

Accurate shear estimation with fourth-order moments

Andy Park ¹★, Xiangchong Li ^{1,2} and Rachel Mandelbaum ¹

¹Department of Physics, McWilliams Center for Cosmology and Astrophysics, Carnegie Mellon University, 5000 Forbes Ave, Pittsburgh, PA 15213, USA

²Brookhaven National Laboratory, Bldg 510, Upton, New York 11973, USA

Accepted 2025 January 7. Received 2025 January 7; in original form 2024 November 21

ABSTRACT

As imaging surveys progress in exploring the large-scale structure of the Universe through the use of weak gravitational lensing, achieving sub-per cent accuracy in estimating shape distortions caused by lensing, or shear, is imperative for precision cosmology. In this paper, we extend the Fourier power function shapelets (FPFS) shear estimator using fourth-order shapelet moments and combine it with the original second-order shear estimator to reduce galaxy shape noise. We calibrate this novel shear estimator analytically to a sub-per cent level-accuracy using the AnaCal framework. This higher order shear estimator is tested with realistic image simulations, and after analytical correction for the detection/selection bias and noise bias, the multiplicative shear bias $|m|$ is below 3×10^{-3} (99.7 per cent confidence interval) for both isolated and blended galaxies. Once combined with the second-order FPFS shear estimator, the shape noise is reduced by ~ 35 per cent for isolated galaxies in simulations with Hyper Suprime-Cam and Vera C. Rubin Observatory Legacy Survey of Space and Time observational conditions. However, for blended galaxies, the effective number density does not significantly improve with the combination of the two estimators. Based on these results, we recommend exploration of how this framework can further reduce the systematic uncertainties in shear due to point spread function leakage and modelling error, and potentially provide improved precision in shear inference in high-resolution space-based images.

Key words: gravitational lensing: weak – techniques: image processing – cosmology: observations.

1 INTRODUCTION

Weak gravitational lensing refers to the small but coherent distortions of distant galaxies' light profiles due to intervening massive foreground matter between the source galaxies and observers. Analysing the statistics of this distortion can reveal the distribution of matter and the large-scale structure, yielding deep insights into the fundamental physics of the Universe and its evolution (see Kilbinger 2015 for a review of weak lensing). Measuring the coherent pattern of distortions, or shear, in the observed image is one of the most effective ways to constrain the fundamental physics of the Universe. Upcoming Stage IV imaging surveys like the Vera C. Rubin Observatory Legacy Survey of Space and Time (LSST; LSST Science Collaboration 2009; Ivezić et al. 2019), *Euclid* (Laureijs et al. 2011), and *Nancy Grace Roman Space Telescope* High Latitude Imaging Survey (Akeson et al. 2019) will cover a large sky area and observe more than billion source galaxies, allowing us to make unprecedented per cent-level cosmic shear measurements.

However, the magnitude of this tiny distortion, which causes per cent-level changes in the ellipticities of the observed galaxy images, is, on average, ~ 10 per cent of the root mean square (RMS) of intrinsic galaxy shapes. Hence, an accurate measurement of weak-lensing shear is complex and needs to calibrate several sources of systematic effects that complicate the process of characterizing

galaxy shapes, including the point spread function (PSF) from atmospheric effects and telescope optics (e.g. Liaudat, Starck & Kilbinger 2023); noise bias from image noise due to the non-linearity in shear estimator (e.g. Refregier et al. 2012); model bias due to unrealistic assumptions about galaxy morphology (e.g. Bernstein 2010); selection bias due to sample selection (Kaiser, Squires & Broadhurst 1995) and detection (Sheldon et al. 2020); and biases from blending and deblending of galaxy light profiles (Li et al. 2018). To achieve per cent-level bias in cosmological parameters despite systematics (see Mandelbaum 2018, for a review of systematics in weak lensing), the upcoming surveys require that the residual systematics after all corrections should be well below the statistical uncertainty. This requires that residual systematic uncertainties in shear measurements be below one part per thousand (The LSST Dark Energy Science Collaboration 2018).

Several recent shear measurement techniques have been developed that aim to reduce the reliance on calibration using external simulations. This is desirable because the simulations may not be sufficiently realistic to achieve the level of precision needed for cosmological inference. These methods that aim for unbiased shear inference include METADETECTION (Sheldon et al. 2023), a numerical self-calibration method; and BFD (Bernstein et al. 2016), a Bayesian approach to shear estimation. There has been significant efforts to develop purely analytical shear estimators (Li et al. 2018, 2024b; Li, Li & Massey 2022b; Li & Mandelbaum 2023; Li, Mandelbaum & The LSST Dark Energy Science Collaboration 2024a) in order to achieve sub-per cent accuracy without relying on

* E-mail: chanhyup@andrew.cmu.edu

any calibration from external image simulations. The Fourier power function shapelets method (FPFS hereinafter) can achieve a sub-per cent accuracy by correcting for noise bias, selection and detection bias. In short, FPFS uses set of shapelet modes (Bernstein & Jarvis 2002; Refregier 2003; Massey & Refregier 2005), detection modes and other linear observables after PSF deconvolution to carry out the detection, selection, and measurement of the galaxy ellipticity and response to shear. It uses the first-order derivatives (Jacobian) of the ellipticity with respect to these linear observables and the shear responses of shapelet to obtain the shear response of the ellipticity. To correct for the noise bias, one can use the second-order derivative (Hessian matrix) of the ellipticity and the covariance matrix of the measurement error on the observables (Li et al. 2024b), or add pure image noise to the galaxy image with a carefully chosen noise correlation function to derive an analytical noise bias correction (Li et al. 2024a). This innovative analytical calibration technique for shear estimation, designated AnaCal (Li & Mandelbaum 2023; Li et al. 2024b) is more than a hundred times faster than the current benchmark METADETECTION algorithm, documented in Sheldon et al. (2023).

Lensing shear causes the estimates of the ellipticities of distant galaxies to change. This effect is quantified statistically by observing the characteristics of a galaxy that undergoes simple transformations under shear. One way to parameterize the shape is based on second-order moments of the galaxy image, which captures the orientation and elongation of the object. These moments have been widely adopted in weak-lensing studies due to their robustness and relative insensitivity to image noise. However, higher order moments, while more sensitive to image noise and harder to model, offer the potential to capture additional information about the galaxy's response to shear. They probe finer structural details that are not accessible through second moments alone, suggesting they may contain complementary information about the shear. The original FPFS implementation primarily focused on second-order shapelet moments to construct the ellipticity of galaxies, and achieves a shear estimation bias below 0.3 per cent in the presence of blending. In this work, we extend the FPFS framework by incorporating fourth-order shapelet moments to define the ellipticity. We use image simulations to test the accuracy of our new shear estimator after analytically correcting for detection and selection bias. We then take two FPFS shear estimators, second- and fourth-order, and combine them to maximize the shear signal and reduce the overall shape noise. Our goal is to quantify how much additional information is provided by the fourth-order moments; the fourth-order shear estimator, which is independent of the second-order, can also be used to cross-comparison and improve the systematic control.

This paper is organized as follows. In Section 2, we give a brief overview of the FPFS shear estimator within the AnaCal framework, the fourth-order shear estimator, and introduce a method to combine two different shear estimators. In Section 3, we present galaxy image simulations that we used to test the accuracy of our new fourth-order shear estimator. In Section 4, we show the result of our analysis and quantify the reduction of shape noise by combining two shear estimators. Finally, in Section 5, we summarize our results and future outlook.

2 METHOD

In this section, we briefly review the FPFS shear estimator developed in Li et al. (2018, 2022b) and calibrated with the AnaCal framework implemented in Li & Mandelbaum (2023) and Li et al. (2024b). We then extend FPFS to a higher order shear estimator.

The distortion of galaxy shapes, or shear, caused by foreground inhomogeneous mass distribution can be described by a locally linear transformation (or the Jacobian matrix) as

$$A = \begin{pmatrix} 1 - \gamma_1 & -\gamma_2 \\ -\gamma_2 & 1 + \gamma_1 \end{pmatrix}, \quad (1)$$

where the component γ_1 quantifies the amount of stretching of the image along the horizontal direction and γ_2 quantifies the stretching of the image in the direction at an angle of 45° with the horizontal direction. We use a complex spinor to represent shear as $\gamma = \gamma_1 + i\gamma_2$, where i is the imaginary number unit. In this paper, we set the lensing convergence to zero to simplify the notation. The shear is typically estimated by measuring the galaxy's ellipticity, a spin-2 observable, which negates under a 90° rotation (see appendix B of Li & Mandelbaum 2023 for more details). In the weak-lensing limit, shear is on the order of a few per cent or less ($|\gamma| \lesssim 0.02$), making the shear signal much smaller than the shape noise due to galaxy intrinsic shapes, so that a large ensemble of galaxies is needed to infer shear.

2.1 Galaxy detection

Before measuring shear, detection, and selection of galaxies can introduce biases that affect the shear estimation, hence it is essential to derive their shear response for an accurate shear estimation. Galaxy detection from images in FPFS uses four detection modes (v_i , where $i = 0 \dots 3$) for each pixel in the image to characterize the difference in the value of the pixels with respect to the nearby pixels in four directions. These nearby pixel detection modes are then used to identify peaks that are served as 'peak candidates'. The corresponding selection bias from carrying out the detection process using these peak modes is analytically corrected using the shear responses of the pixel values (Li et al. 2024a, 2022b). For a galaxy profile $f(\mathbf{x})$, with \mathbf{x} denoting image in real space, we define the detection modes for every pixel i as

$$v_i = \iint d^2k \psi_i^*(\mathbf{k}) \frac{f^p(\mathbf{k})}{p(\mathbf{k})}, \quad (2)$$

where $f^p(\mathbf{k})$ is the observed (PSF-convolved, noisy) image in Fourier space and $p(\mathbf{k})$ is the PSF image in Fourier space, and the coordinate centre is set to the centre of this pixel. The detection kernels $\psi_i^*(\mathbf{k})$ for wave number vector $\mathbf{k} = (k_1, k_2)$ are defined in Fourier space as

$$\psi_i^*(\mathbf{k}) = \frac{1}{(2\pi)^2} e^{-|\mathbf{k}|^2 \sigma_h^2 / 2} (1 - e^{i\mathbf{k} \cdot \mathbf{x}_i}), \quad (3)$$

