear response is obtained by Taylor expanding the ellipticity around a small variation in shear:

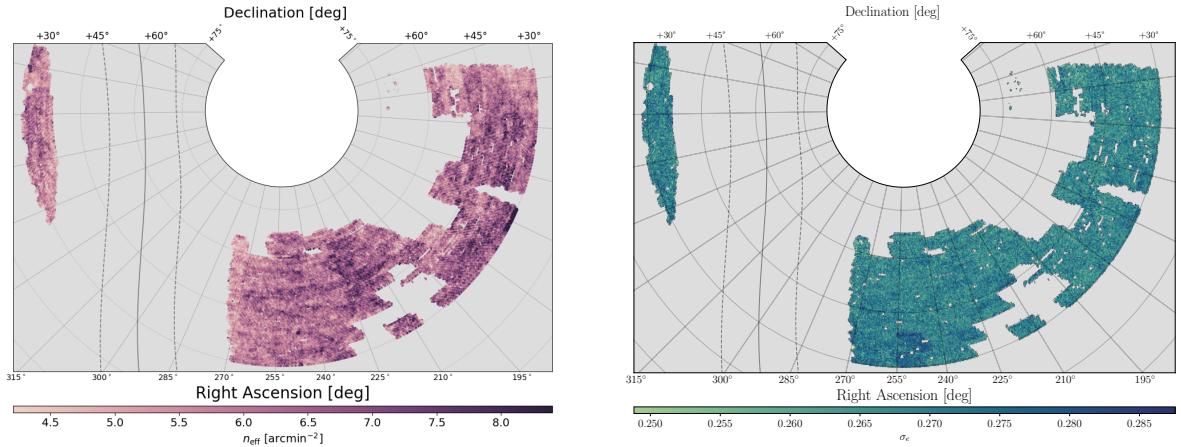
$$e = e|_{\gamma=0} + \left. \frac{\partial e}{\partial \gamma} \right|_{\gamma=0} \gamma + \dots \quad (4)$$

This leads to define at first order the shear response matrix  $R$  as

$$R = \left. \frac{\partial e}{\partial \gamma} \right|_{\gamma=0} \quad (5)$$



**Fig. 3.** Raw averages of galaxies in bins of ratios of galaxy to PSF size (y-axis) and galaxy signal-to-noise ratios (x-axis). The binned quantities, indicated by the color bars, are number counts (*upper left*), weights (*upper right*), shear response matrix diagonal (*lower left*), and per-component shape noise (*lower right*).



**Fig. 4.** Footprint over the Northern Galactic sky in  $N_{\text{eff}}$  and  $\sigma_\epsilon$ .

The response matrix is commonly split into a shear response  $R_\gamma$  and selection response  $R_s$ . The shear response can be estimated per galaxy as

$$R_\gamma \equiv R_{ij} = \frac{e_i^+ - e_j^-}{2\Delta\gamma}, \quad (6)$$

where the  $i, j$  indices refer to the ellipticity components. The selection responses can only be defined for a global population, since the goal is to quantify how cuts affect the mean ellipticity

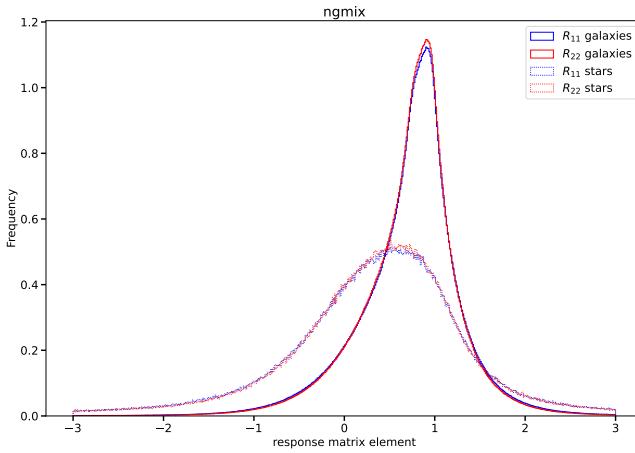
$$R_{\text{sel}} = \frac{\langle e_i^{\text{NS}} \rangle^{S_+} - \langle e_i^{\text{NS}} \rangle^{S_-}}{2\Delta\gamma}. \quad (7)$$

The superscript  $S_\pm$  refers to the artificially sheared galaxy branch the selection cuts are applied to. The NS indices indicate the "No Shear" branch. Calibrated ellipticities are obtained

by averaging over shear responses, since individual realisations can be very noisy. We only use the off-diagonal elements and hence  $e_{\text{cal},i} = e_{\text{meas},i}/\langle R_{ii} \rangle$ . Weights need to be applied consistently in the numerator and denominator when taking these averages. While splitting these two responses gives control and interpretability, Eq. 7 can be noisy. With restricted statistics, inferring both jointly is more robust:

$$R_{\text{tot}} = \frac{\langle e_i^+ \rangle^{S_+} - \langle e_i^- \rangle^{S_-}}{2\Delta\gamma}, \quad (8)$$

This method has been applied in Anbjagane et al. (2025) and Paper V, where the error on the selection response estimated separately was of a similar order as the total error on the multiplicative bias estimate. The selection response in our case is  $R_{s,1} =$  and  $R_{s,2} =$ . While the total shear response is typically below 1,



**Fig. 5.** Shear response of stars and galaxies. See text for a detailed discussion of the non-zero shear response in stars.

the selection response can be positive or negative, meaning selection effects can dilute or boost the total shear. Zuntz et al. (2018) show that cutting on SNR creates a positive selection response while cutting on size causes a negative one.

In practice ShapePipe uses ngmix to measure shapes with Gaussian mixtures. The details of the pipeline are laid out in Guinot et al. (2022). In brief, a guess is made based on a fast moment computation using the HSM algorithm ?. The minimiser uses regularization priors and 6 parameters are fitted to each galaxy:  $[\Delta x, \Delta y, e_1, e_2, r_{50}, F]$  where  $\Delta x, \Delta y$  represent a global offset accounting for some distance between the SExtractor detection centre and the best fit one. The ellipticity components are  $e_1, e_2, r_{50}$  is the half-light radius and  $F$  is the flux. All epochs are then fitted at once, each with their specific PSF. An issue was discovered during the pipeline validation process, which was the lack of coherent offsetting between the different epochs. This issue is described in detail in paper II and its consequences are studied using image simulations.

The shear response distributions in Fig. 5 show that the galaxy shear responses are very consistent for  $R_{11}$  and  $R_{22}$ . The distribution peaks around 1 but has a larger tail towards the left hand side, leading to the mean values  $R_{11} = R_{22} =$ . The origin of the slight discrepancy between  $R_{11}$  and  $R_{22}$  is unclear, but one hypothesis is that our  $51 \times 51$  pixel postage stamp can cause some slight asymmetry between the horizontal and vertical axis propagating into  $e_1$  and  $e_2$  respectively. The large non-zero shear response in our star sample is directly explainable by the offset problem mentioned in the previous paragraph. Stars are point-like objects and their shear response should average to 0. Paper II describes how this issue was both reproduced and corrected in simulated images by fixing the lack of coherent offsetting. This improvement will be applied in upcoming improved UNIONS shape measurements processing.

### 3.2.2. Residual multiplicative bias from image simulations

In Paper II, an extensive set of simulated images are built to validate the shape measurement pipeline and to quantify residual multiplicative biases tied to blends. These simulations use a realistic catalogue of galaxies based on N-body simulations in Li et al. (2023a) and galaxy properties are assigned via a semi-analytic model Lagos et al. (2018). The observation strategy of UNIONS in the  $r$ -band is reproduced with high fidelity and PSFs are directly drawn from the model estimated on the data. Galax-

ies are then placed in two different configurations. First, they are placed on a grid, since from the literature, no bias source exceeding  $m = 0.005$  is expected when using Metacalibration on isolated galaxies. This serves as a dual validation of our shape measurement pipeline and of the image simulations. While it is possible that two errors compensate, the fact that we interface a validated image simulation suite developed by the Kilo Degree Survey collaboration with a well-tested shape measurement code mainly used by the Dark Energy Survey gives us confidence that cancelling-out errors are unlikely, in particular at the relatively coarse precision of  $\sigma_m \sim (10^{-2})$  we currently try to achieve. This step revealed two shortcomings of the pipeline which require corrective multiplicative bias calibrations. The total shear bias when galaxies are placed on a grid is found to be, indicating a trustworthy implementation of the Metacalibration method. Due to the size of these error bars, an uncertainty on the multiplicative bias of  $\sigma_m = 0.01(?)$  is included in the cosmological analyses of Papers IV & V.

### 3.2.3. PSF leakage calibration

In addition to the weighted additive and multiplicative biases, residual PSF leakage remains in the measured ellipticity. This will be quantified in Sect. 5 using various validation techniques to assess the amplitude of the PSF leakage. Here, we review the method to correct PSF leakage at the object level following Li et al. (2023a). We define the measured ellipticity of each object  $i$  as follows,

$$e_i^{\text{obs}} = \epsilon_i^s + \gamma_i + \alpha e_i^{\text{PSF}}, \quad (9)$$

where  $\epsilon_i^s$  corresponds to the stochastic component of the ellipticity,  $\gamma_i$  is the shear component,  $e_i^{\text{PSF}}$  is the interpolated PSF ellipticity at the position of object  $i$  and  $\alpha$  is the associated leakage coefficient. All three terms are complex spin-2 complex quantities. We model the leakage  $\alpha_i$  as a single scalar for both ellipticity components, which corresponds to the PSF leakage contribution as a simple rescaling of the PSF ellipticity.

