Mitigating the effects of undersampling in weak lensing shear estimation with metacalibration

Arun Kannawadi\textsuperscript{1,2}, Erik Rosenberg\textsuperscript{3}, Henk Hoekstra\textsuperscript{2}

\textsuperscript{1}Department of Astrophysical Sciences, Princeton University, 4 Ivy Lane, Princeton, NJ 08544, USA
\textsuperscript{2}Leiden Observatory, Leiden University, PO Box 9513, 2300 RA, Leiden, the Netherlands
\textsuperscript{3}Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK

Accepted 2021 January 21. Received 2021 January 6; in original form 2020 October 8

ABSTRACT

Metacalibration is a state-of-the-art technique for measuring weak gravitational lensing shear from well-sampled galaxy images. We investigate the accuracy of shear measured with metacalibration from fitting elliptical Gaussians to undersampled galaxy images. In this case, metacalibration introduces aliasing effects leading to an ensemble multiplicative shear bias about 0.01 for Euclid and even larger for the Roman Space Telescope, well exceeding the missions’ requirements. We find that this aliasing bias can be mitigated by computing shapes from weighted moments with wider Gaussians as weight functions, thereby trading bias for a slight increase in variance of the measurements. We show that this approach is robust to the point-spread function in consideration and meets the stringent requirements of Euclid for galaxies with moderate to high signal-to-noise ratios. We therefore advocate metacalibration as a viable shear measurement option for weak lensing from upcoming space missions.

Key words: gravitational lensing: weak – methods: observational – cosmology: observations

1 INTRODUCTION

Weak gravitational lensing is a powerful tool to study the growth of large-scale structure in the Universe (see Hoekstra & Jain 2008; Kilbinger 2015; Mandelbaum 2018, for reviews). The shearing of the images of distant galaxies by the gravitational tidal field of intervening structures is now routinely measured in ever larger surveys (e.g. The Dark Energy Survey Collaboration 2005; de Jong et al. 2013; Aihara et al. 2018). The statistics of the resulting correlations in galaxy shapes can be directly compared to models of structure formation, thus constraining cosmological parameters (e.g. Troxel et al. 2018; Hikage et al. 2019; Asgari et al. 2021). The precision will improve dramatically in the next decade with the commencement of a number of large surveys, referred to as Stage IV surveys by the Dark Energy Task Force (Albrecht et al. 2006). Of these, Euclid\textsuperscript{1} (Laureijs et al. 2011) and the Nancy Grace Roman Space Telescope\textsuperscript{2} (formerly known as WFIRST; Spergel et al. 2015; Ake son et al. 2019) are space-based missions, whereas the Rubin Observatory Legacy Survey of Space and Time\textsuperscript{3} (LSST; LSST Science Collaboration et al. 2009; Ivezić et al. 2019) will carry out observations using a new 8m Simonyi Survey Telescope on the ground. Importantly, these experiments are designed with gravitational lensing as a primary objective, so that various systematic errors could be addressed at the early stages through optimal survey design and observational strategies.

Several factors complicate the extraction of the cosmological lensing signal from observed galaxy images, but the most dominant one is the smearing of the images by the point-spread function (PSF), which distorts the shape of the galaxies and blends galaxies along similar lines of sight that may carry different shear signals. Correcting for the effects of the PSF, even for isolated galaxies, in the presence of pixel noise leads to biased estimates of galaxy shapes, and hence to biased estimates of shear (e.g. Hirata & Seljak 2003). Quantifying and correcting for these biases accurately is still an active research. Numerous algorithms, of several flavours, have been developed to estimate shear with minimal biases. Broadly speaking, there are two classes of shape measurement methods (but see Simon & Schneider 2017, which unifies them in the absence of blending).

(i) Model-fitting methods assume a parametric form for the surface brightness profile. The model is convolved with the PSF and the best fit to the data is determined. Various implementations have been developed, most notably lensfit (Miller et al. 2007, 2013), IMSHAPE (Zuntz et al. 2013), and ngmix\textsuperscript{4} (Sheldon 2015).

(ii) Moment-based methods measure a few low-order moments of the images, with a compact weight function centred at each source

\textsuperscript{1}https://www.euclid-ec.org
\textsuperscript{2}https://roman.gsfc.nasa.gov/
\textsuperscript{3}https://www.lsst.org/
\textsuperscript{4}https://github.com/esheldon/ngmix
A comprehensive list of the methods along with their baseline performance can be found in Mandelbaum et al. (2015). The performance of a shape measurement algorithm is typically assessed using simulated images of galaxies (see Plazas 2020, for a recent review). Through such simulations, a series of community-wide blind challenges have provided important insights into the factors that affect the performance of these algorithms (Bridle et al. 2009; Kitching et al. 2011; Mandelbaum et al. 2014). As shown by Hoekstra et al. (2015) and explored further in Hoekstra et al. (2017), the inferred bias depends on the input parameters of the image simulations, such as the number density of galaxies and stars beyond the detection limit, but also the morphological parameters of the bright input galaxies. The presence of undetected faint sources also changes the noise properties, thereby affecting the galaxy shapes (Gurvich & Mandelbaum 2016; Eckert et al. 2020). Thus, inferring the actual shear bias in any observed data set requires dedicated image simulations that match the data as closely as possible, both in terms of galaxy properties and observational conditions (Fenech Conti et al. 2017; Mandelbaum et al. 2018; Zuntz et al. 2018; Kannawadi et al. 2019). These image simulations currently rely on data from the Hubble Space Telescope for obtaining the parent galaxy catalogue, which may become too restricted in terms of depth and cosmic variance (Kannawadi et al. 2015) to use to calibrate shear for Stage IV surveys. Furthermore, these dedicated simulations need to capture the clustering of galaxies at the required depth to evaluate the bias from blending (Samuroff et al. 2018; Euclid Collaboration et al. 2019; Kannawadi et al. 2019). Since most shear measurement methods exhibit biases that are a function of galaxy properties, any uncertainty in the true galaxy population, particularly at the faint end, translates to an uncertainty in shear calibration, which may be the dominant contribution to the overall error budget (e.g. Hildebrandt et al. 2020).