where $\mathbf{x}_i = (x_i, y_i) = (\cos(i\pi/2), \sin(i\pi/2))$ are position vectors to nearby pixels with lengths equal to the image pixel side length and orientations pointing towards the four directions separated by $\pi/2$. The shear response of these detection modes is given by

$$v_{i;\alpha} = \iint d^2k \psi_{i;\alpha}^* \frac{f^p(\mathbf{k})}{p(\mathbf{k})}, \quad (4)$$

where the subscript ' α ' denotes the partial derivative with respect to one component of the shear, γ_α . The shear response of each detection kernel can be written as a combination of shapelet basis and is given by

$$\begin{aligned} \psi_{i;1}^* &\equiv \frac{\partial \psi_i^*}{\partial \gamma_1} = \frac{1}{(2\pi)^2} e^{|\mathbf{k}|^2 \sigma_h^2 / 2} (k_1^2 - k_2^2) \sigma_h^2 (1 - e^{i(k_1 x_i + k_2 y_i)}) \\ &\quad - \frac{1}{(2\pi)^2} e^{|\mathbf{k}|^2 \sigma_h^2 / 2} (ix_i k_1 - iy_i k_2) e^{i(k_1 x_i + k_2 y_i)}, \end{aligned} \quad (5)$$

$$\psi_{i;2}^* \equiv \frac{\partial \psi_i^*}{\partial \gamma_2} = \frac{1}{(2\pi)^2} e^{i(k_1^2 \sigma_h^2/2 + 2k_1 k_2) \sigma_h^2} (1 - e^{i(k_1 x_i + k_2 y_i)}) - \frac{1}{(2\pi)^2} e^{i(k_1^2 \sigma_h^2/2 + 2k_1 k_2) \sigma_h^2} (i y_i k_1 + i x_i k_2) e^{i(k_1 x_i + k_2 y_i)}. \quad (6)$$

2.2 FPFs shapelet modes

FPFS uses polar shapelet modes (Massey & Refregier 2005) to construct various galaxy properties, including flux, size, and shape. These polar shapelet modes are constructed by projecting the observed noisy galaxy image after PSF deconvolution on to a set of Gaussian-weighted orthogonal functions (Li et al. 2018). The FPFs complex polar shapelet modes are defined as

$$M_{nm} \equiv \iint d^2k \tilde{\chi}_{nm}^*(\mathbf{k}) \frac{f^p(\mathbf{k})}{p(\mathbf{k})}, \quad (7)$$

where $\tilde{\chi}_{nm}$ is the Fourier transform of the polar shapelet basis function characterized by a radial quantum number ‘ n ’ and an angular quantum number, or spin number, ‘ m ’. Polar shapelet basis functions χ_{nm} are defined as

$$\chi_{nm}(\mathbf{x} | \sigma_h) = (-1)^{(n-|m|)/2} \left\{ \frac{[(n-|m|)/2]!}{[(n+|m|)/2]!} \right\}^{1/2} \times \left(\frac{|\mathbf{x}|}{\sigma_h} \right)^{|m|} L_{\frac{n-|m|}{2}}^{|m|} \left(\frac{|\mathbf{x}|^2}{\sigma_h^2} \right) e^{-|\mathbf{x}|^2/2\sigma_h^2} e^{-im\theta}, \quad (8)$$

where $L_{\frac{n-|m|}{2}}^{|m|}$ are the Laguerre polynomials and σ_h is the smoothing scale of shapelets and the detection kernel. n can be any non-negative integer and m is an integer between $-n$ and n in steps of two.

Under a Fourier transform, the shapelet basis changes as

$$\chi_{nm}(\mathbf{x} | \sigma_h) \rightarrow \tilde{\chi}_{nm}(\mathbf{k}) = i^n \chi_{nm}(\mathbf{k} | 1/\sigma_h), \quad (9)$$

so the amplitude of the shapelet basis in Fourier space is the same as in real space but the scale is the inverse of that in real space (Refregier 2003). Typically, σ_h is set so that it is greater than the scale radius of the PSF in configuration space so that the deconvolution does not amplify the noise on small scales, or large $|k|$. In this work, we do not adapt σ_h to the size of the galaxy light profile. However, in real observations, we can split the survey into smaller patches and have a different smoothing scale σ_h for each patch.

When a galaxy image is distorted under the influence of shear, a finite number of independent shapelet modes are coupled (separated by $|\Delta n| = 2$ and $|\Delta m| = 2$). This causes the sheared shapelet modes to be a linear combination of a finite number of other shapelet modes. Since the FPFs shapelet modes are computed after deconvolving the galaxy with the PSF model, the PSFs do not bias shear estimation as long as we have a good PSF model at the positions of the galaxies.

2.2.1 Galaxy flux and galaxy size

The zeroth-order shapelet mode, M_{00} , is the value of the central peak of the smoothed image with the smoothing scale of σ_h . We follow Li & Mandelbaum (2023) to use M_{00} to quantify the brightness of galaxies, and FPFs flux is defined as

$$F = \frac{M_{00}}{\iint d^2k |\tilde{\chi}_{00}(\mathbf{k})|^2 / |p(\mathbf{k})|^2}, \quad (10)$$

where the denominator is the square of the L^2 norm of the resmoothing kernel. We can use the flux and convert it to FPFs magnitude

defined as

$$m_F = m_{\text{zero}} - 2.5 \log(F), \quad (11)$$

where m_{zero} is the zero point of the survey. The value of m_{zero} is 27 for Hyper Suprime-Cam (HSC) coadded images (Bosch et al. 2018) and 30 for LSST coadded images (Ivezić et al. 2019).

The galaxy size is conventionally measured using second-order Gaussian weighted moments (see e.g. Hirata & Seljak 2003). Following Li & Mandelbaum (2023), we use the combination of spin-0 shapelet moments to define the size as

$$\iint d^2x f \left(\frac{|\mathbf{x}|}{\sigma_h} \right)^2 e^{-|\mathbf{x}|^2/2\sigma_h^2} = M_{00} + M_{20}, \quad (12)$$

and define the FPFs resolution as

$$R_2 = \frac{M_{00} + M_{20}}{M_{00}}. \quad (13)$$

Note that this definition of resolution given by Li & Mandelbaum (2023) is conceptually similar to the resolution defined in Hirata & Seljak (2003). Under a shear distortion, the shapelet moments M_{00} and M_{20} change from their intrinsic values \bar{M}_{00} and \bar{M}_{20} as

$$M_{00} = \bar{M}_{00} + \sqrt{2}(\gamma_1 \bar{M}_{22c} + \gamma_2 \bar{M}_{22s}), \\ M_{20} = \bar{M}_{20} + \sqrt{6}(\gamma_1 \bar{M}_{42c} + \gamma_2 \bar{M}_{42s}), \quad (14)$$

where we use M_{nm} and $M_{\text{nm}s}$ to denote the real and imaginary part of the complex shapelet mode M_{nm} . We use their linear shear response to derive the shear response of the galaxy detection/selection and to correct for the detection/selection bias in Section 2.3.

2.3 Detection and selection weights

We use the weight functions introduced in Li & Mandelbaum (2023) for galaxy detection and selection. For the galaxy detection process, we apply cuts on peak modes, and for the galaxy selection process, we apply cuts on magnitude and resolution. The analytical shear responses of the hard cuts are noisy and unstable, especially when applying multiple cuts to the galaxy properties, as a hard selection weight is discontinuous and not differentiable at the selection boundary. For this work, instead of applying hard cuts on observables, we use truncated sine functions (see equation 45 of Li & Mandelbaum 2023) since these are differentiable and are more stable. The selection weight used to select the galaxy sample is given by

$$w(\mathbf{v}) = T_0^{\text{sel}}(M_{00}) T_2^{\text{sel}}(M_{00} + M_{20}) \prod_{i=0}^3 T^{\text{det}}(v_i), \quad (15)$$

where T_0^{sel} is used to select bright galaxies with high-signal-to-noise ratio (SNR), T_2^{sel} is used to select well-resolved large galaxies, and T^{det} is used to define the cut on peak modes. Since we have calculated the shear response of the detection modes (Section 2.1) and the shapelet modes used to quantify the flux and size (Section 2.2), we can calculate the derivative of the detection and selection weight function to shear.

2.4 Fourth-order shear estimator

Li & Mandelbaum (2023) use shapelet modes M_{00} , M_{20} , and M_{22} (hereafter second-order moments) and their shear responses to compute the galaxy flux, size, and shape (ellipticity). In this work, we use the same detection and selection defined in Li & Mandelbaum (2023) and Li et al. (2024a) and extend the second-order shapes adopted by Li & Mandelbaum (2023) to fourth-order shapes.

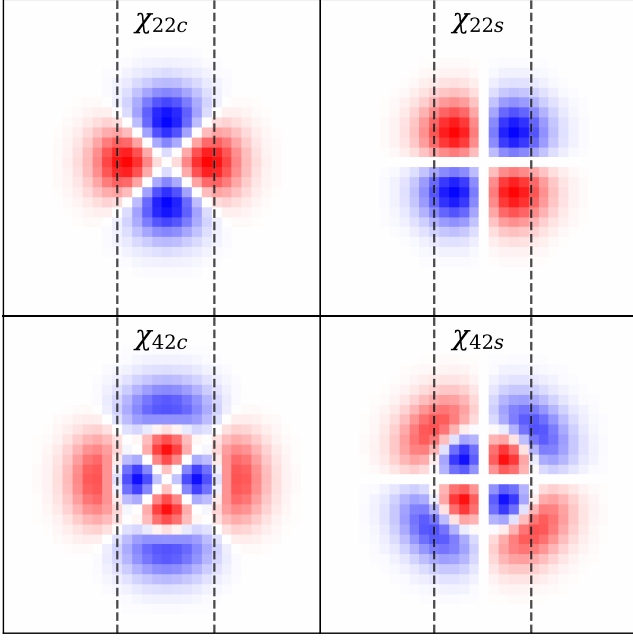


Figure 1. The real and imaginary components of the spin-2 second-order and fourth-order shapelet basis. The fourth-order shapelets are sensitive to scales larger and smaller than that of the second moments, as referenced by the dashed lines. The colour scale assigned to each basis function spans the interval $[-A, A]$, with A representing the maximum absolute value of the corresponding basis function.