The correction of leakage consists of two consecutive steps to estimate  $\alpha_i$  for each object. A first step consists of binning the galaxies by and signal-to-noise ratio,  $\nu_{\text{SNR}}$ , and resolution,

$$\mathcal{R} = \frac{r^{\text{PSF}}}{r^{\text{PSF}} + r^{\text{obs}}}, \quad (10)$$

where  $r^{\text{PSF}}(r^{\text{obs}})$  is the PSF (galaxy) half-light radius. We build a  $20 \times 20$  grid in  $(\nu_{\text{SNR}}, \mathcal{R})$  and compute the PSF leakage coefficient  $\alpha$  using a linear regression in each bin. Fig. 6 shows the binned  $\alpha$  parameter. Faint and small galaxies tend to have a negative leakage contribution.

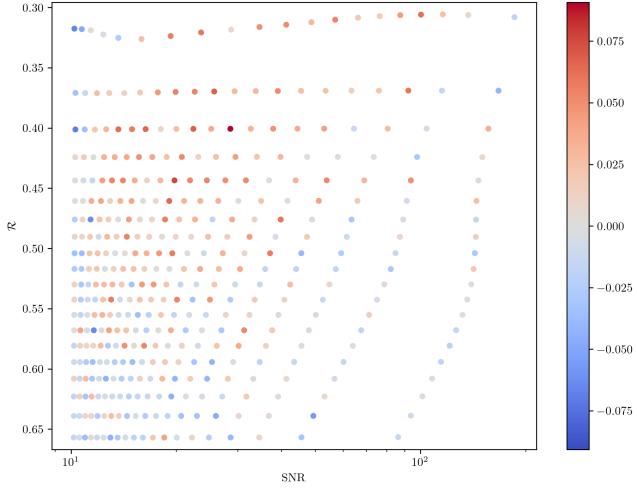
Next, we estimate the global leakage trend in the  $(\nu_{\text{SNR}}, \mathcal{R})$  plane by fitting the regressed and binned  $\alpha$  with the fitting function

$$\alpha_{\text{trend}}(\nu_{\text{SNR}}, \mathcal{R}) = a_0 + a_1 \nu_{\text{SNR}}^{-2} + a_2 \nu_{\text{SNR}}^{-3} + b_1 \mathcal{R} + c_1 \mathcal{R} \nu_{\text{SNR}}^{-2}. \quad (11)$$

We then correct the observed ellipticities by subtracting the fitted leakage,

$$e_i^{\text{obs,tmp}} = e_i^{\text{obs}} - \alpha_{\text{trend}}(\nu_{\text{SNR},i}, \mathcal{R}_i) e_i^{\text{PSF}}. \quad (12)$$

After that first correction, some residual leakage might remain. To remove this residual leakage further, we perform another suite of linear regressions to estimate, per bin, the residual leakage  $\alpha_{\text{res}}$ . We then remove for each object its residual leakage based



**Fig. 6.** PSF leakage in the measured galaxy ellipticities as a function of SNR and resolution  $\mathcal{R}$  Eq. (10).

on the bin it is assigned to, similar to Eq. (12). The overall leakage removed per object writes

$$\alpha_i = \alpha_{\text{trend}}(\nu_{\text{SNR},i}, \mathcal{R}_i) + \alpha_{\text{res},i}, \quad (13)$$

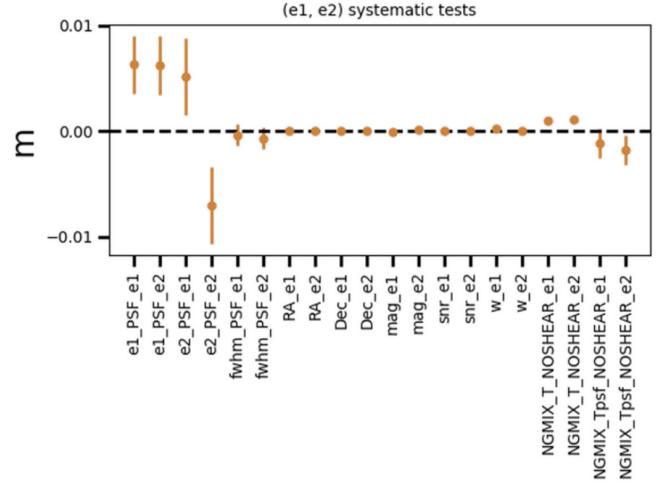
where  $\alpha_{\text{res},i}$  is the value of the residual leakage in the  $(\mathcal{R}, \nu_{\text{SNR}})$  bin of object  $i$ . By construction, the measured leakage per bin is close to zero after subtraction. Sect. 5 will show the impact of this correction on the leakage using different estimators.

### 3.2.4. Size, flux, and other biases

To test the robustness of the shear estimate we perform a simple linear regression as a function of observable parameters. In Fig. 7 the multiplicative factor  $b$  is represented  $\gamma_{\text{obs}} = b \langle \epsilon_{\text{obs}} \rangle + c$ . While  $b$  can be interpreted as a multiplicative bias with respect to a given observable quantity, we use the notation  $b$  to distinguish it from the global calibrated bias determined from image simulations described in 3.2.2. For most quantities — magnitude, SNR, and RA/DEC — the ellipticity does not correlate. The weights  $w$  also do not correlate with ellipticities, a positive consequence of the use of gridded weights from Eq. 2. The input PSF size, whether measured with moments (fwhm\_PSF) or as the Gaussianized PSF size from ngmix used for the deconvolution, lightly anti-correlates with the ellipticity components at the  $1\sigma$  level. Galaxy size estimated from the Gaussian profile during shape measurement, by contrast, correlates with ellipticity at the  $XX\sigma$  level with a small but non-negligible value. Unfortunately the image simulations do not have the statistical power to inform on this trend. Gatti et al. (2021) found similar trends in DES-Y3, albeit with a different size dependence: large galaxies appear more affected in UNIONS, while smaller galaxies show boosted shear in DES-Y3. The correlation for PSF ellipticities are particularly strong. We describe different ways to recover and correct for this PSF-leakage in 5.2. The values are respectively XX YY, indicating strong deviations from 0.

### 3.3. Tile and CCD diagnostics

To gain confidence in the absence of systematics a common diagnostic is to look at the ellipticity distributions as a function of the core processing units. In our case the core levels are the



**Fig. 7.** Bias when fitting a linear relation  $e_i(X) = b e_i(X) + d$  where  $X$  represents FHP: plot needs better label (also we could reserve the name  $m$  and  $c$  bias to the calibratable bias and call these something else? (Gatti chose  $b$  for  $m$ )

focal plane, where the image is taken, and the tile, when the exposures are combined for detection. We start by looking at the average ellipticity binned in focal plane position and tile position in Figs. 8&9. In the case of the focal plane, each galaxy contributes to multiple pixels, appearing at various positions across different exposures. Qualitatively no pattern seems to emerge in either of the plots. When looking at the structure of the PSF in B.1 and the level of leakage, one could have imagined seeing some trends related to the PSF field. The scatter on each plot also appears under control. Having found no global trends visually, we estimate the reduced  $\chi^2$  for each histogram as a quantitative check. Assuming pure Poisson noise and testing consistency with zero, the reduced  $\chi^2$  values are

- CCD level:  $\chi^2(e_1)/\text{d.o.f.}=20525.7/18000=1.140$
- $\chi^2(e_2)/\text{d.o.f.}=20175.1/18000=1.121$
- Tile level:  $\chi^2(e_1)/\text{d.o.f.}=2565.6/2500=1.026$
- $\chi^2(e_2)/\text{d.o.f.}=2450.4/2500=0.980$ .

At the tile level the reduced  $\chi^2$  values are close to 1, while at the CCD level they differ slightly. We attribute this to the PSF-leakage, as the PSF displays a coherent pattern across the focal plane, as shown in Fig. B.1. Therefore this small excess should be captured by the PSF-leakage scheme detailed in Sect. 5.2

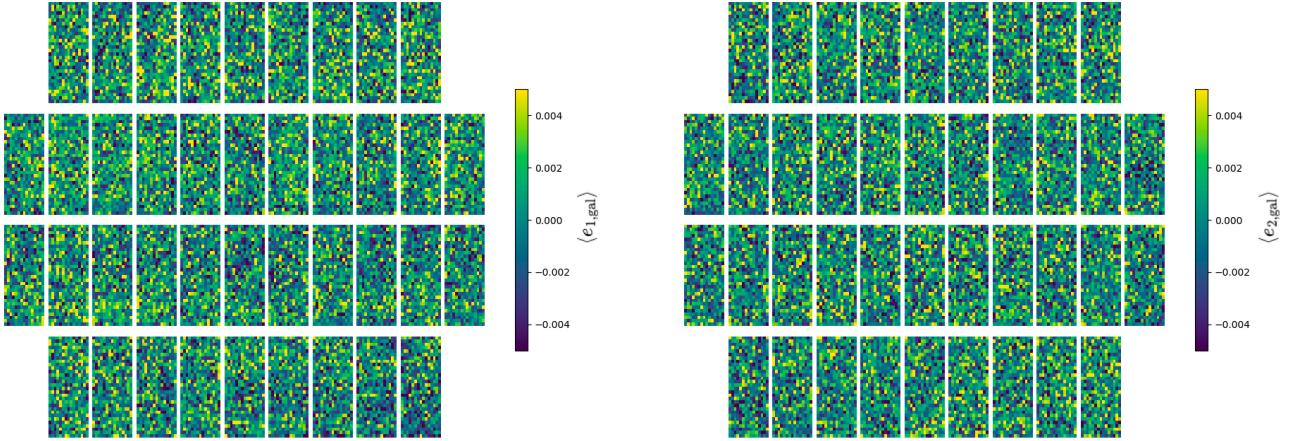
## 4. Position-shape correlation function null tests

### 4.1. Tangential shear around processing units

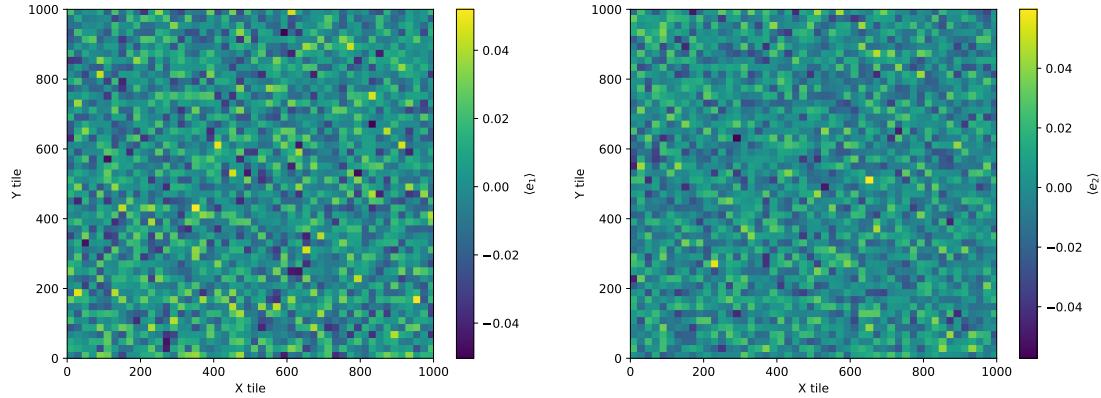
A common additional null-test is to investigate the galaxy-galaxy lensing signal around processing-unit centres. Figure 10 shows  $\gamma_t$  and  $\gamma_x$  around tile and CCD centres. The reduced  $\chi^2$  values are consistent with unity, and no spatial pattern is visible, indicating an absence of additive systematics at the tile and CCD level. The error bars are shot-noise only for the tile-centred correlation functions; the CCD-centred ones use a jackknife process to account for survey-wide variability.