Recently, Huff & Mandelbaum (2017) and Sheldon & Huff (2017) proposed a method known as metacalibration (cf. Section 2.2), that is sufficiently unbiased for any (shear independent) subset of galaxies. Following a similar proposal earlier by Kaiser (2000), metacalibration estimates the bias in the shear measurement (using any shear measurement algorithm) from the data themselves. This method has been validated on image simulations mimicking ground-based seeing conditions to subper cent accuracy in shear bias (Huff & Mandelbaum 2017; Sheldon & Huff 2017) and has been applied to Dark Energy Survey - Year 1 data (Zuntz et al. 2018). The cosmological constraints obtained in Troxel et al. (2018) from a galaxy shape catalogue measured with m3SHAPE and calibrated with image simulations are consistent with those obtained from a shape catalogue measured with NOMIX calibrated on the data with metacalibration. This demonstrates the ability of the metacalibration algorithm to calibrate shear to an accuracy to better than 1 per cent without the need for ultra-realistic galaxy populations in the simulations. However, image simulations are nevertheless required to validate that shear measurement satisfies the stringent requirements of the Stage IV surveys (e.g. Sánchez et al. 2020). For instance, the Euclid mission requirements document (Laureijs et al. 2011) mandates that the shear be recovered better than 0.2 per cent to realise the promised figure of merit. At such increased levels of accuracy, one has to worry about systematic effects that were largely ignorable in the past. These include, but are not limited to, instrumental effects such as the brighter-fatter effect (Antilogus et al. 2014; Coulton et al. 2018), charge transfer inefficiency (Rhodes et al. 2010; Israel et al. 2015) and astrophysical effects such as the presence of colour gradients in galaxy profiles due to wavelength dependence of a diffraction limited PSF (Voigt et al. 2012; Semboloni et al. 2013; Carlsten et al. 2018; Er et al. 2018) and blending of galaxies (Dawson et al. 2016; Hoekstra et al. 2017; Euclid Collaboration et al. 2019). An updated version of metacalibration with object detection included (metadetection; Sheldon et al. 2020) provides a way forward to account for some of the blending related biases.

A potential limitation for metacalibration in the case of space-based missions is pixelization. While the space-based surveys enjoy a better and more temporally stable PSF than their ground-based counterparts5, their PSFs are typically not Nyquist-sampled in any given exposure, despite the detectors having physically smaller pixels. This can lead to small galaxies being poorly sampled as well. As metacalibration involves reconstructing the galaxy profile from a discretized image (cf. Section 2.2), when the galaxy image is undersampled, the interpolation step will be unable to recover the galaxy profile faithfully. In this paper, using image simulations, we investigate the level of residual biases as a result of having an undersampled PSF. In our companion paper by Hoekstra et al. (2021), we further validate our conclusions with more realistic simulations. While both of these papers adopt Euclid as a reference, our results are fairly generic and are also applicable to the High-Latitude Imaging Survey\(^6\) (HLS), the weak lensing programme of the Roman Space Telescope with the aim of studying dark energy parameters.

This paper is organized as follows. In Section 2, we provide an overview of the metacalibration procedure and how pixelization can degrade the performance of the procedure. We explain the computational methods used in this study in Section 3 and show our main results on shear bias in Section 4. Finally, strategies for mitigating this bias are explored in Section 5 and a general discussion of our results follows in Section 6.

2 OVERVIEW OF SHAPE MEASUREMENTS AND METACALIBRATION

We describe the effects of pixelization in the context of quadrupole moments, since they are more amenable to analytical calculations than model parameters, although we believe both classes of methods should be affected in a similar manner for unblended sources (see Simon & Schneider 2017, which unifies both frameworks). We will defer the explicit treatment of model-fitting methods on undersampled images to a later work.

2.1 Shape measurements

Mathematically, the quadrupole moment is a symmetric rank-2 tensor that characterizes the distribution of an object around a chosen point, usually the centroid. Despite allowing for an unbiased estimate of the shear, one cannot use unweighted moments in the

---

5 Even ground-based PSFs may be undersampled under extremely good seeing conditions, which ironically then have to be rejected (e.g. Bosch et al. 2018)

6 https://www.roman-hls-cosmology.space/
The presence of noise. One must use weighted moments instead, which inevitably leads to biases that have to be calibrated out. The components of the quadrupole moments\(^7\) are given by
\[
Q_{ij} = \int d^2x x_i x_j W(x) f(x),
\]
where \(x_i\) and \(x_j\) are the components of \(x\) for \(i, j \in \{1, 2\}\) with its centroid at the origin and \(W(x)\) could in principle be an arbitrary non-negative function that is sufficiently localized around \(x = 0\) to suppress the noise at large distances. Bernstein & Jarvis (2002) suggest for maximum signal-to-noise (SNR) the use of adaptive moments, where the weight function is the best-fitting elliptical Gaussian function of the galaxy itself.

The complex ellipticity in terms of unweighted quadrupole moments is then given by the combination \(Q_{11} - Q_{22} + 2iQ_{12}\), with a normalizing factor of overall size, given by either \(Q_{11} + Q_{22}\) or \(Q_{11} + Q_{22} + 2\sqrt{Q_{11}Q_{22} - Q_{12}^2}\) so that the ellipticity is invariant under flux scaling and length scaling. Following the notation of Bartellmann & Schneider (2001), we refer to the former as \(\chi\)-type ellipticity (also known as \textit{distortion}) and the latter as \(\epsilon\)-type ellipticity, and they transform\(^8\) differently under a lensing shear as follows:

\[
\epsilon = \epsilon^{(\text{int})} + g \quad \text{and} \quad \chi = \chi^{(\text{int})} + 2g + g^2 \frac{\epsilon^{(\text{int})}}{1 + g^2}. 
\]

where \(\epsilon^{(\text{int})}\) and \(\chi^{(\text{int})}\) denote the intrinsic ellipticity of a galaxy and \(g = g_1 + ig_2\) is another complex quantity referred to as the (reduced) lensing shear (Schneider & Seitz 1995; Seitz & Schneider 1997). Thus, the ellipticity of a galaxy is a noisy estimate of the shear, with the intrinsic shape being the noise here. Invoking the isotropy assumption, \(\langle \epsilon^{(\text{int})} \rangle = \langle \chi^{(\text{int})} \rangle = 0\), we can see that the ellipticity can be considered as a one-point estimator of shear since

\[
\langle \epsilon \rangle = g \quad \text{and} \quad \langle \chi \rangle = 2 \langle \chi^{(\text{int})} \rangle g + O(|g|^2). 
\]

Metacalibration with undersampling\(^3\) hence imperfect) due to its perturbative nature, acting as yet another source of bias.

For small values of shear, which is usually the case in weak gravitational lensing, we can capture the deviation of the estimated shear \(\tilde{\epsilon}\) from the true shear \(\epsilon_{\text{true}}\) from the linear bias model (Heymans et al. 2006; Massey et al. 2007) with a two-component additive bias \(\epsilon\) and a \(2 \times 2\) multiplicative bias tensor \(M\) as

\[
\tilde{\epsilon} = \epsilon_{\text{true}} + \epsilon + M \epsilon_{\text{true}},
\]

with

\[
\epsilon = (c_1, c_2)^T, 
\]

and

\[
M = \begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix}.
\]

Typically, the off-diagonal elements \(m_{12}\) and \(m_{21}\) are negligible and \(m_{11} = m_{22}\). Thus, \(M\) is treated as a rotationally invariant scalar in practice. However, we will use a more generic \(2 \times 2\) tensor in this study. We also denote the shear in bold font as a two-component spinor instead of a complex number for convenience.