In Fig. 1, we show the spin-2 second-moment and fourth-moment basis functions. The fourth-moment basis is more sensitive to pixels with radius both larger and smaller than that of the second moment. Within the same polar angle, the sensitivities of fourth moments to smaller and larger radii have opposite signs, making M_{42} sensitive to the difference in spin-2 behaviour between pixels with small and large radii. Quantitatively, a shapelet model contains information mainly between the minimum and maximum scales (Massey & Refregier 2005) defined as

$$\theta_{\min} = \frac{\sigma_h}{\sqrt{n+1}} \quad \text{and} \quad \theta_{\max} = \sigma_h \sqrt{n+1}, \quad (16)$$

where σ_h is the shapelets smoothing scale and n is the shapelets order. We refer readers to Appendix A for a detailed analysis of the shapelet moments of a Gaussian profile and their sensitivity to shear.

Using the shear responses of shapelet modes defined in Massey & Refregier (2005), the shapelet modes are given by

$$M_{42c} = \bar{M}_{42c} + \frac{\sqrt{6}}{2} \gamma_1 (M_{20} - M_{60}) - \sqrt{5} \gamma_1 M_{64c} - \gamma_2 \sqrt{5} M_{64s} \quad (17)$$

$$M_{42s} = \bar{M}_{42s} + \frac{\sqrt{6}}{2} \gamma_2 (M_{20} - M_{60}) + \sqrt{5} \gamma_2 M_{64c} - \gamma_1 \sqrt{5} M_{64s}, \quad (18)$$

where M_{nm} represents the sheared shapelet modes and \bar{M}_{nm} represents the intrinsic shapelet modes in the absence of shear. As described in Li et al. (2018), shear can be inferred from equations (17) and (18) by taking the expectation values on both sides. Assuming that there is no preferential direction of the randomly selected galaxy ensemble, the expectation values of the intrinsic spin-2 and spin-4 shapelet modes on the right-hand side of the equations reduce to zero, i.e. $\langle \bar{M}_{42} \rangle = \langle \bar{M}_{64} \rangle = 0$. It is the population variance of these

spin-2 and spin-4 shapelet modes that causes the shape noise of the shear estimator.

Li et al. (2018) introduced a normalizing scheme to re-weight the shapelet modes dominated by bright galaxies and reduce the shape noise of the shear estimation. The dimensionless FPFs fourth-order ellipticity is defined as

$$e_1 + ie_2 = \frac{M_{42}}{M_{00} + C^{(4)}}, \quad (19)$$

In Li & Mandelbaum (2023), they use M_{22} in the numerator of equation (19) to define the spin-2 ellipticity. In this work, we use M_{42} to define the fourth-order spin-2 ellipticity. The FPFs weighting parameter $C^{(4)}$ adjusts the relative weight for galaxies with different brightness. Note that the optimal value of the weighting parameter $C^{(n)}$ is different for each order of the estimator, as different shapelet moments are sensitive to different scales.

2.5 Analytical shear calibration

In this subsection, we outline how to correct for the noise bias in shear due to the pixel noise in the images. The previous generation of AnaCa1 corrected the noise bias by computing the second- and higher order derivatives of the non-linear observables (Li et al. 2024b). This required taking derivatives of smoothstep functions with smoothness parameters, which led to significant fluctuations in higher order derivatives Li et al. (2024a). Sheldon & Huff (2017) proposed a numerical recipe to correct for noise bias by adding additional noise to the already-noisy image. Li et al. (2024a) analytically prove the method is free of noise bias and adopts the analytical version within the AnaCa1 framework. This paper follows Li et al. (2024a) to analytically correct for noise bias in shear estimation. Following Sheldon & Huff (2017), we introduce an additional layer of noise to the image that shares the same statistical properties after being rotated counterclockwise by 90° , with the rotation defined in the space prior to PSF convolution. The addition of this noise layer effectively cancels out the spin-2 anisotropies present in the original noise image after deconvolution. Note that Sheldon & Huff (2017) is a ‘finite-difference’ version of the noise bias correction which is not the same as the ‘analytical’ version in Li et al. (2024a). This paper uses the ‘analytical’ version, not the ‘finite difference’ version. The shear response of the renoized ellipticity can be measured after adding the simulated noise as (adopting Einstein notation)

$$\langle \tilde{\tilde{R}}_\alpha \rangle = \left\langle \frac{\partial (\tilde{\tilde{w}} \tilde{\tilde{e}}_\alpha)}{\partial \tilde{\tilde{v}}_i} (\tilde{\tilde{v}}_{i,\alpha} - 2\delta v'_{i,\alpha}) \right\rangle, \quad (20)$$

where the subscript α denotes each component of the shear and $\delta v'_i$ is the measurement error of the i th linear observable from the additional noise. We use a double tilde ($\tilde{\tilde{v}}$) to denote linear observables with doubled image noise. The shear estimator is then

$$\hat{\gamma}_\alpha = \frac{\langle \tilde{\tilde{w}} \tilde{\tilde{e}}_\alpha \rangle}{\langle \tilde{\tilde{R}}_\alpha \rangle} + \mathcal{O}(\gamma^3), \quad (21)$$

where $\tilde{\tilde{e}}$ is the ellipticity observed after adding the additional image noise. This shear estimator is free from noise bias and is accurate to second order of shear. It is worth noting that the *renoizing* approach does not require any computation of noisy second- and high-order derivatives, and does not include higher order terms from image noise which are present in the original version of FPFs. However, a limitation of this method is that we need to double the image noise

before running detection and source measurement. In this work, we used this *renoizing* approach to analytically correct for the noise bias.

2.6 Combining shear estimators

Given the second- and fourth-order shear estimators, we combine these estimators to maximize the SNR of the shear estimator and minimize the noise in the shear estimation, including intrinsic shape noise and measurement error from image noise. The combined shear estimate is defined as

$$\hat{\gamma}_{\alpha,t} = \mu \hat{\gamma}_{\alpha,2} + (1 - \mu) \hat{\gamma}_{\alpha,4}, \quad (22)$$

where α is the component of the shear and the subscripts 2 and 4 are the order of the shear estimator. The statistical uncertainty in the combined shear estimation is then

$$\begin{aligned} \text{Var}(\hat{\gamma}_{\alpha,t}) &= \mu^2 \sigma_{\gamma_{\alpha,2}}^2 + (1 - \mu)^2 \sigma_{\gamma_{\alpha,4}}^2 \\ &\quad + 2\mu(1 - \mu) \text{Cov}(\hat{\gamma}_{\alpha,2}, \hat{\gamma}_{\alpha,4}), \end{aligned} \quad (23)$$

where $\sigma_{\gamma_{\alpha}}$ are the standard deviation of each estimator, and ρ_{α} is the correlation between the two estimators. The shear estimator weight μ is defined so that it minimizes the variance of the combined estimate as

$$\begin{aligned} \mu &= \arg \min_{\mu} \text{Var}(\hat{\gamma}_{\alpha,t}) \\ &= \frac{\sigma_{\gamma_{\alpha,4}}^2 - \text{Cov}(\hat{\gamma}_{\alpha,2}, \hat{\gamma}_{\alpha,4})}{\sigma_{\gamma_{\alpha,2}}^2 + \sigma_{\gamma_{\alpha,4}}^2 - 2 \text{Cov}(\hat{\gamma}_{\alpha,2}, \hat{\gamma}_{\alpha,4})} \\ &= \frac{\sigma_{\gamma_{\alpha,4}}^2 - \rho_{\alpha} \sigma_{\gamma_{\alpha,2}} \sigma_{\gamma_{\alpha,4}}}{\sigma_{\gamma_{\alpha,2}}^2 + \sigma_{\gamma_{\alpha,4}}^2 - 2 \rho_{\alpha} \sigma_{\gamma_{\alpha,2}} \sigma_{\gamma_{\alpha,4}}}. \end{aligned} \quad (24)$$

In the case of no correlation between the two estimators, $\rho_{\alpha} = 0$ and the weight μ is equivalent to inverse variance weighting. If both estimators give an unbiased shear estimate, i.e. $\langle \hat{\gamma}_{\alpha,2} \rangle = \langle \hat{\gamma}_{\alpha,4} \rangle = \gamma_{\alpha}$ when averaged over an ensemble of galaxies, then the combined estimator also gives an unbiased shear estimate.

3 SIMULATIONS AND TESTS

In this section, we describe mock astronomical images that we produce with known input shears to test the performance of our fourth-order shear estimator and compare it against the second-order shear estimator developed in Li & Mandelbaum (2023). We corrected for noise bias, selection bias, and detection bias using methods developed in Section 2. The estimated shear, $\hat{\gamma}_{\alpha}$, is related to the input shear, γ_{α} , as

$$\hat{\gamma}_{\alpha} = (1 + m_{\alpha})\gamma_{\alpha} + c_{\alpha}, \quad (25)$$

where the subscript α denotes each component of the shear and m_{α} (multiplicative bias) and c_{α} (additive bias) are used to test the accuracy of the shear estimator (Heymans et al. 2006; Huterer et al. 2006). We also combine second- and fourth-order shear estimators as derived in equation (22) and study the reduction in the shape noise.

3.1 Simulation setup

We use the same simulation setup as used in Li & Mandelbaum (2023). In summary, we generate two sets of data: image simulations with isolated galaxies, where each realization contains multiple postage stamps that each have one galaxy, and blended galaxy image simulations, where galaxies are randomly distributed. We vary the seeing, pixel scale, and image noise level to match those of the HSC and LSST surveys.

Table 1. Table for simulation setups. The noise variance for each simulation is set to the mean noise variance in the 3-yr HSC data and the 10-yr LSST observation, respectively.

Variable	HSC	LSST
Pixel scale	0.168 arcsec	0.20 arcsec
Seeing	0.60 arcsec	0.80 arcsec
Moffat profile exponent	3.5	2.5
Magnitude zero point	27.0 mag	30.0 mag
Noise variance	3.6×10^{-2}	3.5×10^{-1}
PSF ellipticity	$(e_1 = 0.02, e_2 = -0.02)$	

The pixel scale is set to 0.168 arcsec (0.2 arcsec) for HSC (LSST) like simulation. The PSF for these simulations is modelled with a Moffat (1969) profile,

$$p_m(\mathbf{x}) = \left[1 + c \left(\frac{|\mathbf{x}|}{r_p} \right)^2 \right]^{-n}, \quad (26)$$

where c and r_p are set such that the full width half-maximum (FWHM) of the Moffat PSF is 0.60 arcsec (0.80 arcsec) matching the median seeing of actual HSC *i*-band (expected LSST *i*-band) images. The exponent n is 3.5 for HSC simulations (Li et al. 2022a) and 2.5 for LSST simulations (Sheldon et al. 2023). Table 1 summarizes the values used for each survey setup. To introduce an anisotropic PSF, we shear the PSF so that it has ellipticity ($e_1 = 0.02, e_2 = -0.02$). We show a small region of one simulated image for both isolated and blended simulations in Fig. 2.