### 4.2. Tangential shear around stars

Stars form a very bright class of objects present across the whole sky. While they are necessary to calibrate the PSF-model, their



**Fig. 8.** Average calibrated ellipticity components ( $e_1, e_2$ ) as a function of position on the focal plane. Due to the complex dither pattern, each galaxy appears at multiple positions on the focal plane. The spread is small and no pattern emerges. **FHP: need for quantitative chi2?**



**Fig. 9.** Mean ellipticity components ( $e_1, e_2$ ) as a function of position on the tile. No complex pattern emerges indicating an absence of additive systematic at the tile level.

intense **brightness** can result in halos around them. While the effect of inclusion or exclusion of galaxies from these halo-regions can be tested with masking, a good verification is to compute the galaxy-galaxy lensing signal around stars. This **allows us to test whether** any unsubtracted light from the stars coherently affected galaxies during the shape measurement process. In Figs. 12 & 11 we compute the tangential and cross signal around GAIA selected stars and around our own PSF selected stars. What

We measure the tangential and **cross-component** shear around stars, as null tests. First, we use the UNIONS PSF validation stars as position sample. The shear around this sample is shown in Figure 11.

Next, we consider GAIA star samples, split up into three magnitude bins. The result is shown in Fig. 12.

#### 4.3. Stars around galaxies

In Zhang et al. (2024), a framework is developed to capture the additive bias  $f$ . This can not be understood as a null test since its result will most likely deviate from 0. Indeed

## 5. PSF estimation and leakage quantification

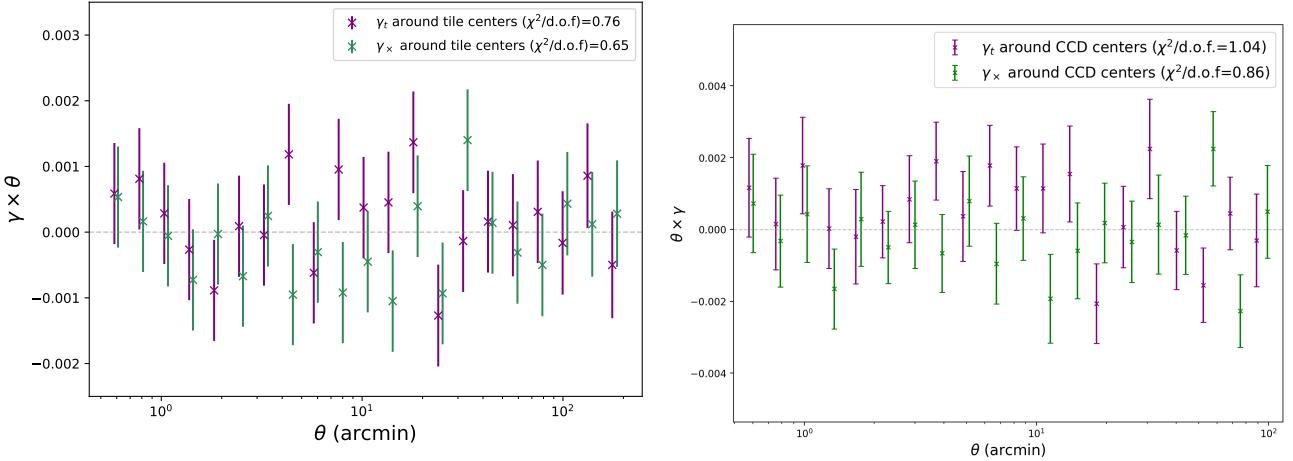
We validate the PSF model, assess PSF leakage, quantify the impact of the object-wise leakage correction (Section 3.2.3), and estimate the amplitude of the PSF systematics additive bias on the two-point correlation function.

### 5.1. PSF stars

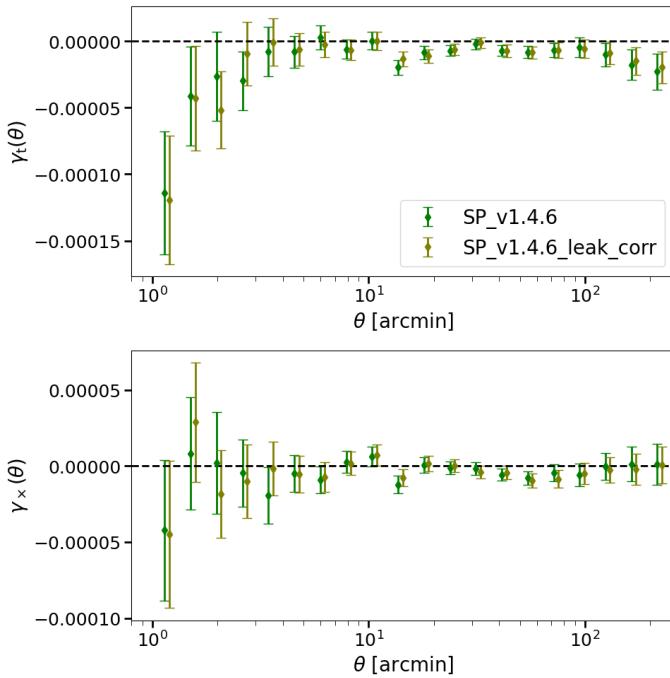
#### 5.1.1. Star sample selection and properties

To identify stars we rely on the size-magnitude space, using the mode of the size distribution with a  $\pm 0.2$  pixel tolerance. Objects with an apparent magnitude  $m_r \in [18, 22]$  are kept in the star sample. The lower cut excludes objects which can be affected by the brighter-fatter effect, while the upper bound is meant to exclude slightly **extended, fainter** galaxies. Due to the noisiness of the shape measurement process, some stars make it into the final galaxy sample, biasing the shear measurements. The induced shear bias is quantified in Paper II. The stars are split into two samples, the training sample on which the PSF model is built contains 80% while the validation makes up the remaining 20% which are used for the diagnostic work below.

Figure 14 shows the distribution of PSF model residuals at reserved star positions. The ellipticity residuals are centred on



**Fig. 10.** Galaxy-galaxy lensing signal around tile centers (left) and CCD centers (right). All reduced  $\chi^2$  indicate compatibility with the null hypothesis, indicating that the data-set passes this null test.

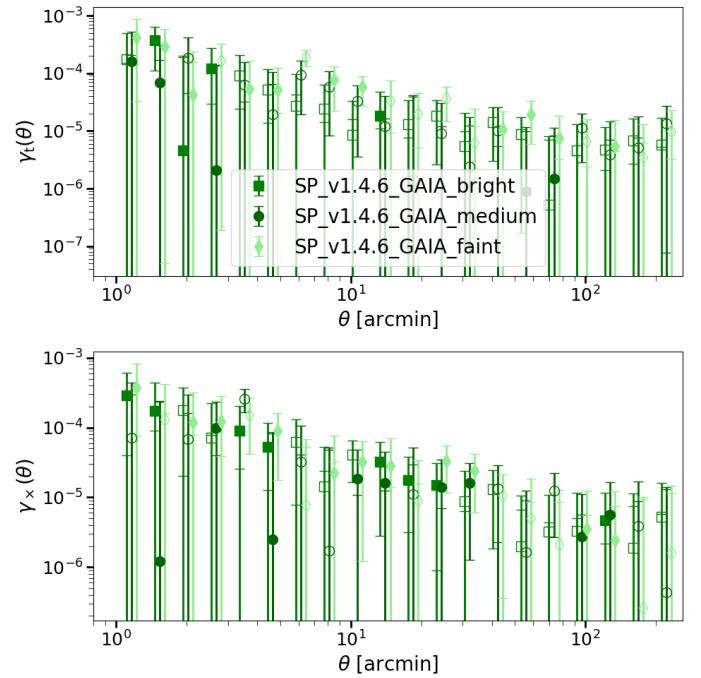


**Fig. 11.** Tangential (top panel) and cross-component (bottom) shear of galaxies around UNIONS PSF validation stars.

zero with a small spread. The size residuals, while also peaking at zero, are skewed towards negative values, corresponding to an overestimation of the PSF size. Zhang et al. (2023) introduces those moments to quantify PSF model errors beyond ellipticity and size. It is computed from higher-order moments computed on the PSF and star postage stamps following

$$M_{pq} = \frac{\int dx dy u^p v^q \omega(x, y) I(x, y)}{\int dx dy \omega(x, y) I(x, y)}, \quad (14)$$

where  $(x, y)$  are pixel coordinates, and  $(u, v)$  are standardised coordinates, defined in a coordinate system where the second-moment covariance is the identity matrix. This transformation de-correlates higher-order from second-order moments. See Zhang et al. (2023) for details. Additionally,  $\omega(x, y)$  is a Gaussian weight function matched to the PSF size, and  $I(x, y)$  is the