\subsection{2.2 Metacalibration algorithm}

The multiplicative bias may be seen as an average response of the shapes of a population of galaxies to a coherent shear. The \textsc{metacalibration} algorithm attempts to estimate empirically the responsivity of each galaxy in the data. It involves deconvolving the observed image by the PSF \(P\) to obtain the (noisy) galaxy profile, shearing the galaxy by a small amount \(|g|| \sim 0.01\) and reconvolving by a slightly larger PSF given by \(P(\tilde{g}/|\tilde{g}|^{-1}x)\) to obtain the galaxy image under a different shear condition. Mathematically, we can describe the newly obtained image as the inverse Fourier transform of

\[
\tilde{I}(k|g) = \tilde{\Gamma}^* (k) \tilde{\epsilon} \frac{\tilde{J}(k)}{\tilde{P}(k)},
\]

where \(\tilde{\epsilon}\) is the shearing operator and \(\tilde{J}(k)\) stands for the Fourier transform of any real-space function \(J(x)\). The per-galaxy responsivity to shear, a \(2 \times 2\) tensor \(R\) defined as

\[
R := \frac{\partial \epsilon}{\partial g} \bigg|_{g=0} = \epsilon^+ - \epsilon^-, \quad \frac{\partial \chi}{\partial g} \bigg|_{g=0} = -\epsilon^+ + \epsilon^-.
\]

where \(\epsilon^+\) and \(\epsilon^-\) are small artificial (stimuli) shears applied to the deconvolved image through the shearing operator \(\tilde{\epsilon}\), and \(\epsilon^+\) and \(\epsilon^-\) are the corresponding measured ellipticities. Here, we simply denote the ellipticity in spinor form by \(\epsilon\) and do not make a distinction between the \(\epsilon\)- and \(\chi\)-type ellipticities as the multiplicative factor is naturally captured by the shear response. The shear responsivity includes much of the raw multiplicative bias in the estimator and the unbiased shear estimate in the absence of PSF anisotropy is

\[
\langle \epsilon \rangle = \langle R \rangle^{-1} \langle \epsilon \rangle.
\]

Note that in this estimator, all the galaxies carry equal weight and it ignores the uncertainty in shape estimates. Depending on the level of noise, we might want to downweight some galaxies in comparison to others but in a manner that does not depend on the ellipticity itself. The generalization of the shear estimator given in equation (11) in the presence of galaxy weights \(w\) is

\[
\tilde{\epsilon} = \langle w R \rangle^{-1} \langle w \epsilon \rangle.
\]
2.3 Effects of pixelization

The effect of sampling and aliasing in the images can be best understood in the Fourier domain. In the absence of noise and sky background, the observed image is the inverse Fourier transform of \( \tilde{I}(k) = \tilde{G}(k) \tilde{P}(k) \), where \( \tilde{G}(k) \) and \( \tilde{P}(k) \) are the Fourier transforms of the galaxy and the PSF respectively. In practice, \( \tilde{I}(k) \) has to be computed at discrete values of \( k \) from a discrete (pixelated) image of finite extent. The GaSim implementation of the METACALIBRATION algorithm involves computing a discrete approximation to the Fourier transforms by first constructing a continuous function by interpolating the pixel values, and taking the Fourier transform assuming a periodic boundary condition, which is arguably better than using discrete Fourier transforms.

Note that the quantities above, especially \( \tilde{I}(k) \), are not discrete Fourier transforms of the pixelized image but a discrete approximation to the Fourier transform of \( \tilde{I}(k) \) obtained by interpolating from the discrete image. Due to the discretization factor, there exists a \( k_{\text{max}} \) such that \( \tilde{P}(k) \equiv 0 \) for all \( k \) with magnitude \( |k| > k_{\text{max}} \). This holds true irrespective of the central obscuration in the telescope aperture. Thus, ignoring the read noise, the observed image is band-limited. For interpolation using the sinc function, which is perfect for band-limited signals, the interpolated image is formally the inverse Fourier transform of

\[
\tilde{I}_{\text{interp}}(k) = \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} \tilde{I}(k + n \Delta_k) \Theta(\Delta_k - |k_1|) \Theta(\Delta_k - |k_2|),
\]

(13)

where \( n = (n_1, n_2) \), \( \Theta(\cdot) \) is the Heaviside step function, and \( \Delta_k \) is the periodicity in Fourier space due to starting from discrete pixel values. For square pixels of physical size \( a \), \( \Delta_k = 2\pi/a \).

According to the Nyquist-Shannon sampling theorem, in order to reconstruct a continuous function from discrete samples by interpolation, the (spatial) sampling frequency has to be more than twice the maximum frequency present in the original signal, also referred to as the Nyquist frequency. This follows from equation (13): if \( \tilde{I}(k) = 0 \) for \( |k| > \frac{\Delta_k}{2} \), then only the contributing term is that of \( n = 0 \) and \( \tilde{I}_{\text{interp}}(k) \equiv \tilde{I}(k) \). If the above condition is not satisfied, then frequencies higher than the Nyquist frequencies are misinterpreted as lower frequencies, a phenomenon well-known as aliasing. If we perform operations such as applying a lensing shear, the images obtained by shearing \( \tilde{I}_{\text{interp}} \) would be systematically different from that obtained from \( \tilde{I}(k) \), were it to be accessible.

We define\(^9\) the sampling factor \( Q \) as the ratio of the Nyquist scale (reciprocal of Nyquist frequency) to the pixel scale. An image is said to be Nyquist-sampled if \( Q > 1 \), critically sampled if \( Q = 1 \) and undersampled if \( Q < 1 \). An undersampled image is said to be strongly undersampled if \( Q < 0.5 \), where all frequency modes are aliased. Incidentally, if \( Q < 0.5 \), the true profile of the object cannot be recovered from any set of four differed exposures (see Lauer 1999, for example).

If the PSF is Nyquist-sampled then by definition, \( \Delta_k > 2k_{\text{max}} \) for all galaxies and there is no signal present if \( n \neq 0 \). \( \tilde{I}_{\text{interp}}(k) \) has only a small contribution only from the (subdominant) detector noise, whose effect is to change the inferred property of the noise characteristics. For an undersampled PSF, in addition to the detector noise, there is a finite contribution from the signal as well. This results in \( I'(k|g) \) in equation (9) being different from \( \tilde{I}(k) \) even when \( g = 0 \). While the PSF is not Nyquist-sampled, it can nevertheless be modelled very well from multiple star images, so we will assume that an oversampled image of the PSF is available which does not suffer from aliasing when interpolated between pixels. In this paper, we refer to the bias in shear estimate arising due to aliasing as ‘aliasing bias’. We refrain from referring to it as ‘pixelization bias’, in order to avoid confusing this with other sources of bias such as discretisation of a continuous image.