When testing the accuracy of our fourth-order shear estimator, we use the same simulation used in Li & Mandelbaum (2023): for each type of simulation, we have two sets of galaxy images with the same realization of image noise generated; one distorted with ($\gamma_1 = 0.02, \gamma_2 = 0$) and the other with ($\gamma_1 = -0.02, \gamma_2 = 0$) as introduced in Pujol et al. (2019). We used the ring test setup (Massey et al. 2007) by having our galaxy sample contain orthogonal galaxies with the same morphology and brightness but with the intrinsic major axes rotated by 90°. This allows us to cancel out the intrinsic shape noise in our simulation.

3.2 Isolated image simulations

We test the performance of fourth-order shear estimator relative to the second-order estimator using isolated galaxy image simulations. The goal of this is to understand how the estimator performs when blurring effects are absent, focusing on the impact of shear, PSF size, and image noise. For isolated galaxy image simulations, galaxy images are generated using the publicly available software *GalSim* (Rowe, Jarvis & Mandelbaum 2014) and use the COSMOS HST parametric galaxy catalogue (Mandelbaum et al. 2019) with limiting magnitude $F814W = 25.2$ as the input galaxy catalogue (see section 3.1 of Li & Mandelbaum 2023, for more details). Note that at this limiting magnitude for the input catalogue, some of the lower SNR galaxies that are expected in the isolated simulations may be missing from the analysis. Galaxies are rendered in 64×64 pix postage stamp images after the shear distortion and PSF convolution. Each realization (subfield) of the simulated image contains 100×100 postage stamps, and after including orthogonal galaxy pairs as described in Section 3.1, each subfield contains 5×10^3 orthogonal galaxy pairs. Each orthogonal galaxy is placed adjacent to its unrotated galaxy in the same simulation. We simulate 4000 subfields with different realizations for the noise image, the galaxy sample, and random galaxy rotation.

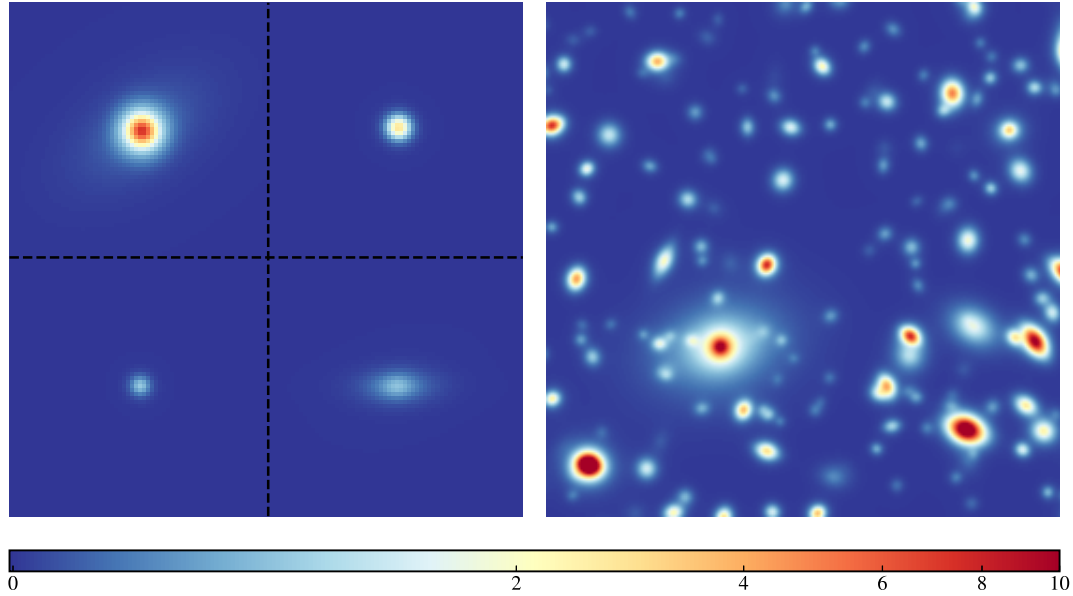


Figure 2. The left panel shows a $128 \times 128 \text{ pix}^2$ (equivalent to $0.36 \times 0.36 \text{ arcmin}^2$) stamp image of the isolated galaxy image simulation with HSC seeing, where the dotted black lines show the boundaries of the $64 \times 64 \text{ pix}^2$ stamps. The right panel shows a random cut-out coadded image of $240 \times 240 \text{ pixel}^2$ (equivalent to $0.8 \times 0.8 \text{ arcmin}^2$) of *griz* bands of the LSST-like blended galaxy image simulation (Sheldon et al. 2023). Both images were produced before adding noise.

3.3 Blended image simulations

Blended galaxy simulations are essential for understanding the estimator’s performance in realistic observational conditions, where the blending of multiple objects occurs. We also tested our fourth-order estimator on blended galaxy image simulations and compared it against the second-order estimator. For blended galaxy image simulations, we use the open-source package *descwl-shear-sims*¹ (Sheldon et al. 2023), which includes the survey parameters for the simulation, such as noise and PSF. The package uses the input galaxy catalogue from *WeakLensingDeblending*.² The output of the package is a calibrated image with a subtracted background and estimated noise variance of the image. The simulated galaxies include bulge, disc, and active galactic nucleus (AGN) components. The bulge and the disc have varying fluxes and the AGN is represented as a point source located at the galaxy centre. To include some of the complexity of realistic galaxy light profiles, the isophotes of these simulated galaxies are not strictly elliptical. Galaxy positions are randomly distributed across the image, without accounting for spatial clustering.

Each simulated image contains about 10^5 input galaxies and covers 0.12 deg^2 , corresponding to a galaxy number density of about 230 per square arcmin. However, not all of these galaxies are detectable at the depths achieved by HSC and LSST. When testing the accuracy of each shear estimator on blended galaxies, we also adopted the ring test to cancel out shape noise. For these blended simulations, the rotated galaxy is stored in a different realization. The blended galaxy image simulations are simulated with *descwl-shear-sims* (Sheldon et al. 2023), and the setup is the same as Li et al. (2024b).

¹<https://github.com/LSSTDESC/descwl-shear-sims>

²<https://github.com/LSSTDESC/WeakLensingDeblending>

For LSST-like simulations, we use *griz* bands with the same galaxy profile and PSFs in each band without dithering. The noise variance level matches the expectations for a coadded image based on LSST 10-yr observations, and each band has a different noise variance. The image is coadded in these four bands with inverse variance weighting based on the sky background noise. The multiband coadds, then, have a well-defined PSF (Mandelbaum et al. 2023) since each band has a spatially constant weight across the images. These four bands were chosen as it is likely that LSST cosmic shear analysis will use observations in *griz* bands. For the HSC-like setup, we used the image noise variance in the HSC *i* band. HSC primarily uses the *i*-band for cosmic shear due to its better seeing conditions compared to other bands.

4 RESULTS

In this section, we present the results of testing the precision and accuracy of our fourth-order FPFS shear estimator with isolated galaxies (Section 4.1) and blended galaxies (Section 4.2) as described in Section 3. We also compare the results against the second-order FPFS shear estimator and combine the two shear estimators to quantify the reduction in shape noise.

4.1 Isolated galaxies

In this subsection, we focus on isolated galaxies simulated within postage stamps as described in Section 3.2.

4.1.1 Precision test

We first tested the performance of our fourth-order estimator and compared the uncertainty on this new estimator to that of the second-order estimator. We then computed how much information is gained, or how much uncertainty is reduced, by combining the

two estimators. When analysing the uncertainty of each estimator, we use 100 out of 4000 subfields to save computational time. For each subfield, we include the intrinsic shape noise in our analysis by excluding a single randomly chosen galaxy from each orthogonal pair when both are detected. In addition, we deliberately positioned the centre of each galaxy at the centre of the postage stamp. We forced the pipeline to do measurements based on the known centre without running any detection and selection process during the image processing. This is equivalent to setting the selection weight function, w in equation (15), to 1 for all galaxies. We use the results from this process to set the FPFS weighting parameter $C^{(n)}$ in the denominator of the spin-2 ellipticity estimator in equation (19).

For each subfield i , the estimated shear is computed as

$$\hat{\gamma}_{\alpha,i}^{(n)} = \frac{\sum_{j \in i} \tilde{w}_j^{(n)} \tilde{e}_{\alpha,j}^{(n)}}{\sum_{j \in i} \tilde{R}_{\alpha,j}^{(n)}}, \quad (27)$$

where $n \in \{2, 4\}$ is of the order of the shear estimator, α is the component of the shear, and the summation is over all galaxies j in subfield i . Note that the weight is also order-dependent as the weight parameter $C^{(n)}$ is different for each order. Forcing the measurements at the known centres and not applying any selection criteria is equivalent to setting the effective weight in the shear estimator $\tilde{w}_j^{(n)} = 1$.

We calculate the statistical uncertainty in the shear estimation from a population of subfields of galaxies as

$$(\sigma_{\gamma}^{(n)})^2 = \left\langle \frac{1}{2} \left((\hat{\gamma}_1^{(n)})^2 + (\hat{\gamma}_2^{(n)})^2 \right) \right\rangle N, \quad (28)$$

where N is the number of galaxies in each subfield. With each subfield having 100×100 postage stamps, but with one galaxy in each orthogonal pair excluded, $N = 5 \times 10^3$. The total uncertainty is a combination of uncertainty due to galaxies' intrinsic shape noise, $\gamma_{\text{RMS}}^{(n)}$, and uncertainty due to noise in the galaxy images, $\sigma_e^{(n)}$. We assume that they add in quadrature, such that

$$(\sigma_{\gamma}^{(n)})^2 = (\gamma_{\text{RMS}}^{(n)})^2 + (\sigma_e^{(n)})^2. \quad (29)$$

To obtain the galaxies' intrinsic shape noise, $\gamma_{\text{RMS}}^{(n)}$, we use simulated galaxies with zero image noise and the same weighing parameter value $C^{(n)}$ for each estimator to measure the FPFS ellipticity. To save computational time, we only measure the first component of the ellipticity and calculate the RMS across our sample of subfields. To obtain the measurement uncertainty, $\sigma_e^{(n)}$, we measure the total uncertainty using a noisy realization of the galaxy image, again using only the first component of the ellipticity. We then subtract e_{RMS} in quadrature following equation (29) to get the measurement uncertainty for each shear estimator.