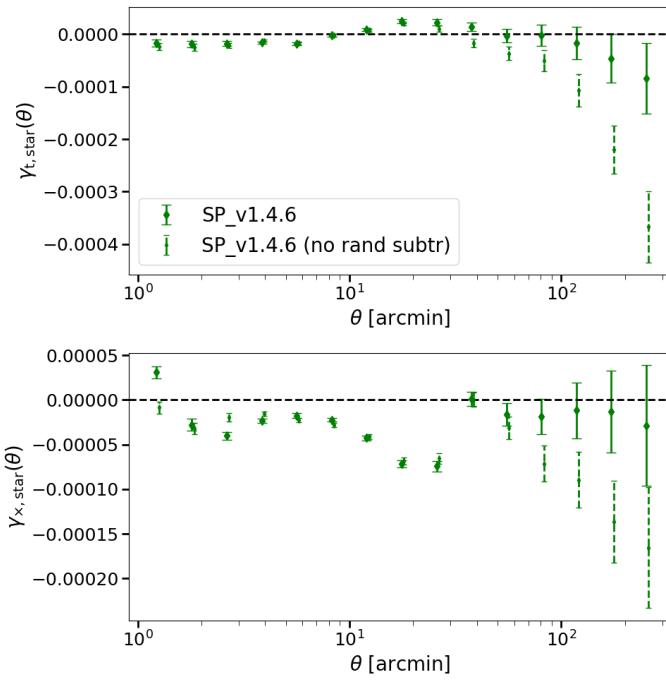
**Fig. 12.** Tangential (top panel) and cross-component (bottom) shear of galaxies around GAIA stars.

light profile in the image coordinate system. A spin-2 quantity can be built from fourth-order moments as

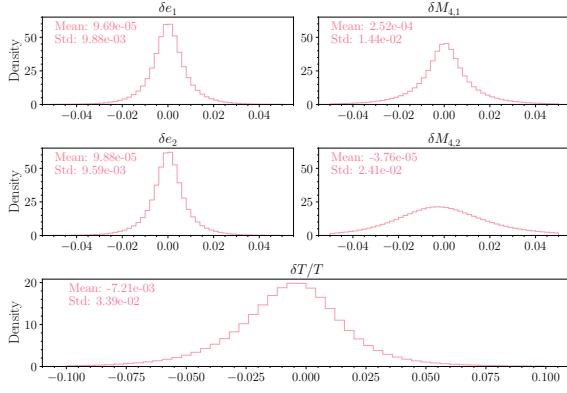
$$M^{(4)} = (M_{40} - M_{04}) + 2i(M_{13} + M_{31}). \quad (15)$$

The distributions of both fourth-order moment components are centred on zero with a small spread, with the second component showing a larger variance than the first one. Overall, the PSF model appears to perform well on reserved stars.

As explained in 5.1.1, stars are selected in a precise magnitude range  $r \in [18, 22]$ . One common issue for weak lensing surveys when choosing calibration stars is the brighter-fatter effect: as more charges are trapped in a given pixel, the pixel gets electrically charged and electrons are repelled to neighboring pixels, effectively changing the potential lines and making the central pixels smaller. On Megacam this effect is fairly small Astier et al. (2013); Guyonnet et al. (2015) since the thinned CCD's



**Fig. 13.** Tangential (top panel) and cross-component (bottom) shear of UNIONS PSF validation stars around galaxies.



**Fig. 14.** Residuals of the PSF evaluated at reserved star positions, shown as histograms. Top left panels: First and second ellipticity components. Top right panels: First and second fourth-moment components, defined in Eq. (15). Bottom panel: Normalised size  $\delta\tilde{T}^{\text{PSF}} = (T^* - T^{\text{PSF}})/T^{\text{PSF}}$ . [MK: I guess the size is hlr? Change in plot.]

are back-illuminated. The magnitude–radius plot in Guinot et al. (2022) shows a very constant stellar locus down to  $r \sim 16.5$ .

## 5.2. PSF contamination of galaxy shapes

As in Sect. 3.2.3, we model PSF contamination to galaxy shear estimates as additive, linear biases. In addition to the leakage  $\alpha$  term introduced in Eq. (9), we quantify contributions related to PSF mismodeling with residuals in ellipticity,  $\delta e^{\text{PSF}} = e^* - e^{\text{PSF}}$ , and in normalised size,  $\delta\tilde{T}^{\text{PSF}} = (T^* - T^{\text{PSF}})/T^{\text{PSF}}$ . Here, quantities with the superscript ‘\*’ denote measurements on stars, while ‘PSF’ refers to PSF model predictions at the position of the star. The star and PSF quantities must be measured with the same methods; otherwise, the residuals mix PSF mismodelling with

potential ellipticity and size mismatches between shape measurement techniques. We use adaptive moments with the galsim HSM implementation for both.

Our linear model for the PSF systematic ellipticity is

$$e^{\text{PSF, sys}} = \alpha e^{\text{PSF}} + \beta \delta e^{\text{PSF}} + \eta \delta T^{\text{PSF}}, \quad (16)$$

such that

$$e^{\text{obs}} = \epsilon^s + \gamma + e^{\text{PSF, sys}}. \quad (17)$$

The parameters  $\alpha$ ,  $\beta$  and  $\eta$  modulate the three PSF systematics terms. Following (Paulin-Henriksson et al. 2008), the PSF size is redefined as  $\delta T^{\text{PSF}} = e^{\text{PSF}} \delta \tilde{T}^{\text{PSF}}$ . The ellipticity factor makes  $\delta T^{\text{PSF}}$  a spin-2 field, and Eq. (17) is thus well-defined.

### 5.2.1. Galaxy-PSF correlation functions

Galaxy-PSF cross-correlation functions are powerful tools for validating the PSF model and the shape measurement process. These cross-correlations vanish in an ideal case; in practice, they are routinely used to estimate the amplitude of the additive bias in the shear correlation function due to PSF systematics. With the error model from Eq. (16) they can be used to distinguish between leakage and mismodelling.

A first step is to compute PSF and PSF residual correlations at star positions, known as  $\rho$ -statistics (Rowe 2010; Jarvis et al. 2016; Gatti et al. 2021). They allow for the comparison of the performance of different PSF models. Here we use them to estimate the amplitude of the leakage bias. Figure 15 shows the different  $\rho_i$  with  $i \in [0, \dots, 5]$  corresponding to the different cross-correlations in Eq. (16). The shaded region corresponds to  $0.5\rho_+^{\text{sys}}$ . We estimate the amplitude from  $\xi_+(\vartheta)$  and its signal-to-noise ratio following Mandelbaum et al. (2018). The  $\rho$ -statistics associated to leakage tend to exceed those requirements on large scales, and  $\rho_2$  shows an excess on small scales. The excess is modelled in Paper IV’s configuration-space cosmological analysis. The different  $\rho$  correlation functions are:

$$\rho_0(\vartheta) = \langle e^{\text{PSF}} e^{\text{PSF}} \rangle(\vartheta); \quad \rho_1(\vartheta) = \langle \delta e^{\text{PSF}} \delta e^{\text{PSF}} \rangle(\vartheta); \quad (18)$$

$$\rho_2(\vartheta) = \langle e^{\text{PSF}} \delta e^{\text{PSF}} \rangle(\vartheta); \quad \rho_3(\vartheta) = \langle \delta T^{\text{PSF}} \delta T^{\text{PSF}} \rangle(\vartheta); \quad (19)$$

$$\rho_4(\vartheta) = \langle \delta e^{\text{PSF}} \delta T^{\text{PSF}} \rangle(\vartheta); \quad \rho_5(\vartheta) = \langle e^{\text{PSF}} \delta T^{\text{PSF}} \rangle(\vartheta). \quad (20)$$

Despite these excesses, the amplitude of the correlations alone does not reveal the contribution of each  $\rho$ -statistic to the PSF systematic bias. One must first estimate the values of  $\alpha$ ,  $\beta$ , and  $\eta$ . We perform this estimation introducing Galaxy-PSF correlations also known as  $\tau$ -statistics (Hamana et al. 2020; Giblin et al. 2021a; Gatti et al. 2021; Zhang et al. 2023; Guerrini et al. 2025). They are defined as

$$\tau_0(\vartheta) = \langle e^{\text{gal}} e^{\text{PSF}} \rangle(\vartheta); \quad \tau_2(\vartheta) = \langle e^{\text{gal}} \delta e^{\text{PSF}} \rangle(\vartheta); \quad (21)$$

$$\tau_5(\vartheta) = \langle e^{\text{gal}} \delta T^{\text{PSF}} \rangle(\vartheta). \quad (22)$$

Figure 16 shows the three  $\tau$ -statistics with and without applying the object-wise leakage correction described in Section 3.2.3. The object-wise correction reduces the amplitude of  $\tau_0$  and  $\tau_5$ , which is expected.  $\tau_0$  directly relates to the amount of leakage and  $\tau_5$  is strongly correlated with  $\tau_0$ . This correlation becomes clear in the constraints on  $\alpha$ ,  $\beta$ , and  $\eta$  below.