A distinctive signature of aliasing is that it affects \( g_1 \) preferentially more than \( g_2 \). We motivate this generic result using symmetry. Aliasing introduces an oscillatory pattern along the axes, modulated by the image \( I(x) \). If, on average, the image is isotropic but for a small \( g_2 \), the contributions from the \( n \neq 0 \) term in the integral for \( Q_{12} \) are zero. However, if there is a net \( g_1 \), then the horizontal and vertical axes are not treated on the same footing. The oscillations do not cancel each other, contributing to \( Q_{11} - Q_{22} \). This leads to a non-zero \( m_1 \), but not \( m_2 \). In an image with discrete pixels, the subpixel offset breaks this symmetry, which leads to a non-zero but small \( m_2 \) in practice.

3 METHOD

Our goal is to evaluate the impact that aliasing would have for measuring shear using METACALIBRATION from planned Stage IV surveys. Specifically, we study the performance of METACALIBRATION on simulated galaxy images, some of which may be poorly sampled. We use the specifications of Euclid in our fiducial simulation set-up, and extend conclusions to the Roman Space Telescope by considering corresponding sampling factors. The Euclid telescope has an entrance pupil that is 1.2 m in diameter. Euclid will cover about 15,000 deg\(^2\) of the sky, thereby imaging about two billion galaxies with an SNR greater than 10. The images for shape measurements are captured using the Euclid visible (VIS) instrument\(^10\), with a single broad band filter covering a wide range of wavelengths from 550 nm to 900 nm and at a resolution of 0.1’’ pixel\(^{-1}\). The PSF is wavelength dependent set by the diffraction limit, and as a result, each galaxy is convolved by a different PSF determined by its spectral energy distribution (SED).

3.1 Image simulations

Postage stamp images of isolated galaxies are simulated with the publicly available GaSim\(^11\) package (Rowe et al. 2015). We model the surface brightness profiles of galaxies using Sérsic profiles (Sérsic 1968), which have well-defined ellipticities. GaSim comes equipped with a COSMOSCatalog, which is a catalogue of best-fitting Sérsic parameters to galaxies in the COSMOS survey (Koekemoer et al. 2007; Leauthaud et al. 2007; Scoville et al. 2007). The best-fitting parameters were obtained using the method described in Lackner & Gunn (2012). This serves as our parent galaxy catalogue in estimating shear biases (cf. Section 4). The distribution of the sizes of the galaxies is shown in Figure 1. Distributions of other galaxy properties such as apparent magnitudes, best-fitting Sérsic parameters can be found in figs 11-13 in the appendix of Mandelbaum et al. (2014). Our parent galaxy catalogue is likely to be different from catalogues from Euclid or Roman Space Telescope due to different selection functions. Hence, the shear biases derived

\(^9\) Our definition differs from that of Shapiro et al. (2013); Plazas et al. (2016) by a factor of two.

\(^10\) https://sci.esa.int/s74EDxW

\(^11\) https://github.com/GalSim-developers/GalSim
in this study are not directly applicable to those surveys but are only indicative.

Since the PSF is diffraction-limited, its size depends on the colour of the galaxy itself. For the sake of simplicity, we use a monochromatic Airy function as the PSF for all galaxies, corresponding to a wavelength of 800 nm (unless otherwise mentioned) and arising from a circular aperture of diameter 1.2 m, with a central obscuration factor of 0.3. We exploit the perfect knowledge of this input PSF, assuming that the PSFs in the real data can be modelled to sufficient accuracy. An imperfect knowledge of the PSF will introduce bias for all shear measurement methods, and quantifying this for METACALIBRATION is outside the scope of this paper.

Euclid will image a given field using the VIS instrument with three or four dithered exposures, and the HLS with the Roman Space Telescope will include several dithered visits, over multiple passbands however. Therefore, we render four individual exposures for each galaxy. The exposures have the same PSF but differ by uniform subpixel offsets of half a pixel in each direction. The galaxy itself has a random subpixel shift to avoid consistently sampling at the peak of the light profile. We also include a 90°-rotated copy (about the same randomly offset centre) for each galaxy to beat down the intrinsic shape noise (Massey et al. 2007). The mapping between sky coordinates and pixel coordinates is linear, given by a constant pixel scale of 0.1 pixel⁻¹. Since the goal of this paper is to study the bias just due to pixelization and having undersampled PSFs, we do not add any pixel noise to the images for most simulations in this paper. The rationale is that it is sufficient to show on noiseless simulations that aliasing bias is significant. Adding pixel noise complicates the study by inducing correlations in the noise and also contaminates the estimate of aliasing bias with noise bias.

Furthermore, we do not expect any significant coupling between aliasing and other detector-level systematics. The quantities $I_{\text{interp}}(k)$ and $\hat{I}(k)$ (see equation 13) differ systematically only for those values of $k$ for which $\Delta_k < |k| < k_{\text{max}}$. Instrumental effects that mix the different $k$-scales can transfer the spurious power in these high-frequency modes to lower frequencies, which, if not corrected for accurately, could bias the flux and shape measurements. However, we expect such imperfect corrections to bias the measurements regardless of whether the PSF is Nyquist-sampled or not. Hence, our simulations also assume ideal detector behaviour.

3.2 Metacalibration procedure

Although the PSFs in space telescopes are temporally stable, the variation of the PSF over the focal plane could mean that different exposures of the same galaxy have different PSFs. For this reason, ideally, one might want to apply the metacalibration procedure to the individual exposures with corresponding PSFs, rather than applying it to a coadded image. For convenience, we refer to the output images of metacalibration as metacalibrated exposures. We apply the metacalibration algorithm (see Section 2.2) to each exposure with five artificial shears of 0, ±0.01 and ±0.01, generating five metacalibrated exposures per original exposure.

To mimic the metacalibration procedure that will be done on the real data, we do not use the perfectly known, smooth PSF profile, but render the PSF image by convolving the obscured-Airy profile with a top-hat pixel response function corresponding to a pixel scale $p = 0\text{"}1$ pixel⁻¹ and oversampling the image by a factor of five. A smooth PSF model is then constructed from this oversampled PSF image, which is used for the deconvolution step. The PSF interpolated from a finite-sized image is not strictly band-limited. To suppress any spurious power at high spatial frequencies, we use the band-limited Airy PSF for the deconvolution step. Following the prescription in Huff & Mandelbaum (2017), we dilate the analytical PSF prior to its convolution with the pixel response function and use an oversampled image of the same for shape measurements. The choice of re-convolving with a band-limited, preferably analytical, function should be made for the real data to suppress any artefacts arising from the deconvolution step.