The statistical uncertainty of FPFS also depends on the FPFS weighting parameter, $C^{(n)}$, in equation (19) (Li & Mandelbaum 2023). Here, we assume image noise fields that are homogeneous but that have pixel-to-pixel correlations. This is a reasonable approximation for faint, small galaxies observed in ground-based surveys, where noise is primarily due to sky background fluctuations and where measurements may be made on a coadd with pixel-to-pixel noise correlations. However, this assumption may not fully capture the noise properties in upcoming Stage IV imaging surveys. Further testing is required to evaluate the robustness of this assumption to achieve the precision needed for percent-level cosmology. In Fig. 3, we show each source of uncertainty as a function of $C^{(n)}$ for the second-order (blue) and fourth-order (orange) estimators on simulated isolated galaxies.

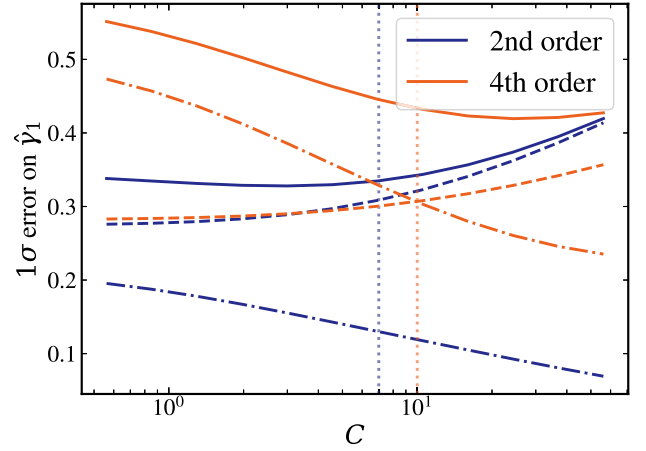


Figure 3. The 1σ statistical uncertainty on a single component of the estimated shear $\hat{\gamma}_1$ for individual isolated galaxies (solid lines) as a function of the weighting parameter, $C^{(n)}$, in equation (19). For each second and fourth order estimator, the total uncertainty has contributions due to image noise (dotted lines) and intrinsic shape noise (dash-dotted lines). The vertical dashed lines show the values of $C^{(n)}$ used for each estimator throughout the paper.

The galaxies' intrinsic shape noise (dashed) for both estimators increases as a function of $C^{(n)}$, whereas the measurement uncertainty (dash-dotted) for both decreases as a function of $C^{(n)}$ since a larger value of $C^{(n)}$ adds more weight to the limited number of bright galaxies with higher SNR. The measurement uncertainty for the fourth-order estimator is larger than that of the second-order for our simulation setup as the fourth-order shapelet basis amplifies the noise. From this result, we set $C^{(2)} = 7$ for the second-order shear estimator and $C^{(4)} = 10$ for the fourth-order shear estimator for the rest of the paper.

Once we fixed the value of $C^{(n)}$ for each shear estimator, we relaxed the forced measurement and allowed the pipeline to carry out detection and selection during the image processing. For the rest of this section, during the image production step and before running the FPFS detection process, we shift the galaxy centroid with random sub-pixel offsets. Following Li & Mandelbaum (2023), we run the FPFS peak detection algorithm and use the detected peaks as the centroid. Using the detected peaks to measure the galaxy properties, we then apply flux- and resolution-based selection criteria in equation (15) to assign selection weights to the detected galaxy sample. After analytically correcting for noise bias, selection bias, and detection bias, we show the shape measurement uncertainty from the second-order (blue) and fourth-order (orange) estimator on isolated galaxies in Fig. 4. This figure shows that for isolated galaxy samples, combining the two shear estimators reduces the statistical uncertainty in shear inference by ~ 30 per cent. This result suggests that for isolated galaxies, we can get significantly more information if we estimate the shear using second and fourth-order moments and combine them using equation (23). A 30 per cent reduction in shape noise is equivalent to the increase in sample size that would be achieved by expanding the survey area by 70 per cent.

4.1.2 Adaptive versus non-adaptive moments

We next compare each FPFS shear estimator against the widely used reGAUSS shear estimator method (Hirata & Seljak 2003). This comparison is motivated by the differences in how these methods

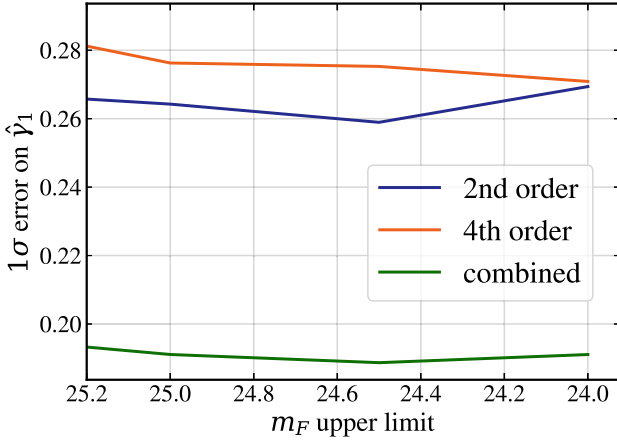


Figure 4. The 1σ statistical uncertainty on shear measurement $\hat{\gamma}_1$ for individual isolated galaxies as a function of the upper limit of FPFs magnitude (m_F) for the galaxies included in the measurement. The uncertainty for each second and fourth order estimator includes contributions due to image noise and intrinsic shape noise. Equation (22) is used to combine the two shear measurements.

use moments of the galaxy profiles to measure ellipticity. One key distinction between the two is the use of adaptive versus non-adaptive approaches. In the non-adaptive method, as implemented in FPFs, a fixed kernel (or weight function) is applied uniformly across all galaxies, regardless of their morphology. The adaptive method used by reGauss optimizes the weight function for each galaxy based on its specific shape and size. This optimization improves the precision of the shear estimation by effectively accounting for variations in galaxy morphology to match the weight functions to the galaxy images. We use the GALSIM implementation of reGauss to correct for the PSF when estimating galaxy shapes. An important output of the reGauss estimator is the components of the spin-2 ellipticity of each galaxy:

$$(e_1, e_2) = \frac{1 - (b/a)^2}{1 + (b/a)^2} (\cos 2\phi, \sin 2\phi), \quad (30)$$

where b/a is the axis ratio and ϕ is the position angle of the major axis with respect to sky coordinates. Another important output of reGauss is the resolution factor R_2 :

$$R_2 = 1 - \frac{T_{\text{PSF}}}{T_{\text{gal}}}, \quad (31)$$

where T_{PSF} is the trace of the second moment matrix of the PSF and T_{gal} is the trace of the second moment matrix of the observed PSF-convolved galaxy image. The resolution factor quantifies how well the galaxy image is resolved relative to the PSF; $R_2 \sim 1$ for well-resolved objects and $R_2 \sim 0$ for poorly resolved images. In general, the positively and negatively sheared galaxy images have different resolution factors ($R_2^+ \neq R_2^-$).

We select well-resolved galaxies to avoid a large multiplicative bias correction at a low-resolution factor for reGauss (Li et al. 2022a). To ensure a consistent sample between FPFs and reGauss, we apply the following criteria for each subfield: only keep galaxies that satisfy $R_2^{(+,-)} > 0.3$ and $|e^{(+,-)}| < 2$. Applying the resolution factor and the ellipticity criterion reduces the sample to $\sim 4 \times 10^3$ galaxies out of 5×10^3 galaxies in each subfield. When comparing the two shear estimation methods, we only retain galaxies that were detected in both methods, ensuring the same sample population for

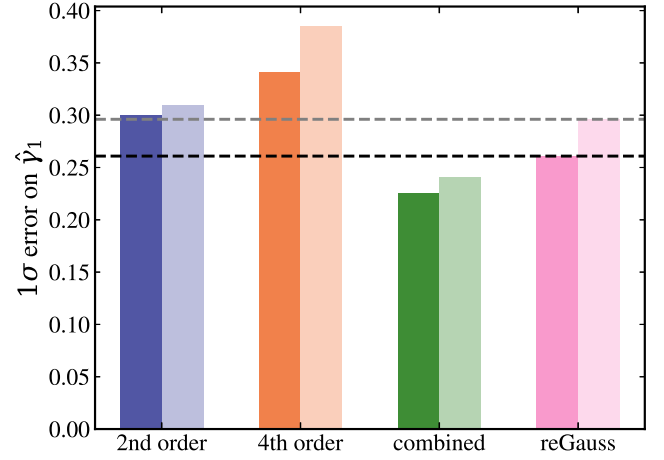


Figure 5. The 1σ statistical uncertainty on shear measurement of individual galaxies from each estimator; second and fourth order, combining the two FPFs estimators, and from reGauss method. The uncertainty values of reGauss are drawn in dotted horizontal lines to simplify comparison. For each estimator, the left bar (deep) only has contributions from the intrinsic shape noise and the right bar (shallow) has contributions from intrinsic shape noise and image noise.

both FPFs and reGauss. This helps isolate differences in precision between the methods without introducing sample bias.

For an isotropically oriented galaxy ensemble distorted by some known constant shear, the shear responsivity can be estimated:

$$\hat{R}^{(R)} = \frac{\langle \widehat{w^{(R)} e_1^{(R)}}^+ - \widehat{w^{(R)} e_1^{(R)}}^- \rangle}{2\gamma}, \quad (32)$$

where the superscript ‘+’ (‘-’) refers to measurements of galaxies with positive (negative) shear γ applied, R denotes quantities estimated using reGauss method, and $w^{(R)}$ is the inverse variance weight defined as

$$w^{(R)} = \frac{1}{(\gamma_{\text{RMS}}^{(R)})^2 + (\sigma_\gamma^{(R)})^2}, \quad (33)$$

Using the reGauss method, the shear is then estimated as

$$\hat{\gamma}_{\alpha,i}^{(R)} = \frac{\sum_{j \in i} w_j^{(R)} e_{\alpha,j}^{(R)}}{R_{\alpha,j}^{(R)}}, \quad (34)$$

To obtain $\gamma_{\text{RMS}}^{(R)}$, we use the noise-free isolated galaxy image simulation and set the weights $w = 1$. We use equation (34) to estimate the shear and take the variance across multiple subfields to get the uncertainty. This process gives $\gamma_{\text{RMS}}^{(R)} = 0.26$, the typically used value for intrinsic shape noise. To obtain the total uncertainty including intrinsic shape noise and measurement error, we use noisy realizations of the images. We set $\gamma_{\text{RMS}}^{(R)} = 0.26$ and use the value $\sigma_\gamma^{(R)}$ returned from GALSim.