The measurement of those two types of correlation functions

allows us to estimate the parameters of the PSF ellipticity error model  $\Omega = (\alpha, \beta, \eta)^T$ . Using notations from Guerrini et al. (2025), we perform it solving:

$$\tau = R\Omega + \Sigma, \quad (23)$$

where  $\tau$  is the flattened data vector of the three  $\tau$ -statistics and  $R$  is a matrix built from  $\rho$ -statistics. These quantities are related via the relations

$$\tau_0(\vartheta) = \alpha\rho_0(\vartheta) + \beta\rho_2(\vartheta) + \eta\rho_5(\vartheta); \quad (24)$$

$$\tau_2(\vartheta) = \alpha\rho_2(\vartheta) + \beta\rho_1(\vartheta) + \eta\rho_4(\vartheta); \quad (25)$$

$$\tau_5(\vartheta) = \alpha\rho_5(\vartheta) + \beta\rho_4(\vartheta) + \eta\rho_3(\vartheta). \quad (26)$$

We apply the methodology developed in Guerrini et al. (2025) and estimate  $\Omega$  using a semi-analytical covariance matrix and sampling with a least-square estimator. Constraints on  $\Omega$  are shown in Figure 17. The object-wise correction again reduces the  $\alpha$  parameter. The  $(\alpha, \eta)$ -plane shows a strong degeneracy between those parameters. This relates to the multiplication by  $e^{\text{PSF}}$  in the definition of the size residuals. This correlates the leakage and size error terms in such a way that a quantitative measurement of the leakage must be evaluated in the  $(\alpha, \eta)$ -plane rather than using the 1D posterior on  $\alpha$ . [SG: Should we add a plot of the uncorrelated fit to quote a value for the leakage?]

We propagate these estimates to assess the amplitude of the additive bias due to PSF systematics contributing to the two-point correlation functions  $\xi_{\pm}(\theta)$ . We refer to this additive systematic bias as  $\xi_{\pm}^{\text{sys}}(\theta)$ . It is estimated from the previously established  $[\alpha, \beta, \eta]$  coefficients as

$$\begin{aligned} \xi_{\text{PSF,sys}}(\vartheta) = & \alpha^2\rho_0(\vartheta) + \beta^2\rho_1(\vartheta) + \eta^2\rho_3(\vartheta) + \\ & 2\alpha\beta\rho_2(\vartheta) + 2\alpha\eta\rho_5(\vartheta) + 2\beta\eta\rho_4(\vartheta). \end{aligned} \quad (27)$$

This additive contribution should be negligible compared to the cosmological signal. We assess it from the ratio  $\xi_{\pm}^{\text{sys}}(\theta)/\xi_{\pm}(\theta)$ , selecting scales where the cosmological signal dominates. Figure 18 shows the ratio of the systematic contribution,  $\xi_{\pm}^{\text{sys}}(\theta)$ , with respect to the cosmological signal. With a threshold of 10%, the leakage correction clearly reduces the PSF systematic contribution to  $\xi_{\pm}(\theta)$  on all scales.

### 5.2.2. Object-wise PSF leakage

### 5.2.3. Scale-dependent PSF leakage

Omitting the  $\beta$  and  $\eta$  PSF residual terms from Eq. (16), the scale-dependent leakage is estimated from  $\tau_0(\vartheta)$  and  $\rho(\vartheta)$  as

$$\alpha(\vartheta) = \frac{\tau_0(\vartheta)}{\rho_0(\vartheta)}. \quad (28)$$

This function was introduced in Bacon et al. (2003) as a PSF diagnostic function. Following Section 5.2.1, the PSF systematic additive contribution to the two-point correlation function is

$$\xi_{\pm}^{\text{sys}}(\vartheta) = \alpha^2(\vartheta)\rho_0(\vartheta) = \frac{\tau_0^2(\vartheta)}{\rho_0(\vartheta)}. \quad (29)$$

Figures ?? and 19 show, respectively, the scale-dependent leakage and the resulting systematic additive bias. The scale-dependence of the leakage parameter is very mild, remaining mostly constant across scales. This estimator also shows a reduced systematic contribution from leakage after the object-wise

correction. Figure D.1 compares the predicted systematic additive bias estimated with the scale-dependent leakage and from  $\rho$ - and  $\tau$ -statistics of Section 5.2.1. The two estimates match on large scales. The discrepancy on small scales is due to the unmodeled  $\beta$  term in the scale-dependent estimator. The scale-dependent estimator is therefore incomplete, but yields consistent leakage amplitude and impact on the two-point correlation function on large scales.

**Magnitude-dependence** The leakage function  $\alpha(\theta)$  depends on galaxy magnitude. For a faint sample, the leakage is negative, cancelling out some of the leakage of brighter galaxies. UNIONS and DES show very similar behaviour.

### 5.3. Magnitude dependence

[SG: I will produce those plots as soon as possible. (At least  $\tau_0$ ). There is still the question of whether we have to recompute the weights and calibration for each magnitude/size bin? I would say that we don't as we are making a claim on the specific version of ShapePipe we are validating.]

Plot of PSF residuals with magnitude, see Gatti Fig. 6.

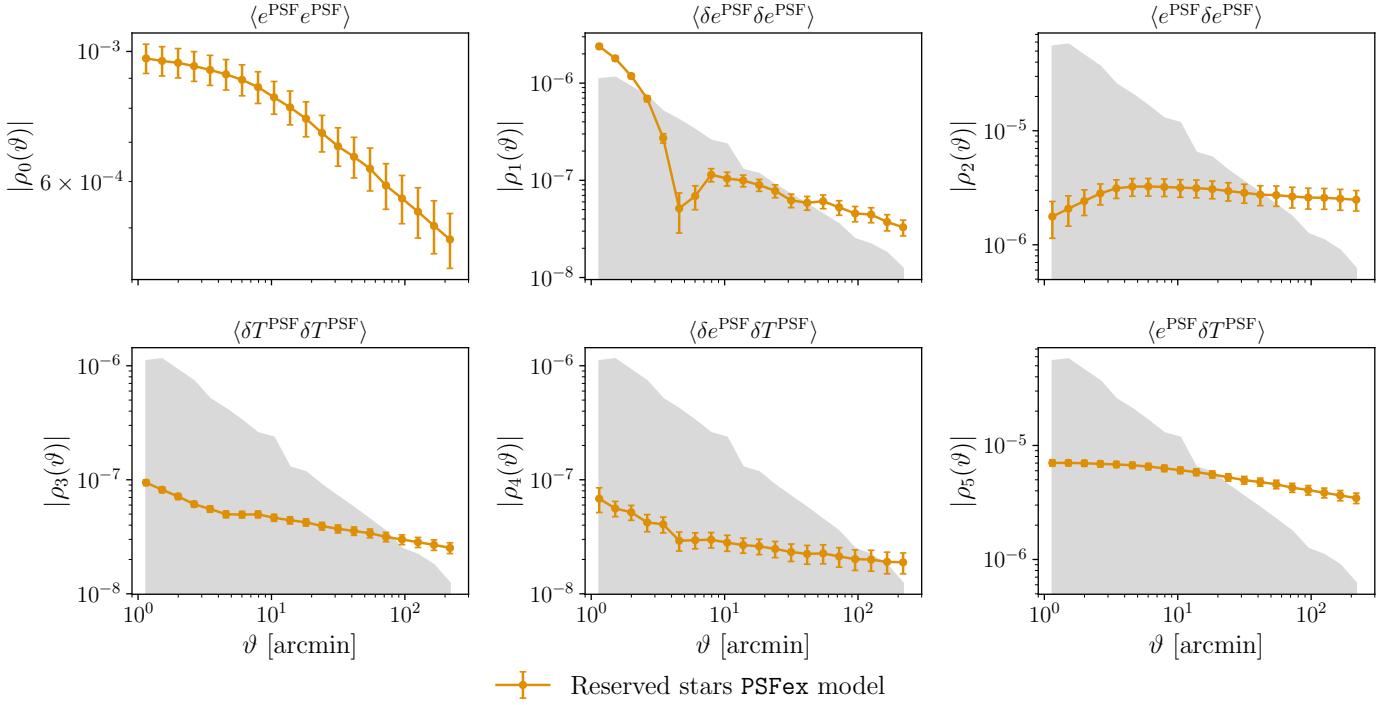
$\tau$ -statistics with different magnitude cuts.

### 5.4. Size dependence

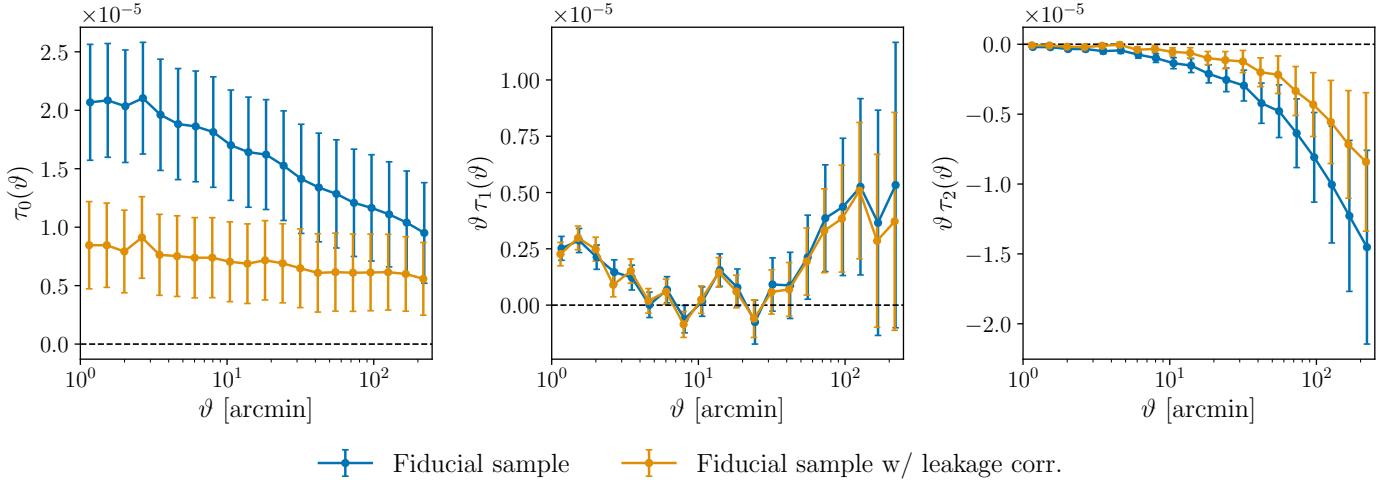
[SG: I added this plot in a previous section where I describe the object-wise leakage correction.] 2D plot of object-wise leakage as function of size and SNR/magnitude, as in KiDS+DES paper.