We find that the quintic interpolation scheme (the default option in GalSim) is far too inaccurate and causes large differences between the ellipticities prior and posterior to interpolation. Switching to a sinc interpolation kernel improves the accuracy, but is extremely slow to render the images after interpolation. We find LanceCon with $N = 50$ as an optimal compromise between speed and accuracy and is our default interpolation kernel. See Appendix A for more details on our choice of interpolation kernel.

In order to further ensure that our results are not affected due to moment calculations from poorly sampled PSF images or shape measurement failures in a particular subpopulation of galaxies, we employ a control branch that is similar to the idea proposed by Pujol et al. (2019). In our control branch, we skip the steps in metacalibration where the galaxy image is interpolated and deconvolved by the PSF. Instead, we include the artificial shear prior to PSF convolution, and then render the image with dilated PSF and measure its shear responsivity. Therefore, by construction, we do not expect our control branch to exhibit aliasing bias. To avoid any subtle selection effects between the metacalibration and control branches (some measurements are failures only in the metacalibration branch), we impose the two to have the same population of galaxies. Note that our pipeline does not include detection and deblending steps, as we simulate images of isolated galaxies. As a result of this simplification, we do not suffer from the object detection bias discussed in Sheldon et al. (2020) and Hoekstra et al. (2021).

3.3 PSF correction

In principle, METACALIBRATION can account for the PSF in the calculation of the per-object responsivity without any explicit PSF-correction. Hoekstra et al. (2021) show that Metacalibration works in a Euclid-like set-up (with constant PSF) without an explicit PSF-correction. However, in the case of variable PSFs it
may prove beneficial to use some simple PSF-correction schemes to detrend some of the well-understood effects of PSF, so that the biases that need to be removed by metacalibration are small to start with. While Sheldon et al. (2020) show that smooth PSF variations, at the levels expected for the LSST, are not an issue, it remains unclear whether galaxy-by-galaxy PSF variation in Euclid due to their SEDs could be similarly handled. Hence, we employ a few simple PSF-correction methods and study their sensitivity to metacalibration. We provide a brief overview of the methods used in this paper.

3.3.1 Gaussian fit
This method approaches both the PSF and the galaxy as elliptical Gaussians. We use a simple single-Gaussian fitting method whereby we fit an elliptical Gaussian to the PSF and to the galaxy image. The difference in the covariance matrices of the two Gaussians gives the equivalent of unweighted moments for the intrinsic galaxy. Due to its simplistic assumptions, this method is expected to suffer from severe model bias (Bernstein 2010; Voigt & Bridle 2010). However, in combination with metacalibration this has been shown to achieve residual biases within a per cent on real data (Sheldon & Huff 2017; Zuntz et al. 2018).

The fitting methodology employed above is different from those methods that fit a PSF-convolved galaxy model to the image and maximize the likelihood to find the best-fitting parameters. However, in the absence of blending and missing pixels, model-fitting methods and moment-based methods are equivalent (Simon & Schneider 2017). We will therefore not concern ourselves with the exact details of the model-fitting approach.

3.3.2 Regaussianisation
A natural extension of the idea above is captured in the Regaussianisation method (Hirata & Seljak 2003). It is a perturbative method whereby a small non-Gaussianity in the PSF is accounted for and the intrinsic ellipticity of the galaxy is obtained from the second-order moments, with correction terms from fourth-order radial moments included (see Bernstein & Jarvis 2002, for details). We used the GalSim implementation of the algorithm in this paper.

3.3.3 KSB
The KSB method, named after the authors of Kaiser et al. (1995), is one of the pioneering shape measurement methods completely based on image moments. We use the GalSim implementation of this method described in Appendix C of Hirata & Seljak (2003). In this implementation, the ellipticity estimator is calculated from the quadrupole moments measured with a circular Gaussian weight function, whose size is typically (but not necessarily) matched to observed galaxy. The PSF moments are also calculated using the same weight function as the galaxy following Hoekstra et al. (1998). From the spatial derivatives of the weight function, polarizability tensors are calculated and are used to remove the effect of PSF anisotropy and obtain shear.

It is worth noting that all of these methods are inexpensive computationally, both in terms of memory and computing cycles, in relation to metacalibration (see Table A1). Also, note that among the three methods listed above, only KSB has a freedom in the choice of the weight function. We will utilise this freedom in Section 5.

Figure 2. Relationship between input and output shears from different PSF-correction methods for a random subset of 10 000 galaxies. The true shear values range from -0.04 to 0.04 in steps of 0.01 in one component, while the other component is fixed at zero. The data points are offset slightly along the horizontal direction for clarity. The filled points are the shear recovered from the metacalibration branch and the open points are from the control branch. The solid curves show the best-fitting cubic polynomial and the dashed-dotted lines are the best-fitting straight lines (for $g_1$ only), with the best-fitting parameters listed in Table 1.

The PSF-correction routines in GalSim require that both the galaxy and PSF moments are calculated from images with the same pixel scale. This limitation prevents us from using the perfectly known PSF model. The PSF moments (and moments of small galaxies) calculated from images at native pixel scale have sampling errors large enough to cause ensemble shear biases of the order of 0.01, but are greatly mitigated when the PSF image is oversampled by a factor of two or more (e.g. Hoekstra et al. 2021). To be consistent with these oversampled PSF images, we interleave the four metacalibrated exposures to obtain a high-resolution coadded image, with an effective pixel scale of 0.05 pixel$^{-1}$ using the galsim.utilities.interleaveImage routine (Kannawadi et al. 2016). This step plays the role of a more complicated coaddition technique that may be used in practice and represents the best case scenario. A typical coadded image will not be sampled this well. In addition to the increased resolution, the advantage of interleaving (as in any coaddition) is that it allows us to use shape information of galaxies from multiple images simultaneously, as in multi-epoch model fitting methods.

4 Quantifying aliasing bias
For a (semi-)realistic population of galaxies taken from the CO-SMOS sample (see Section 3.1), we show how well we can recover the cosmological shear signal. As mentioned in Section 2.1, we relate the recovered shear to the true shear via a generalized linear shear bias model given in equation (6).