In Fig. 5, we compare the 1σ statistical uncertainty on the shear measurement of individual galaxies from different shear estimation methods for 100 subfields. For each shear estimation method, we present the intrinsic shape noise (shaded) and the total uncertainty (no shade), which includes intrinsic shape noise and measurement error. We use $C^{(2)} = 7$ and $C^{(4)} = 10$ for the second-order and fourth-order AnaCal estimator, respectively, which optimizes the total uncertainty in Fig. 3. Equation (22) is used to combine the two AnaCal estimators and equation (23) to quantify the uncertainty of the combined estimator. The dashed horizontal lines are drawn to

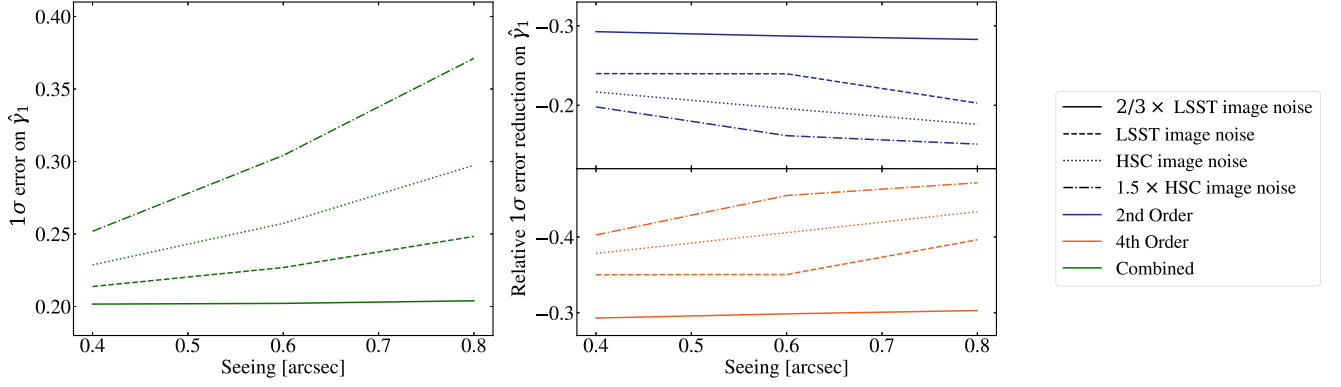


Figure 6. The 1σ statistical uncertainty per galaxy on shear measurement $\hat{\gamma}_1$ for isolated galaxies as a function of the PSF size using different image noise levels and shear estimators. The uncertainty for each second and fourth order estimator includes contributions due to image noise and intrinsic shape noise. Equation (22) is used to combine the two shear measurements. Left panel shows the 1σ error on $\hat{\gamma}_1$ using the combined shear estimator. Right panel shows the ratio of statistical uncertainty from the second-order and fourth-order estimators compared to the combined estimator. The noise variance for HSC is approximately twice that of the LSST noise variance. The results indicate that the combined shear estimator consistently provides lower statistical uncertainty relative to individual shear estimators across varying seeing size and image noise levels.

indicate the values for reGauss. In general, reGauss performs better than individual FPFs estimators, or in other words the adaptive moments method has lower uncertainty than the non-adaptive moments method. However, when combining the second- and fourth-order FPFs estimators, we get ~ 30 per cent reduction in the shape noise compared to that of reGauss.

4.1.3 Varying seeing and image noise

In the previous subsections, we used the HSC PSF size (0.60 arcsec) and the average noise in the i band for isolated image simulations, and found an ~ 30 per cent reduction in shape noise when including the fourth moment-based shear estimator along with the second moment-based estimator. In this subsection, we vary the PSF size and image noise to see test the dependence of this finding on the observational conditions. In Fig. 6, we show the uncertainty in each shear estimator and combined for four different PSF seeing sizes and image noise levels. ‘ $1.5 \times \text{HSC}$ ’ corresponds to an image noise level with a noise variance that is 1.5 times the average noise variance of HSC in i -band coadds, and ‘ $2/3 \times \text{LSST}$ ’ corresponds to the noise variance that is $2/3$ of the expected noise variance from 10 yr of LSST coadds. The reduction in shape noise is consistent for all different seeing sizes and image noise levels. The curve for ‘ $2/3 \times \text{LSST}$ ’ is basically flat because for a low image noise level, we can perfectly deconvolve the image with a PSF without amplifying noise on small scales. As illustrated in the right panel of Fig. 6, the ratio of uncertainty from second/fourth-order shear estimators to that of combined is roughly independent of PSF size and image noise level, and the fourth-order estimator exhibits a greater susceptibility to image noise than the second-order estimator. This suggests that for isolated galaxies, combining the two independent FPFs shear estimators consistently reduces statistical uncertainty, making it the preferred approach.

4.1.4 Accuracy test

In this subsection, we test how accurate our fourth-order estimator is compared to the second-order estimator. The multiplicative and additive biases in our shear estimation in equation (25) are measured

as introduced in Sheldon et al. (2020) as

$$m_\alpha = \frac{\langle \widehat{we}_\alpha^+ - \widehat{we}_\alpha^- \rangle}{0.02 \langle \widehat{R}_\alpha^+ + \widehat{R}_\alpha^- \rangle} - 1, \quad (35)$$

and

$$c_\alpha = \frac{\langle \widehat{we}_\alpha^+ + \widehat{we}_\alpha^- \rangle}{\langle \widehat{R}_\alpha^+ + \widehat{R}_\alpha^- \rangle}, \quad (36)$$

where 0.02 is the value of constant shear applied to each galaxy image and \widehat{we}_1 and \widehat{R}_1 are the first components of the weighted ellipticity and its shear responsivity after noise bias correction, respectively. When quantifying the multiplicative and additive biases, we test γ_1 and γ_2 separately; when testing γ_1 we set $\gamma_2 = 0$ and vice versa. The quantities with superscript ‘+’ are estimated from images distorted by positive shear, ($\gamma_1 = 0.02$ and $\gamma_2 = 0$) or ($\gamma_1 = 0$ and $\gamma_2 = 0.02$) and superscript ‘-’ are estimated from images distorted by negative shear, ($\gamma_1 = -0.02$ and $\gamma_2 = 0$) or ($\gamma_1 = 0$ and $\gamma_2 = -0.02$). It is worth noting that galaxies in each orthogonal galaxy pair and galaxies with different applied shears (positive and negative γ_1) are selected and weighted independently. We then apply the shear estimator to this selected sample to evaluate our corrections for detection and selection biases.

For the accuracy test, we used the HSC seeing size and image noise level and used 4000 subfields with galaxy orthogonal pairs to approximately cancel out the intrinsic shape noise. The amplitudes of the multiplicative (additive) biases for each FPFs estimator are shown in the top (bottom) panel of Fig. 7. We found that the multiplicative biases for both estimators are well within the LSST 10-yr requirements, and we do not find a significant additive bias.

4.2 Blended galaxies

In this subsection, we discuss the results of the precision test (as shown for isolated galaxies in Section 4.2.1) and the accuracy test (Section 4.2.2) by applying the second- and fourth-order AnaCal shear estimators on the blended simulated galaxy images as described in Section 3.3. In Li et al. (2024b), they demonstrated that the second-order AnaCal shear estimator yields unbiased results for the accuracy test. The results presented here confirm that the second-order shear estimator tested in this section is consistent with these

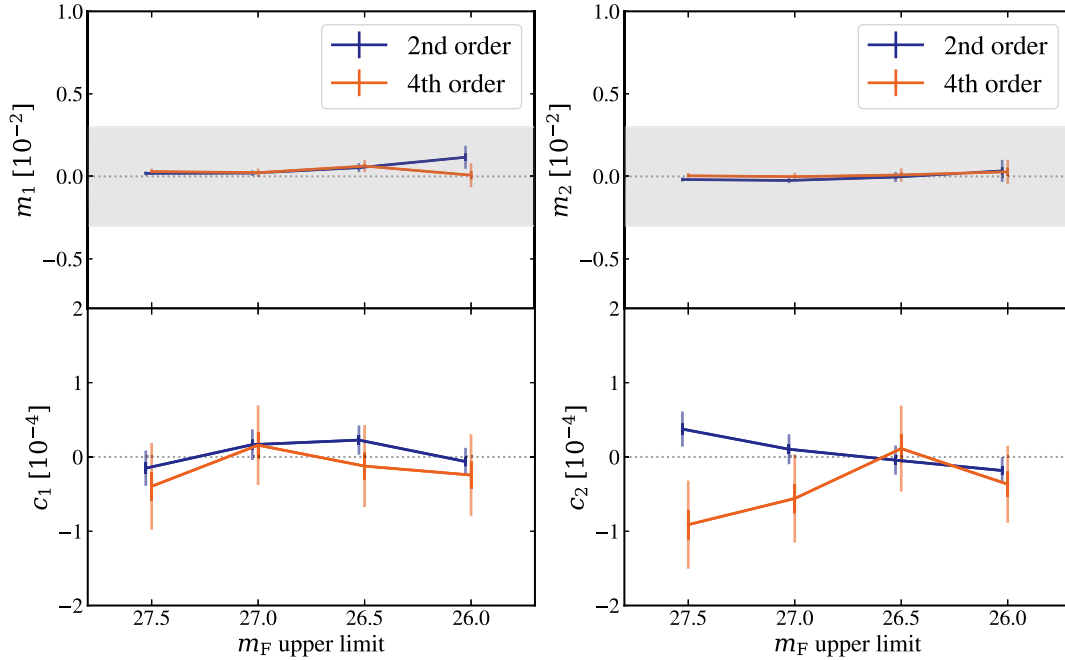


Figure 7. The multiplicative bias (upper panel) and additive bias (lower panel) of the AnaCal shear estimator on isolated galaxies using an HSC-like configuration for each order of the shear estimator. The error bars show the 1σ and 3σ uncertainties, respectively, and the grey-shaded region is the LSST 10-yr requirement on the multiplicative shear bias (The LSST Dark Energy Science Collaboration 2018). In both panels, the blue points are slightly shifted to differentiate the error bars. Both estimators give unbiased shear estimation for each shear component.

previous performance tests of the AnaCal second-order shear estimator.