## 6. Conclusions

We present a robust weak lensing catalogue built on the excellent imaging data of Gwyn et al. (2025), following the methodology of Guinot et al. (2022) and using the ShapePipe pipeline (Farrens et al. 2022). This release covers imaging from  $3500 \text{ deg}^2$  of the sky resulting in an effective area  $A_{\text{eff}} = 2,894 \text{ deg}^2$  and effective number density  $n_{\text{eff}} = 4.96 \text{ arcmin}^{-2}$  for the fiducial v1.4.6.3 catalog. We initially described the pipeline and its different components. To give an understanding of the different UNIONS lensing papers published so far we detail the versioning process adopted by the collaboration, with the key difference being the PSF model used, with MCD applied to v1.3 and PSFEx used in v1.4. Inspecting the ellipticity dependence on various observational quantities revealed strong PSF leakage. In Sect. 5 we lay out the star selection and discuss an involved method based on star-star correlations ( $\rho$ -statistics) and star-galaxy correlation functions ( $\tau$ -statistics) to evaluate a systematic contamination to the two-point galaxy-galaxy ( $\xi_{\text{sys}}$ ). This allows us to propagate the additional PSF signal into the correlation function and determine scale-cuts to use in a cosmological analysis to keep the PSF contributions below an acceptable threshold. To further verify the robustness of the catalogue, we run a suite of null tests. The ellipticity distribution in the two core processing units, tiles & CCDs, shows no correlation with position. Similarly the galaxy-galaxy lensing around tile and CCD centres is consistent with zero. To verify that bright stars do not leave a coherent imprint on galaxy shapes we compute the galaxy-galaxy lensing signal around our selected stars as well as GAIA stars for an independent sample. All are consistent with the null hypothesis, indicating that problematic regions have been masked out. In modern weak lensing surveys the detection of B-modes often serves as a key systematic test to indicate the presence of spurious signals,



**Fig. 15.**  $p$ -statistics measured on the reserved stars of the fiducial lensing sample. Shaded regions correspond to requirements derived following Mandelbaum et al. (2018). Excess on large scales create PSF leakage quantified in Section 5.2.1.



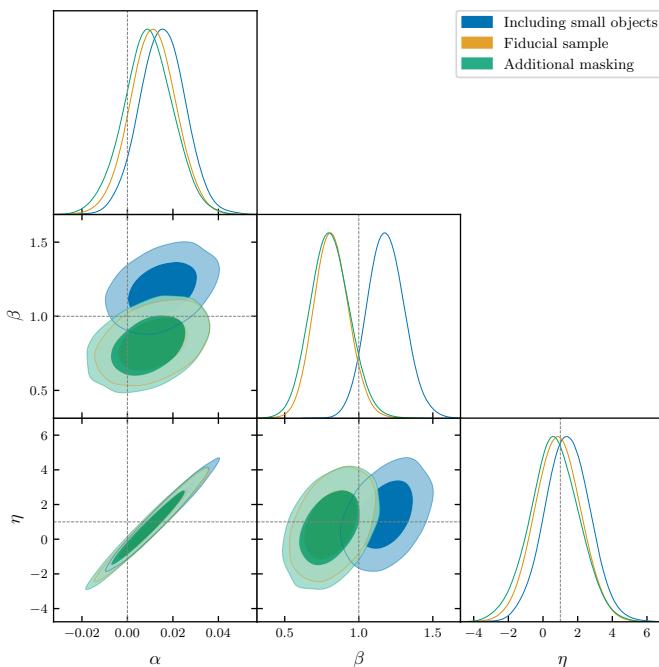
**Fig. 16.**  $\tau$ -statistics measured on the fiducial weak lensing sample (SP\_v1.4.6). Blue curve corresponds to the measurement before the object-wise leakage correction is applied. Yellow curve is obtained with the same objects after applying the object-wise correction described in Section 3.2.3.

since no physical signal is known to produce curl-like correlation functions at the precision levels at which stage-III surveys operate. A dedicated paper (Paper III) explores these B-mode correlations using pure-mode correlation functions  $\xi_{\pm}^{E/B}(\theta)$ , COSEBIS, and harmonic-space angular power spectra  $C_{\ell}^{BB}$ . **FHP: need one or two sentences summarising the paper, should this go into the body of text?.**

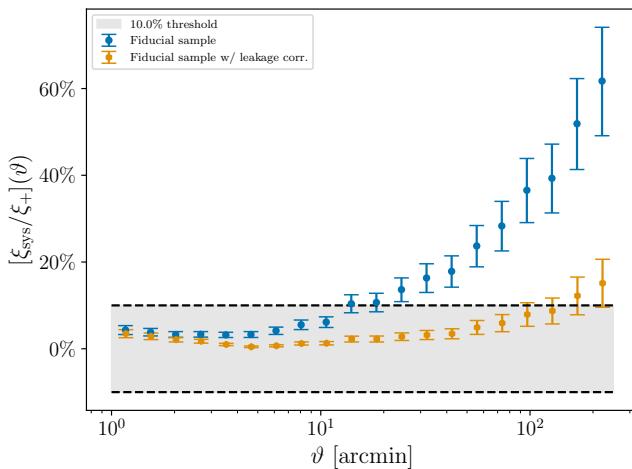
This paper is the first in a suite of works aimed at using this data to obtain cosmological constraints, in particular on the lensing amplitude  $S_8$ . **Each companion paper contributes to this effort:**

- Paper II details the calibration efforts for both shear and redshift, relying on external simulations to validate the methodology and estimate residual biases.
- Paper III presents detailed B-mode calculations, further validating this data-set and exploring the scales on which a robust cosmological analysis is feasible.
- Paper IV shows the first application of this data set to constrain cosmological parameters using cosmic shear measurements.
- Paper V presents harmonic-space cosmological constraints from this catalogue.

While some published works already used this data set, many This catalogue offers strong possibilities of synergies with data sets in the northern sky. The anticipated DESI-DR2 catalog



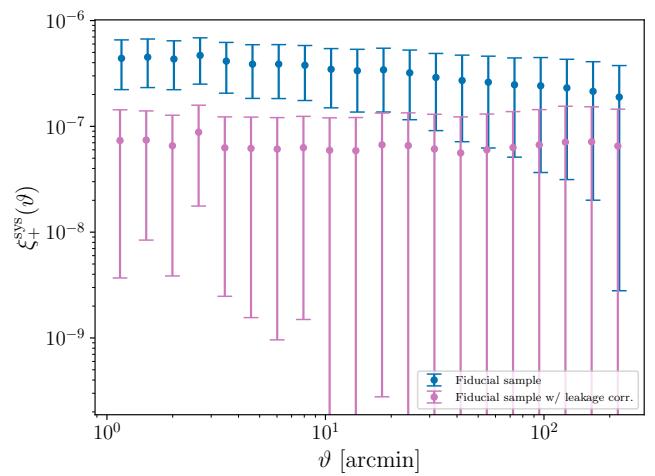
**Fig. 17.** Constraints on PSF parameters  $\alpha$ ,  $\beta$  and  $\eta$  obtained using a semi-analytical covariance matrix and a least-square estimator. Dark (light) shaded colours indicate 68.3% (95.5%) confidence regions.



**Fig. 18.** Ratio of the PSF systematic additive bias  $\xi_{+}^{\text{sys}}(\vartheta)$  and  $\xi_{+}(\vartheta)$  for the object-wise corrected (blue points) and uncorrected (orange) catalogs.

should offer unprecedented overlap with UNIONS. Similarly J-Plus 2nd Jpas While it can be expected that *Euclid* DR2 supersedes this catalogue on the same area in the spring of 2029, according to the current schedule, this leaves an interesting gap to explore cosmological and astrophysical questions relying on lensing data. To extract the full cosmological constraining power from the UNIONS data, it is necessary to add photometric redshifts to gain some insight on the line-of-sight distances of galaxies. These efforts are currently in progress to allow a full tomographical analysis.

**Acknowledgements.** This work was made possible by utilising the CANDIDE cluster at the Institut d’Astrophysique de Paris. The cluster was funded through grants from the PNCG, CNES, DIM-ACAV, the Euclid Consortium, and the Dan-



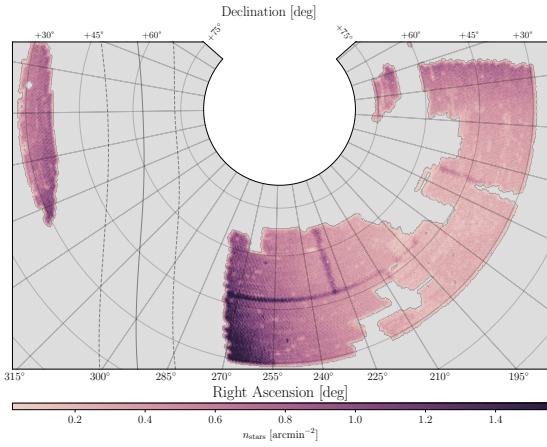
**Fig. 19.** PSF systematic additive bias estimated from the scale-dependent leakage. Blue (pink) points show the case without (with) object-wise leakage correction; error bars are jackknife estimates.

ish National Research Foundation Cosmic Dawn Center (DNRF140). It is maintained by Stephane Rouberol. FHP acknowledges support from CNES.