The primary source of the $e$ term is PSF leakage, i.e., inade-
Table 1. Residual multiplicative bias in shear measured with various PSF-correction methods after applying \textsc{metacalibration}. The uncertainties capture the deviation of the mean shear from the best-fitting function, and do not include the uncertainty in the shear measurement itself. The \(\epsilon\)-terms are from the linear fit were very similar to cubic polynomial fit.

| Algorithm for \(|m_1|\times10^2|\) \(\chi\) \(\epsilon\) from applied \(|m_2|\times10^4|\) \(\chi\) \(\epsilon\) from applied | \(\chi\) \(\epsilon\) from applied | \(\chi\) \(\epsilon\) from applied |
|------------------|------------------|------------------|------------------|
| Gaussian fit     | -0.81 ± 0.01     | -0.8966 ± 0.0002 | -6.420 ± 0.003   | 0.2 ± 2.0        | -0.3 ± 0.3 |
| Regaussanisation | -1.17 ± 0.03     | -1.3380 ± 0.0002 | -11.97 ± 0.01    | 0.5 ± 2.1        | -0.9 ± 1.4 |
| KSB              | -0.853 ± 0.003   | -0.8342 ± 0.0003 | -7.614 ± 0.002   | 0.2 ± 2.5        | 0.2 ± 0.2 |

One of two shear components is non-zero, we recovered (al-

quate correction for PSF ellipticity. As we have chosen a circular

PSF for our study, we do not expect any additive bias terms. We

checked this explicitly by seeing that we recover a null signal (about

10^{-5} or smaller) when the true shear is zero. Additionally, when

only one of two shear components is non-zero, we recovered (al-

most) zero shear for the component set to zero. We also noticed a

mild linear dependence on the non-zero component, corresponding

to \(|m_1|\) \(\approx |m_2|\) \(\lesssim 2 \times 10^{-4}\).

In practice, the PSF will be slightly anisotropic, leading to non-
egligible additive bias terms due to PSF leakage, and also possibly
due to pixelization effects. However, in this study, we are not in-

terested in quantifying the contribution of pixelization to additive

biases as any residual biases can be estimated and removed from the
data themselves. We will therefore focus on the multiplicative bias

terms, and thus the ability to recover non-zero shear. Our results

already indicate that there is no discernible cross-talk between the
two shear components, and therefore the multiplicative bias \(|M|

may be considered to be strictly diagonal.

Figure 2 shows the relationship between the true input shear \(g_1\)

and the recovered shear \(\hat{g}_1\) for \(i = 1, 2\). The presence of a small

higher order (cubic) relation is evident in the \(g_2\) component, which
closely follows the trend in the control branch for \(g_2\) (and \(g_1\)). The

cubic relation is a result of calculating \(\chi\)-type ellipticities as op-
posed to \(\epsilon\)-type ellipticities\(^{12}\) (see equation 5) and is inherent to

the Gaussian fit and \textsc{Regaussanisation} estimators. We notice that

the systematic error in \(g_1\) is dominated by a linear term, which is

practically negligible in the control branch. In Table 1, we list the

multiplicative and additive bias terms obtained by fitting a cubic

function (without a quadratic term) to the points in Fig. 2. In all

cases, the value of \(m_1\) is an order of magnitude larger than \(m_2\), indic-
ating that aliasing affects \(m_1\) more than \(m_2\), which is not surprising
given the symmetry argument in Section 2.3. We also fit a straight

t line to the data points in the upper panel of Fig. 2, whose slopes

are tabulated in Table 1. A comparison of the two \(m_1\) columns for

\(\chi\)-type measurements shows that when the nonlinear terms are

dropped, the amplitude of \(m_1\) is overestimated by a small amount,
by approximately \(m_2\).

Since the biases from higher order shear terms, cross-terms in

\(|M|\) and \(|\epsilon|\) are negligible, we estimate the linear multiplicative bias

as

\[
m_{ij} := \hat{g}_i / |g_i| - 1, \tag{14}
\]

for \(i = 1, 2\) and for small enough \(g_i\). Note that this simple form of the

estimator implies we do not have to perform a linear regression over

multiple true shear values to find the multiplicative term, thereby

\(^{12}\) Although the \(\epsilon\)-type ellipticity appears to be the preferred type as it is

unaffected by higher order shear terms, the quantity under the square root is not guaranteed to be positive for PSF-corrected moments in the presence of noise, which introduces biases.

\[\text{Figure 3. Lower panel: Aliasing bias as a function of minimum half-light radius imposed on the galaxy samples. The filled points show } m_1 \text{ after } \textsc{metacalibration} \text{ and the open points show } m_2. \text{ Both } m_2 \text{ and } m_1 \text{ for the control branch (dashed-dotted lines) show signs of nonlinear bias in } \chi\text{-type measurements. The shaded region represents that } \text{Euclid requirement on } m_1 \text{ to be within } \pm 2 \times 10^{-3}. \text{ Upper panel: The fraction of galaxies in our sample (by number) that survive the minimum size criterion.}
\]

\[\text{Figure 3 shows the cumulative marginal distribution of input galaxy}
\]

\[\text{reducing the volume of simulations required. The uncertainty in } m_1 \text{ is then simply the uncertainty in } \hat{g}_2 \text{ scaled by } 1/|g_1|. \text{ In the following sections, the multiplicative bias is computed from simulations with input shear of } \hat{g}_1 = g_2 = -0.02.\]

\[\text{4.1 Dependence on galaxy sizes}
\]

We demonstrate that the bias observed in the \(g_1\) component is pre-
dominantly due to a subpopulation of small galaxies, whose intrinsic

sizes happen to be of the order of, or smaller than, the pixel size,
and hence undersampled in a single exposure. By progressively

eliminating small galaxies from our sample, we show in Fig. 3

that \(m_1\) converges to zero within the accepted range of accuracy.
Eliminating galaxies smaller than a given threshold corresponds to

a special case of equation (12) with a binary weight \(w(r; r_{\text{min}}) = \Theta(r - r_{\text{min}}), \) where \(\Theta(\cdot)\) is the Heaviside-step function, \(r\) is the

true intrinsic half-light radius of the galaxy (circularized, so that
the weight is independent of ellipticity), and \(r_{\text{min}}\) is the threshold,
which sets the minimum size of the galaxy in our sample. Figure 3
also shows explicitly that the aliasing affects only \(m_1\) and has little
effect on \(m_2\).

The bias for a sample of galaxies depends on its size distri-

bution, as a result of this size dependence. The upper panel of

Fig. 3 shows the cumulative marginal distribution of input galaxy
sizes, marginalized over other galaxy properties such as intrinsic ellipticities, Sérsic indices etc. For the COSMOS galaxy sample we simulated, the bias is dominated by 13 per cent of the galaxies, whose intrinsic radii is smaller than \(0.01\), or 1 native pixel.

Note that we do not have a detection stage in our simplistic set-up and are also insensitive to the absolute level of surface brightness of the galaxy due to absence of a sky background and noise. While we do not apply an explicit magnitude cut in our simulations, our input catalogue is nevertheless magnitude-limited. However, those galaxies in our input catalogue with \(r < 0.01\) were well resolved in the Wide Field Camera of the Hubble Space Telescope with a pixel scale of \(0.05\) pixel\(^{-1}\) (and an effective pixel scale of \(0.03\) pixel\(^{-1}\) after coaddition), but may escape detection or be poorly resolved in Euclid and Roman Space Telescope images which have larger pixel scales. Thus, for future Stage IV experiments from space, an implicit selection effect due to detection may play a natural role in excluding such small galaxies and may suffer less from aliasing than shown here. On the other hand, galaxies are not smooth and have complex morphologies in reality. The size parameter \(r\) merely describes how rapidly the surface brightness drops given a Sérsic index. In practice, even large galaxies that exhibit small-scale features will contribute to aliasing (see Appendix C for shear measured from realistic galaxies).