4.2.1 Precision test

When quantifying the uncertainty on each estimator for noisy simulated blended galaxy images, we use 100 subfields. We measure shear from each image using both the second- and fourth-order FPFs shear estimators and use the scatter in these 100 shear measurements to derive the statistical error. We obtain the per-component statistical uncertainty σ_γ as defined in equation (28) on shear estimation for a region of one square arcmin by normalizing the statistical error according to the area of each simulated image. We follow Li et al. (2024b) to derive the n_{eff} from blended galaxy image simulations as

$$n_{\text{eff}} = \left(\frac{0.26}{\sigma_\gamma} \right)^2 [\text{arcmin}^{-2}], \quad (37)$$

where 0.26 is the per component RMS of intrinsic shape noise using the widely used reGauss shear estimation method (see Fig. 5). We choose to use 0.26 for the shape noise RMS so that we can compare our estimation of the effective number density with other methods that use reGauss. Typically, the galaxy number density is estimated by counting the galaxy number (with weights) after detecting galaxies from images (Chang et al. 2013; The LSST Dark Energy Science Collaboration 2018). Doing so assumes that the statistical uncertainty in shear estimation for each galaxy is uncorrelated. Effectively we compute the equivalent galaxy number density for the shear if it were to be estimated with reGauss.

Motivated by the result of Fig. 6, we show the n_{eff} as a function of SNR cut for each shear estimator for two different survey setups, HSC and LSST, in Fig. 8. Each survey setup follows the respective PSF size and image noise level. The $n_{\text{eff}} \sim 15 \text{ arcmin}^{-2}$ for HSC

setup is 35 per cent smaller than the $n_{\text{eff}} \sim 20 \text{ arcmin}^{-2}$ for a PSF size of 0.6 arcsec from HSC-Y3 (Li et al. 2022a). The value may be smaller because we did not optimize the smoothness parameter for detection. Optimizing this parameter will improve the n_{eff} . The value of $n_{\text{eff}} \sim 39 \text{ arcmin}^{-2}$ for LSST seeing size and image noise level is 5 per cent higher than n_{eff} estimated from Chang et al. (2013). They reported $n_{\text{eff}} \sim 37 \text{ arcmin}^{-2}$ for the expected distributions of observing parameters and all lensing data (r and i bands) before considering blending and masking. The value is estimated in this work is slightly higher due to the use of different selection cuts and different numbers of bands to compute the value; Chang et al. (2013) combined r - and i -bands, whereas we used $griz$ -band coadded images.

We find that for the fourth-order FPFs shear estimator, the effective number density is relatively constant for every SNR lower limit, as the fourth-order amplifies measurement error caused by image noise. This shows that the fourth-order FPFs shear estimator is more sensitive to blending. By combining second- and fourth-order, we find that there is ~ 2 per cent improvement in the effective number density. When estimating shear using blended galaxy samples, the fourth-order shear estimator adds minimal information to the second-order estimator. The sensitivity of the fourth-order estimator to noise and blending highlights its limitations in current ground-based surveys. Space-based images, and deep-field images with higher SNR, offer the potential for improving the effective number density using the second- and fourth-order shear estimators together. We defer this work to future studies.

4.2.2 Accuracy test

We also tested the accuracy of the fourth-order shear estimator on blended galaxies. We used equations (35) and (36) to measure

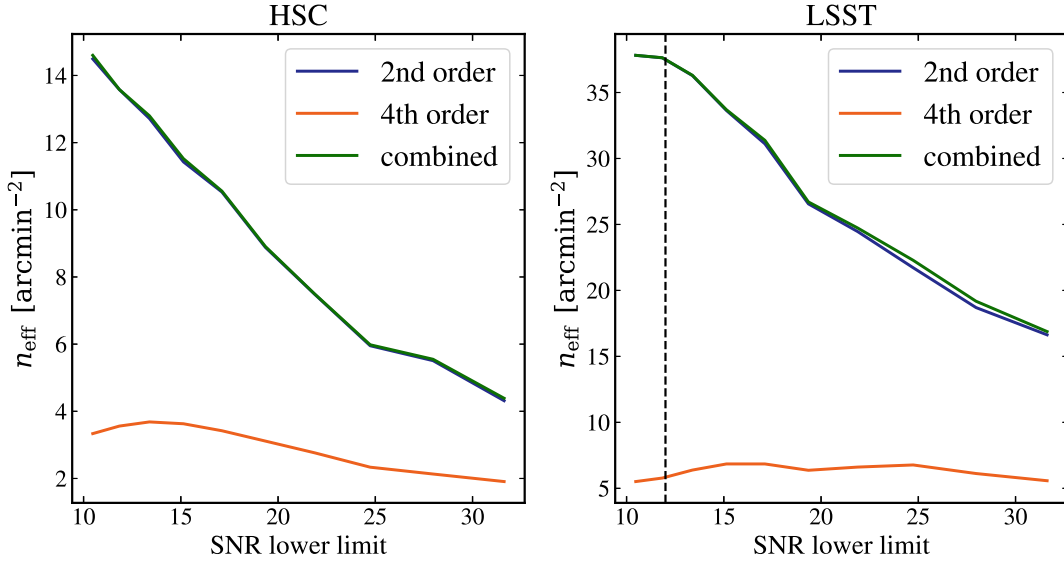


Figure 8. The effective galaxy number density as a function of SNR cut for each AnaCa1 estimator and for each simulation setup (PSF seeing size and image noise level) using 100 subfields with blended galaxies. The statistical uncertainties that go into this effective number density include intrinsic shape noise, image noise, and noise from blending. The vertical dashed line represents the default SNR cut used in this paper.

multiplicative and additive biases in our shear estimator. We include the orthogonal galaxy pairs to approximately cancel out the intrinsic shape noise and use all 5000 subfields to tightly constrain the shear biases. To save computational time, we only test the noise bias correction for the estimation of $\hat{\gamma}_1$ and remove the subscript ($m \equiv m_1$ and $c \equiv c_1$); the multiplicative and additive biases of $\hat{\gamma}_2$ should be comparable to those of $\hat{\gamma}_1$. Fig. 9 shows the multiplicative and additive biases for each AnaCa1 estimator with the LSST 10-yr requirement. The result shows that the fourth-order shear estimator satisfies the LSST 10-yr requirement even in the presence of blending.

5 CONCLUSION

In this work, we extended the perturbation-based shear estimator FPFS, which is part of the new suite of analytic shear estimation algorithms AnaCa1, and developed a higher order FPFS shear estimator. We compared its performance against the second-order shear estimator developed in Li & Mandelbaum (2023). The AnaCa1 framework analytically corrects for the detection and selection bias in shear estimation by deriving the shear response of image pixels. It also uses the *renoizing* approach (Li et al. 2024a) to correct for the noise bias by adding an additional layer of noise to the image. This method does not require any computation of derivatives of noisy non-linear observables. This work focuses on evaluating the accuracy and precision of our new shear estimator on constant-shear simulations, in which all galaxies in isolated and blended image simulations undergo the same shear. The application of this work with redshift-dependent shear will be explored in future research.

Typically, the shear estimation uses the lowest order (specifically, second order) to define spin-2 observables. We used fourth-order FPFS shapelet moments to define our spin-2 observables (e.g. ellipticity) and incorporated the fourth-order shear estimator into the FPFS framework. We tested the precision and accuracy of our new shear estimator with HSC-like and LSST-like simulations on both isolated and blended galaxy image simulations. We find that for both sets of image simulations, the multiplicative shear bias $|m|$ for the fourth-order shear estimator is consistently less than 3×10^{-3}

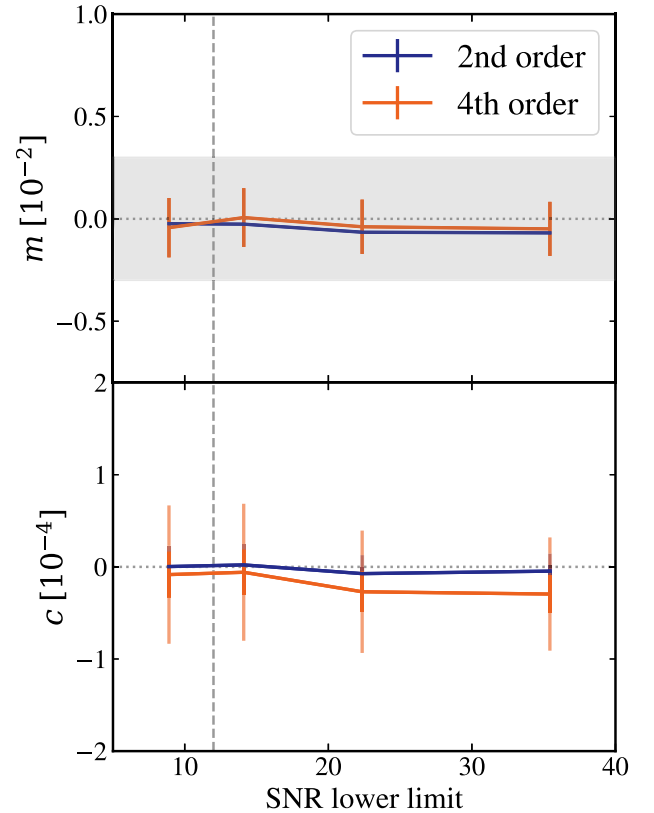


Figure 9. The multiplicative bias (upper panel) and additive bias (lower panel) of the AnaCa1 shear estimator on blended galaxies using an LSST-like configuration for each order of the shear estimator. The error bars show the 1σ and 3σ uncertainties, respectively, and the shaded grey region is the LSST 10-yr requirement on the shear multiplicative bias. The vertical dashed line indicates our default selection SNR > 12 .

(LSST 10-yr requirement on the control of multiplicative bias) within the 3σ uncertainties.

When the fourth-order shear estimator is combined with the second-order shear estimator, we found that the shape noise is reduced by ~ 35 per cent for isolated galaxy image simulations compared to using the second-order estimator alone. For isolated galaxies, the fourth-order shear estimator detects information that is not probed by the second moment, and we found that the shear estimators are complementary. We also found that the effective number density is improved by ~ 2 per cent for blended galaxy image simulation. When blending is introduced, the fourth-order estimator treats the blended galaxies as noise and amplifies their signals causing them to behave like noise. We recommend combining both the second- and fourth-order FPFs shear estimators when estimating shear for high SNR and isolated galaxy ensembles. While this approach shows promise, further testing is required for its application to debledged samples and is left for future work.