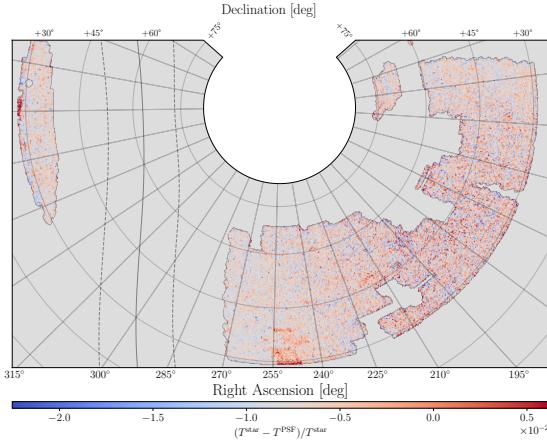
## References

- Abbott, T. M. C., Adamów, M., Aguena, M., et al. 2021, ApJ Supplement Series, 255, 20
- Ahad, S. L. et al. 2025 [[arXiv:2512.14636](https://arxiv.org/abs/2512.14636)]
- Albrecht, A., Bernstein, G., Cahn, R., et al. 2006, arXiv, 0609591
- Anbaganage, D. et al. 2025, Open J. Astrophys., 8, 146158
- Astier, P., El Hage, P., Guy, J., et al. 2013, A&A, 557, A55
- Ayçoberry, E., Ajani, V., Guinot, A., et al. 2023, A&A, 671, A17
- Bacon, D. J., Massey, R. J., Refregier, A. R., & Ellis, R. S. 2003, Mon Not R Astron Soc, 344, 673
- Bernstein, G. M. & Jarvis, M. 2002, AJ, 123, 583
- Bertin, E. 2011, in Astronomical Data Analysis Software and Systems XX, Vol. 442, 435
- Bertin, E. & Arnouts, S. 1996, AAPs, 117, 393
- Boulade, O., Charlot, X., Abbon, P., et al. 2003, in Instrument Design and Performance for Optical/Infrared Ground-based Telescopes, Vol. 4841, 72–81
- Chambers, K. C., Magnier, E. A., Metcalfe, N., et al. 2019, ArXiv e-prints [[arXiv:1612.05560](https://arxiv.org/abs/1612.05560)]
- Cheng, I., Elvin-Poole, J., Hudson, M. J., et al. 2025, arXiv, 2502.00584
- Congedo, G. et al. 2024, Astron. Astrophys., 691, A319
- CosmoStat. 2026
- Cropper, M., Hoekstra, H., Kitching, T., et al. 2013, MNRAS, 431, 3103
- Daley, C. et al. 2026
- DES and KiDS Collaboration, Abbott, T. M. C., Aguena, M., et al. 2023, OpJA, 6
- Erben, T., Hildebrandt, H., Miller, L., et al. 2013, MNRAS, 433, 2545
- Euclid Collaboration, Vinciguerra, S., Bouchè, F., et al. 2025, ArXiv e-prints [[arXiv:2510.04953](https://arxiv.org/abs/2510.04953)]
- Farréns, S., Guinot, A., Kilbinger, M., et al. 2022, A&A, 664, A141
- Gatti, M., Sheldon, E., Amon, A., et al. 2021, MNRAS, 504, 4312
- Giblin, B., Heymans, C., Asgari, M., et al. 2021a, Astronomy & Astrophysics, 645, A105
- Giblin, B., Heymans, C., Asgari, M., et al. 2021b, A&A, 645, A105, arXiv: 2007.01845
- Goh, L. W. K. et al. 2026
- Guerrini, S., Kilbinger, M., Leterme, H., et al. 2025, A&A, 700, A215
- Guerrini, S. et al. 2026
- Guinot, A., Kilbinger, M., Farréns, S., et al. 2022, A&A, 666, A162
- Guyonnet, A., Astier, P., Antilogus, P., Regnault, N., & Doherty, P. 2015, A&A, 575, A41
- Gwyn, S., McConnachie, A. W., Cuillandre, J.-C., et al. 2025, ArXiv e-prints [[arXiv:2503.13783](https://arxiv.org/abs/2503.13783)], comment: submitted to AJ
- Hamana, T., Shirasaki, M., Miyazaki, S., et al. 2020, Publications of the Astronomical Society of Japan, 72
- Hervas Peters, F., Kilbinger, M., Paviot, R., et al. 2025, A&A, 699, A201
- Hervas-Peters, F. et al. 2026

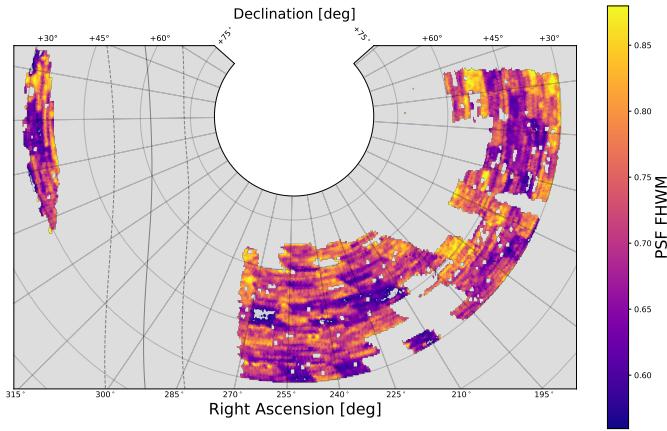
- Heymans, C., Van Waerbeke, L., Miller, L., et al. 2012, *MNRAS*, 427, 146  
 Hirata, C. & Seljak, U. 2003, *MNRAS*, 343, 459  
 Hoekstra, H., Mellier, Y., Van Waerbeke, L., et al. 2006, *ApJ*, 647, 116  
 Hogg, D. W. & Lang, D. 2013, *PASP*, 125, 719  
 Huang, S., Leauthaud, A., Hearin, A., et al. 2020, *MNRAS*, 492, 3685  
 Huff, E. & Mandelbaum, R. 2017, ArXiv e-prints [[arXiv:1702.02600](#)]  
 Jarvis, M., Sheldon, E., Zuntz, J., et al. 2016, *Mon Not R Astron Soc*, 460, 2245  
 Jefferson, J., Omori, Y., Chang, C., et al. 2025, *OpJA*, 8  
 Jeffrey, N. et al. 2021, *Mon. Not. Roy. Astron. Soc.*, 505, 4626  
 Kaiser, N., Squires, G., & Broadhurst, T. 1995, *APJ*, 449, 460  
 Kilbinger, M. 2015, *Rep. Prog. Phys.*, 78, 086901  
 Lagos, C. d. P., Tobar, R. J., Robotham, A. S. G., et al. 2018, *MNRAS*, 481, 3573  
 Li, Q., Kilbinger, M., Luo, W., et al. 2024, *ApJL*, 969, L25  
 Li, S.-S., Hoekstra, H., Kuijken, K., Schaller, M., & Schaye, J. 2025, *Astron. Astrophys.*, 700, A202  
 Li, S.-S., Kuijken, K., Hoekstra, H., et al. 2023a, *A&A*, 670, A100  
 Li, X., Zhang, T., Sugiyama, S., et al. 2023b, *Phys. Rev. D*, 108, 123518  
 Li, X. et al. 2022, *Publ. Astron. Soc. Jap.*, 74, 421  
 Lioudat, T., Bonnин, J., Starck, J.-L., et al. 2021, *A&A*, 646, A27  
 Mandelbaum, R. 2018, *ARA&A*, 56, 393  
 Mandelbaum, R., Miyatake, H., Hamana, T., et al. 2018, *PASJ*, 70, S25  
 Mandelbaum, R., Rowe, B., Bosch, J., et al. 2014, *The Astrophysical Journal Supplement Series*, 212, 5  
 Mandelbaum, R. et al. 2015, *Mon. Not. Roy. Astron. Soc.*, 450, 2963  
 Marmo, C. & Bertin, E. 2008, in *Astronomical Society of the Pacific Conference Series*, Vol. 394, *Astronomical Data Analysis Software and Systems XVII*, ed. R. W. Argyle, P. S. Bunclark, & J. R. Lewis, 619  
 Martin, H. L., Hudson, M. J., Woodfinden, A., et al. 2025, arXiv, 2507.13450  
 Melchior, P., Sutter, P. M., Sheldon, E. S., Krause, E., & Wandelt, B. D. 2014, *Mon. Not. Roy. Astron. Soc.*, 440, 2922  
 Miller, L., Heymans, C., Kitching, T. D., et al. 2013, *MNRAS*, 429, 2858, arXiv:1210.8201  
 Mpetha, C. T., Taylor, J. E., Amoura, Y., et al. 2025, arXiv, 2501.09147  
 Paillassa, M., Bertin, E., & Bouy, H. 2020, *A&A*, 634, A48  
 Paulin-Henriksson, S., Amara, A., Voigt, L., Refregier, A., & Bridle, S. L. 2008, *A&A*, 484, 67  
 Prat, J. & Bacon, D. 2026, in *Weak Gravitational Lensing*, 508–537  
 Robison, B., Hudson, M. J., Cuillandre, J.-C., et al. 2023, *Monthly Notices of the Royal Astronomical Society*, 523, 1614  
 Rowe, B. 2010, *Monthly Notices of the Royal Astronomical Society*, 404, 350  
 Sérsic, J. L. 1963, *Boletín de la Asociación Argentina de Astronomía La Plata Argentina*, 6, 41  
 Sevilla-Noarbe, I., Bechtol, K., Carrasco Kind, M., et al. 2021, *ApJS*, 254, 24  
 Sheldon, E. 2015, *ASCL*, ascl:1508.008  
 Sheldon, E. S., Becker, M. R., MacCrann, N., & Jarvis, M. 2020, *Astrophys. J.*, 902, 138  
 Sheldon, E. S. & Huff, E. M. 2017, *ApJ*, 841, 24  
 Singh, S., Mandelbaum, R., & More, S. 2015, *MNRAS*, 450, 2195  
 Van Waerbeke, L., Mellier, Y., Erben, T., et al. 2000, *Ap*, 358, 30  
 Wright, A. H. et al. 2024, *Astron. Astrophys.*, 686, A170  
 Wright, A. H. et al. 2025 [[arXiv:2503.19441](#)]  
 Xia, Q. et al. 2020, *Astron. Astrophys.*, 633, A89  
 Yamamoto, M., Becker, M. R., Sheldon, E., et al. 2025, arXiv e-prints, arXiv:2501.05665  
 Yoon, M. et al. 2025 [[arXiv:2510.01122](#)]  
 York, D. G., Adelman, J., John E. Anderson, J., et al. 2000, *AJ*, 120, 1579  
 Zhang, T., Li, X., Dalal, R., et al. 2023, *MNRAS*, 525, 2441  
 Zhang, Z., Becker, M. R., & Sheldon, E. S. 2022 [[arXiv:2206.07683](#)]  
 Zhang, Z., Kilbinger, M., Peters, F. H., et al. 2024, *A&A*, 691, A75  
 Zuntz, J., Kacprzak, T., Voigt, L., et al. 2013, *Monthly Notices of the Royal Astronomical Society*, 434, 1604  
 Zuntz, J., Sheldon, E., Samuroff, S., et al. 2018, *MNRAS*, 481, 1149



**Fig. A.1.** Distribution of stars detected in the survey. The density increases towards the galactic plane as expected. This map appears correlated with the PSF diagnostic maps presented below.