A plausible explanation for the bias could be mis-centring, i.e., the offset between the centres of the galaxies and PSF models. While the PSFs are always centred at the centre of the central pixel, the galaxies have random subpixel offsets. The difference in sampling could potentially introduce biases, with smaller galaxies showing larger biases. If this was the dominant source of bias, we would then expect it to be the case even in the control branch where such mis-centring is present. Furthermore, since the PSFs are oversampled by a factor of five, it is highly unlikely that these biases could be arising due to sampling. Finally, we explicitly checked if this is the case by centering all galaxies at the pixel centre and found no significant difference. Hence, we rule out mis-centring as a potential source of the bias.

4.2 Dependence on sampling factor

In order to establish that this bias arises from the undersampling of the images, we study how the shear is recovered as a function of sampling factor \(Q\). We vary the pixel scale to vary \(Q\) and show \(m_1\) as a function of the pixel scale in Fig. 4. The pixel scale in the Euclid VIS detectors is fixed at 0.1 pixel\(^{-1}\) and in the Roman Space Telescope it is 0.11 pixel\(^{-1}\). At a pixel scale of 0.1 pixel\(^{-1}\) (that of Euclid VIS detectors), we see a clear difference in signal obtained using the metacalibration branch and control branch. The uncertainty values are obtained from 1000 bootstrap realizations of galaxy pairs. We see that for Euclid, a residual bias in \(m_1\) of approximately 0.01 after metacalibration is possible, while the control branch shows no biases exceeding the tolerance.

Although our set-up mimics Euclid, it is straightforward to assess the level of bias that would be incurred in a Roman Space Telescope-like setting using scaling arguments. A suite of dedicated image simulations for Roman Space Telescope, one that is more representative than this study is for Euclid, is given in Troxel et al. (2021), albeit without metacalibration. Despite the Roman Space Telescope bandpasses covering the near-infrared wavelengths, the telescope aperture has twice the diameter as that of Euclid and the pixel scale of the H4RG detectors in Roman Space Telescope is only 10 per cent bigger than Euclid VIS detectors. As a result, the Roman Space Telescope PSFs are likely to be equally undersampled or worse than the Euclid PSFs. The sampling factor acts as a powerful summary statistic that is robust to change in the details of the PSF, pixel scale, aberrations etc.

The shape measurements for the HLS are intended to be done in three bandpass filters: \(J\), \(H\), \(F\), whose effective wavelengths are 1290, 1580 and 1840 nm, respectively. The sampling factors for the three filters are approximately 0.504, 0.617, and 0.718 respectively, as indicated by vertical dotted lines in Fig. 4. As shown in Fig. 4, we obtain a bias of about 0.015 and 0.025 for \(F\) and \(H\), bands respectively. For the bluest bandpass \(J\), all of our measurements made using our set-up were failures, even in the control branch. Extrapolating the data points to higher sampling factors gives a bias estimate of about 0.1. These estimates may be approximate, but they are sufficient to demonstrate that the bias exceeds well above the much tighter requirement of \(3.2 \times 10^{-4}\) set in Doré et al. (2018).

These dependencies on the galaxy sizes and sampling ratio, combined with the fact that \(|m_1| > |m_2|\), strongly suggest that the source of bias is aliasing. The shear measured using adaptive Gaussian moments from metacalibrated exposures can be biased at a few per cent level when the PSF is undersampled, thereby requiring large calibration corrections that need to be calculated from image simulations. This negates the fundamental idea behind metacalibration of calibrating shear measurements from the data itself without requiring any (or at least large) corrections from image simulations. As explained in Section 1, the size dependence of aliasing bias introduces a sensitivity to galaxy populations that is otherwise benign for metacalibration when the PSF is well sampled. Even if the aliasing bias can be calculated accurately with sophisticated simulations, it is of practical inconvenience for \(m_1\) and \(m_2\) to have very different values, since the multiplicative bias can no longer be treated as a scalar as they are done in cosmological parameter inference studies (e.g. Hildebrandt et al. 2020). It is therefore expected of metacalibration to provide shear estimates that are sufficiently unbiased to begin with. In the following section, we
explore some strategies to mitigate the pixelization bias internally without requiring external simulations.

5 MITIGATING ALIASING BIAS

While the weight function $W(x)$ in the definition of moments in equation (1) is arbitrary in principle, a careful choice of the weight function has to be made in practice. We showed in the previous section that shapes of best-fitting Gaussians to the galaxies, as well as the ellipticities computed from moments using those Gaussians as weight functions show significant aliasing biases. Thus, the choice of $W(x)$ that maximizes the SNR of the shear estimate is unfortunately not ‘optimal’ due to the bias it incurs when combined with METACALIBRATION. We emphasize that this is not because of a mismatch between the light profiles, and can therefore be expected for all model-fitting methods, which share similar underlying principles (Simon & Schneider 2017). In this section, we explore if moment-based shapes estimated with a different choice of Gaussian parameters could lead to unbiased shear, although with a slightly lower SNR.

5.1 Non-adaptive weight functions

We argue in Appendix B that moments calculated with wide weight functions should be less susceptible to aliasing, as long as the sampling factor is greater than 0.5. We restrict ourselves to the case of circular weight functions only, as elliptical weight functions respond strongly to the applied shear whereas the size of the circular weight function is rather robust to the applied shear. The only free parameter for the circular weight function is its width, which we vary. We investigate three different ways of employing a large weight function in practice:

(i) use a constant weight function, that is sufficiently sampled, for all galaxies (used by Hoekstra et al. 2021);
(ii) scale the width of the adaptive weight function by a constant;
(iii) set a minimum threshold for the width of the weight function.

We emphasize that these choices can be made using the available data themselves, without the need for external simulations. For the remainder of the paper, we focus only on the KSB shear measurement algorithm as the other PSF-correction methods discussed in Section 3.3 do not have the freedom to choose the weight function.

5.1.1 Constant weight function

We begin with the simplest scheme of using wide weight functions: a circular Gaussian with a constant width for all galaxies in our sample. This has the advantage of avoiding calculating the adaptive width for each galaxy and is computationally less expensive. We employ the same circular Gaussian function of width $\sigma_w$ for all galaxies in the sample, and for all artificially sheared images of them, by specifying the `kab_sig_weight` keyword in GALSim. Figure 5 shows how the bias varies as we change $\sigma_w$. A choice of small $\sigma_w$ leads to a residual bias in both $m_1$ and $m_2$ components, with $m_1$ being larger in amplitude. A nonnegligible level of bias is found in $m_2$ for $\sigma_w \lesssim 0.01$. The non-zero $m_2$ is because of accentuated aliasing bias even in the g2 shear component.