An important target for future work is testing this method on spaced-based image simulations, as there may be less blending of the galaxies. In that regime, it would be worth evaluating the improvement in n_{eff} from combining the two orders of shear estimators. This paper focused on evaluating and testing the new shear estimator using constant-shear simulations, in which all isolated and blended galaxies within a single image were subject to uniform shear distortions. MacCrann et al. (2022) showed and tested the performance of METACALIBRATION using simulations with redshift-dependent shear by dividing galaxies into four redshift bins and applying different shear values to each bin. Their findings indicate that the blending of galaxies at different redshifts changes the effective number density distribution as a function of redshift. The implications of redshift-dependent shear within this framework are left for future work.

Another potential extension of this framework involves integrating it with the approach outlined in Zhang et al. (2023), which can detect PSF leakage and modelling errors from all spin-2 quantities contributed by the PSF second- and higher order moments. Combining the second and fourth-order shear estimators with this framework should give more information on the impact of PSF systematics on cosmic shear analyses and how to mitigate their impact on cosmological analysis. A detailed exploration is deferred to future work.

ACKNOWLEDGEMENTS

This work was performed using the Vera cluster at the McWilliams Center for Cosmology at Carnegie Mellon University, operated by the Pittsburgh Supercomputing Center facility. The authors would like to thank Scott Dodelson, Gary Bernstein, and Alexandre Refregier for their useful discussions.

We thank the maintainers of NUMPY (Harris et al. 2020), SCIPY (Virtanen et al. 2020), MATPLOTLIB (Hunter 2007), and GALSIM (Rowe et al. 2014) for their excellent open-source software.

Andy Park, Xiangchong Li, and Rachel Mandelbaum are supported in part by the Department of Energy grant DE-SC0010118 and in part by a grant from the Simons Foundation (Simons Investigator in Astrophysics, Award ID 620789). Xiangchong Li is an employee of Brookhaven Science Associates, LLC under contract no. DE-SC0012704 with the U.S. Department of Energy.

DATA AVAILABILITY

No new data were generated for this work. The code used to analyse simulated galaxy images and obtain shape catalogues is

available from <https://github.com/mr-superonion/xlens/tree/v0.2.2> and the code for shear estimation is available from <https://github.com/mr-superonion/AnaCal/tree/v0.2.2>.

REFERENCES

- Akeson R. et al., 2019, preprint (arXiv:1902.05569)
 Bernstein G. M., 2010, *MNRAS*, 406, 2793
 Bernstein G. M., Jarvis M., 2002, *AJ*, 123, 583
 Bernstein G. M., Armstrong R., Krawiec C., March M. C., 2016, *MNRAS*, 459, 4467
 Bosch J. et al., 2018, *PASJ*, 70, S5
 Chang C. et al., 2013, *MNRAS*, 434, 2121
 Harris C. R. et al., 2020, *Nature*, 585, 357
 Heymans C. et al., 2006, *MNRAS*, 368, 1323
 Hirata C., Seljak U., 2003, *MNRAS*, 343, 459
 Hunter J. D., 2007, *Comput. Sci. Eng.*, 9, 90
 Huterer D., Takada M., Bernstein G., Jain B., 2006, *MNRAS*, 366, 101
 Ivezić Ž. et al., 2019, *ApJ*, 873, 111
 Kaiser N., Squires G., Broadhurst T., 1995, *ApJ*, 449, 460
 Kilbinger M., 2015, *Rep. Prog. Phys.*, 78, 086901
 LSST Science Collaboration 2009, preprint (arXiv:0912.0201)
 Laureijs R. et al., 2011, preprint (arXiv:1110.3193)
 Li X., Mandelbaum R., 2023, *MNRAS*, 521, 4904
 Li X., Katayama N., Oguri M., More S., 2018, *MNRAS*, 481, 4445
 Li X. et al., 2022a, *PASJ*, 74, 421
 Li X., Li Y., Massey R., 2022b, *MNRAS*, 511, 4850
 Li X., Mandelbaum R., The LSST Dark Energy Science Collaboration, 2024a, *MNRAS*, 536, 3663
 Li X., Mandelbaum R., Jarvis M., Li Y., Park A., Zhang T., 2024b, *MNRAS*, 527, 10388
 Liaudat T. I., Starck J.-L., Kilbinger M., 2023, *Front. Astron. Space Sci.*, 10, 1158213
 MacCrann N. et al., 2022, *MNRAS*, 509, 3371
 Mandelbaum R., 2018, *ARA&A*, 56, 393
 Mandelbaum R., Lackner C., Leauthaud A., Rowe B., 2019, *COSMOS real galaxy dataset*, Zenodo
 Mandelbaum R., Jarvis M., Lupton R. H., Bosch J., Kannawadi A., Murphy M. D., Zhang T., LSST Dark Energy Science Collaboration, 2023, *Open J. Astrophys.*, 6, 5
 Massey R., Refregier A., 2005, *MNRAS*, 363, 197
 Massey R. et al., 2007, *MNRAS*, 376, 13
 Moffat A. F. J., 1969, *A&A*, 3, 455
 Pujol A., Kilbinger M., Sureau F., Bobin J., 2019, *A&A*, 621, A2
 Refregier A., 2003, *MNRAS*, 338, 35
 Refregier A., Kacprzak T., Amara A., Bridle S., Rowe B., 2012, *MNRAS*, 425, 1951
 Rowe B., Jarvis M., Mandelbaum R., 2014, Astrophysics Source Code Library, record ascl:1402.009
 Sheldon E. S., Huff E. M., 2017, *ApJ*, 841, 24
 Sheldon E. S., Becker M. R., MacCrann N., Jarvis M., 2020, *ApJ*, 902, 138
 Sheldon E. S., Becker M. R., Jarvis M., Armstrong R., LSST Dark Energy Science Collaboration, 2023, *Open J. Astrophys.*, 6, 17
 The LSST Dark Energy Science Collaboration, 2018, preprint (arXiv:1809.01669)
 Virtanen P. et al., 2020, *Nat. Methods*, 17, 261
 Zhang T. et al., 2023, *MNRAS*, 525, 2441

APPENDIX A: SHAPELET MOMENTS OF GAUSSIAN PROFILES

In this appendix, we provide intuition into the additional information captured by non-adaptive fourth-order moment measurements of a Gaussian profile under shear and compare it against the second-order moment measurement.

Let $f(\mathbf{x}, \sigma_g)$ be a Gaussian profile with scale radius σ_g , which corresponds to the shapelet basis $\chi_{00}(\mathbf{x}|\sigma_g)$. Under shear, the profile transforms as

$$\begin{aligned} \chi_{00}(\mathbf{x}|\sigma_g) \rightarrow \chi_{00}(\mathbf{x}|\sigma_g) &= \bar{\chi}_{00}(\mathbf{x}|\sigma_g) \\ &- \frac{\sqrt{2}\gamma_1}{2} (\chi_{2,-2}(\mathbf{x}|\sigma_g) + \chi_{2,2}(\mathbf{x}|\sigma_g)) \\ &+ \frac{\sqrt{2}\gamma_2}{2} (\chi_{2,-2}(\mathbf{x}|\sigma_g) - \chi_{2,2}(\mathbf{x}|\sigma_g)), \quad (\text{A1}) \end{aligned}$$

where $\bar{\chi}_{00}$ is the intrinsic shapelet basis in the absence of shear. The corresponding shapelet moments M_{nm} of the Gaussian profile can be estimated by projecting the profile to shapelet bases using equation (7).

The generalized Laguerre polynomials are orthogonal over $[0, \infty]$ with a weighting function $x^\alpha e^{-x}$:

$$\int_0^\infty x^\alpha e^{-x} L_n^{(\alpha)}(x) L_m^{(\alpha)}(x) dx = \frac{\Gamma(n + \alpha + 1)}{n!} \delta_{n,m}, \quad (\text{A2})$$

where $\delta_{n,m}$ is the Kronecker delta function. Using the orthogonality relation in equation (A2), if the scale of shapelets and Gaussian is the same ($\sigma_h = \sigma_g$), only the shapelet moments M_{00} , $M_{2,-2}$, and M_{22} are non-zero and fourth-order moments are zero, i.e. $M_{40} = M_{42} = 0$. However, in most cases, $\sigma_h \neq \sigma_g$ or $M_{42} \neq 0$, so shear can be estimated using fourth-order moments. In Fig. A1, we show the spin-2 component of the second- and fourth-order shapelet moments and show that the fourth-order moment is 0 when the scale of the shapelet kernel and Gaussian is the same. For shapelet kernel scale that is smaller than half of the scale of the Gaussian profile (i.e. ratio smaller than 0.5), by comparing the magnitude of each shapelet moment, we find that the fourth-order moments have more information than the

second moment as they can probe smaller radii than the second moment. It is also worth noting that by comparing the magnitude of the shapelet moment, the fourth-order moment has more information on small scales compared to large scales.

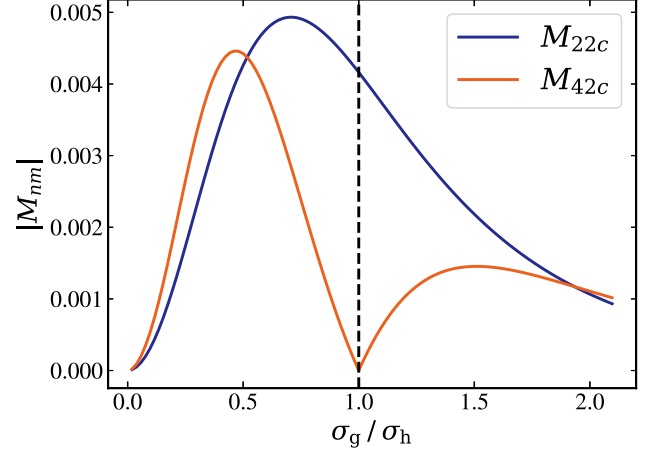


Figure A1. Spin-2 component of the second and fourth-order shapelet moments of a Gaussian profile with input shear $\gamma_1 = 0.02$ as a function of the ratio between the scale of the Gaussian profile to the shapelet kernel. The vertical dashed line indicates when the two scales are identical.

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