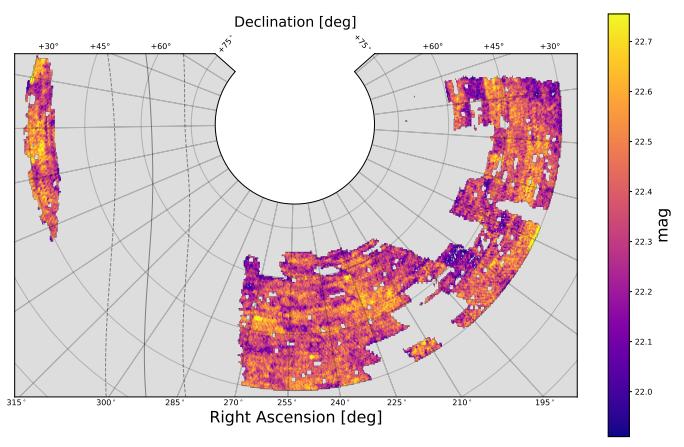
**Fig. A.2.** Distribution of star size residuals from the PSF model estimated on reserve stars.



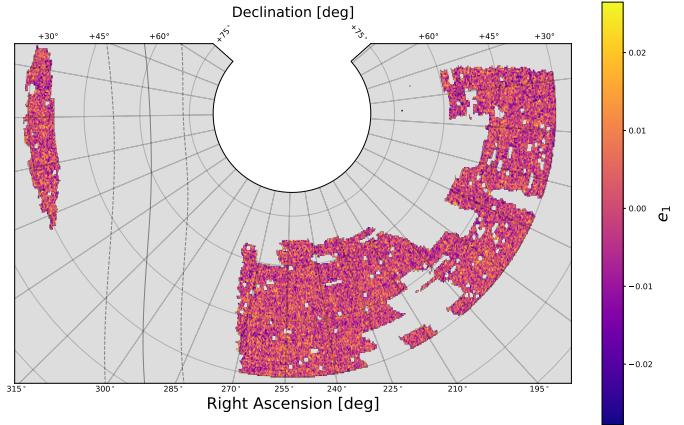
**Fig. A.3.** Caption

## Appendix A: Star & PSF related maps

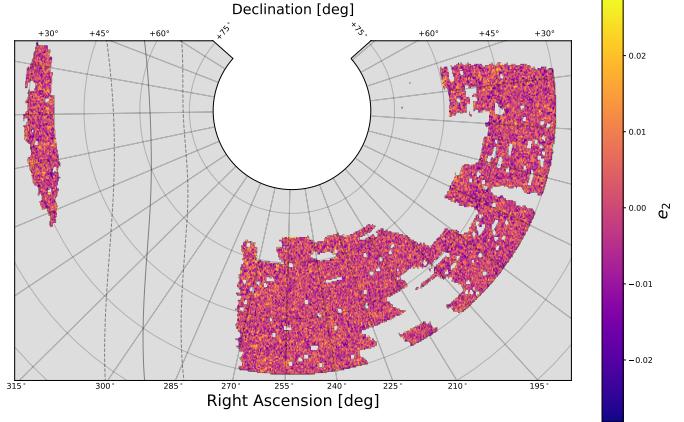
Here we present the sky-coverage of different quantities from our final selected galaxy and star sample



**Fig. A.4.** Caption



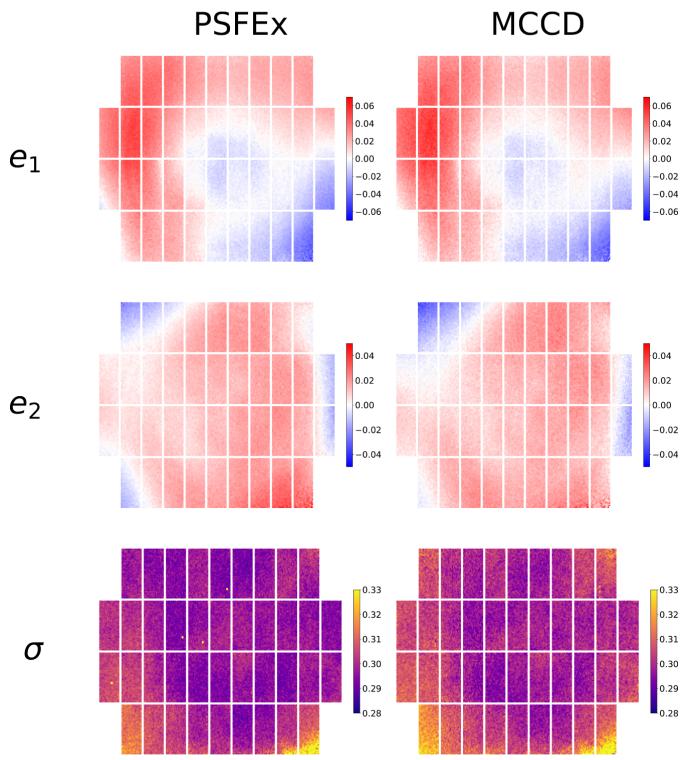
**Fig. A.5.** Caption



**Fig. A.6.** Caption

## Appendix B: Comparing PSF modelling using MCDD to PSFEx

PSFEx is a well-established PSF model widely used in the astronomical community ?. It represents the PSF as a combination of basis vectors estimated directly from stars, known as eigenPSFs. Separate models are constructed for each CCD in the focal plane. The approach relies on matrix factorisation: one matrix captures the fundamental PSF patterns, while the other provides coeffi-



**Fig. B.1.** A comparison of the PSFEx and MCCD models on the focal plane. One can distinguish complex patterns linked to the

clients describing how these patterns combine. Each observed PSF is expressed as a weighted sum of these features. This allows us to obtain a PSF model for any position on the CCD.

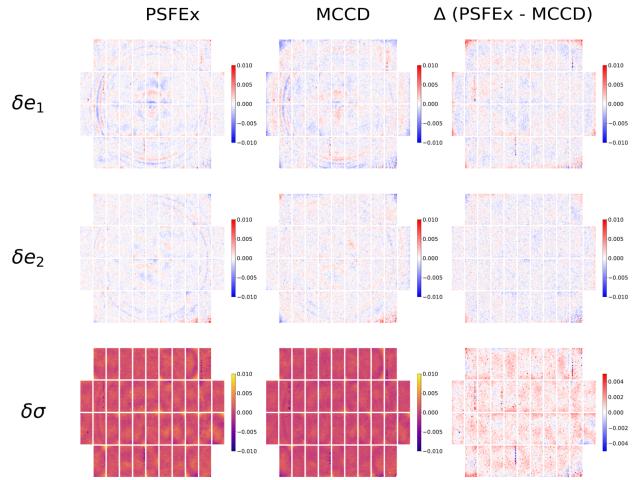
MCCD, described in ???, also relies on eigenPSFs. The model is optimised on the full focal plane. This means the model can be built with more stars, at the cost of an increased complexity, as it needs to capture cross-CCD effects. ? show from simulations that MCCD improves on PSFEx by ~50% in pixel RMSE and ~25% in the prediction of  $e_1$  and  $e_2$ .

To ensure the robustness of our PSF estimation, we created one shear catalogue with MCCD (v.1.3) and one with PSFEx (v1.4). Below we compare PSF model performances and PSF-leakage estimation.

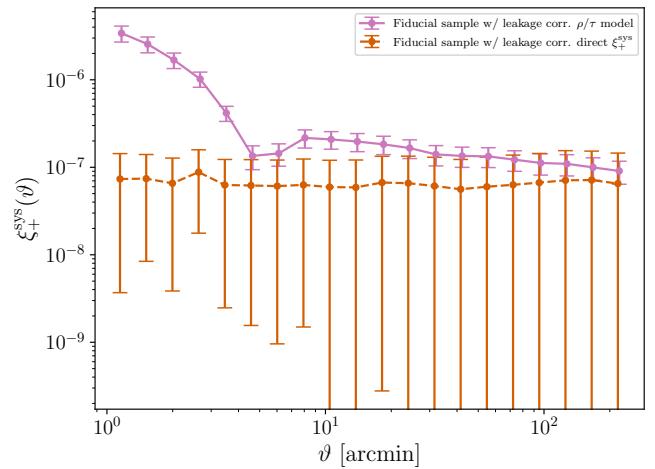
### Appendix C: User-defined catalogues

We create a comprehensive catalogue with all tile detections and all mask entries. The user can then apply any mask and galaxy selection criteria they desire. After applying all cuts, the user needs to calibrate the sample using the metacalibration method and compute the DES-like sample weights. We provide code to carry out those steps.

### Appendix D: Comparison of $\rho$ & $\tau$ derived leakage vs. scale-dependent $\alpha$



**Fig. B.2.** A comparison of the PSFEx and MCCD model residuals on the focal plane. Complex features that neither model could capture are visible. See text for details.



**Fig. D.1.** Comparison of the PSF systematic additive bias estimated from  $\rho$ - and  $\tau$ -statistics and from scale-dependent leakage.