We find that $\sigma_w$ has to be at least 0.015 (for our choice of PSF) for the aliasing bias to be within the allowed uncertainty. For our simulated galaxies, $\sigma_w = 0.15$ ($\sigma_w = 0.2$) is larger than their adaptive sizes for more than 62 (38) per cent of the galaxies. Hoekstra et al. (2021) use $\sigma_w = 0.05$ in their METACALIBRATION set-up (without any PSF-correction) and find no significant aliasing bias with more complex simulations.

5.1.2 Constant scaling

An alternative approach is to scale the width of the adaptive weight function by a constant factor for each galaxy, i.e., $\sigma_w = f \sigma_{\text{adapt}}$. We do this using the `kab_sig_factor` keyword in GALSim. Figure 6 shows how the bias changes as a function of $f$. Interestingly, we notice that the bias changes steeply around $f = 1$ and approaches $|m_1| < 2 \times 10^{-3}$ at around 1.3 and flattens. The smallest bias ($m_1 \approx -6.6 \times 10^{-3}$) was achieved for $f = 1.7$ and the bias increases ever so slightly thereafter. The latter trend is true for the control branch as well, indicating that using a very large weight function causes intrinsic bias (perhaps due to truncation) rather than aliasing bias.

Since the adaptive weights themselves are biased, we investigate if the bias could be further reduced by not recomputing the weight function for the sheared versions of the galaxy. For a given galaxy image, we use the adaptive (circular) moments from the interleaved image to give us a good initial estimate for the width of the weight function. We re-measure galaxy shapes with the KSB method, but now with the same weight function (per galaxy). We find that the estimator is biased more compared to the weight function inferred after the METACALIBRATION procedure. This is understandable as the galaxy is re-convolved with a larger PSF post METACALIBRATION and therefore the former results in a slightly wider weight function compared to the one obtained from the interleaved image, and hence has smaller bias.

5.1.3 Minimum thresholding

Since the use of a constant weight function leads to loss in signal-to-noise for large galaxies that do not contribute to aliasing bias, we
devise a hybrid scheme that treats small galaxies differently from the large ones. We propose that the width of the weight function be \( \sigma_w = \max(\sigma_{\text{min}}, \sigma_{\text{adap}}) \), where \( \sigma_{\text{adap}} \) is the width of the circular adaptive weight function and \( \sigma_{\text{min}} \) is a free parameter to be optimized. We use the KSB algorithm to estimate shear from the ensemble of galaxies using circular Gaussian weight functions with several different values of \( \sigma_{\text{min}} \). Figure 7 indicates the multiplicative bias post-metacalibration as a function of \( \sigma_{\text{min}} \). Beyond \( \sigma_{\text{min}} \) of 0\(^{\prime}\)15, the bias appears to saturate to a small value that is consistent with zero at the 10\(^{-4}\) level, and is in excellent agreement with the control branch as well. Comparing with the results in Fig. 5, it appears that by avoiding Gaussian weight functions smaller than 0\(^{\prime}\)15, we can bring aliasing bias consistent with zero. The exact threshold certainly depends on the size of the PSF (and to some extent the galaxy size distribution), which we analyse next.

### 5.2 Robustness to galaxy colour

Since the diffraction-limited PSF depends on the galaxy colour, the optimal choices for \( \sigma_w \) determined above to mitigate the aliasing bias have to be insensitive over a broad range of PSFs in order for these mitigation strategies to be more generally applicable. In particular, they should hold well for PSFs smaller than the one we have considered. In reality, the PSF is the integral of monochromatic PSFs weighted by the SED of the galaxy and filter throughput. The band-limit is then set by the bluest wavelength determined by the SED of the galaxy and VIS passband. We repeat the simulations of the same population of galaxies with Airy PSFs with effective wavelengths \( \lambda_{\text{eff}} \) of 700 nm and 550 nm. For these shorter wavelength PSFs, we increased the PSF image oversampling from a factor of 5 to 10 to construct a more accurate interpolated PSF model.

We show the results only for the minimum thresholding strategy as we expect it to be the most robust. From Fig. 8, we see only a weak dependence on the effective wavelength, with shorter effective wavelengths showing slightly larger biases for a given minimum \( \sigma_w \) as expected. We found the PSF-dependence to be negligible with other strategies as well. Note that 550 nm is the shortest wavelength that will be covered with the Euclid VIS instrument, and hence this case represents a scenario worse than any realistic worst possible case. It is also reassuring that our mitigation strategy works for strongly undersampled PSFs (\( Q = 0.47 \) for \( \lambda_{\text{eff}} = 550 \) nm).

One additional complication that we only consider briefly in Appendix C is that of realistic galaxy morphologies. Star-forming...
galaxies that suffer more from aliasing due to their blue colour also exhibit features such as knots and spiral arms that boost the power in the high-frequency modes, thus aggravating the impact of aliasing. We show in Appendix C that our mitigating strategies are effective for achromatic galaxies. We expect our mitigation strategies to be effective even in the chromatic case, but this is left for future work.

While the true profiles of the PSFs may be different from the ones considered in the study, they are nevertheless band-limited. This is true in the presence of optical aberrations. Optical imperfections such as guiding errors or internal reflections from dichroic coating result in imperfect knowledge of the PSF, but do not affect the band-limit. Out-of-band wavelengths shorter than 550 nm increase the band-limit by a small amount. On the other hand, non-ideal instrumental effects such as nonlinear and non-uniform pixel response in real-space that are not perfectly corrected for can mix the various frequency modes thereby introducing a small amount of power at arbitrarily high spatial frequencies. Furthermore, the interpolated image from a finite-size stamp of the PSF itself is not strictly band-limited. However, since we deconvolve the image with a perfectly band-limited PSF, these are not major concerns.

Due to the wavelength dependence, the PSF could vary spatially within a galaxy due to the spatial variation of the SED across the galaxy, referred to as colour gradients (Voigt et al. 2012). This introduces shear biases larger than what Stage IV experiments can tolerate when a spatially averaged PSF is used to deconvolve the galaxy image. Recently, Er et al. (2018) showed that metacalibration, by means of deconvolving with a wavelength-averaged PSF, cannot fully eliminate biases due to colour gradients, which alone happens to exceed this requirement (Semboloni et al. 2013). One way around this limitation, at least for moment-based methods, is to use large weight functions that can trade colour gradient bias for noise bias, which can then be removed by metacalibration in principle (Semboloni et al. 2013; Er et al. 2018; Kamath et al. 2020). It is intriguing (and comforting) to see that both aliasing bias and colour gradient biases can be minimised by the same procedure of using a wider weight function, although for different reasons. The low sensitivity to PSFs in Fig. 8 indicates that not only are our strategies independent of the colour of the galaxy, but should be robust to any colour gradients in the galaxy as well.

5.3 Robustness to pixel noise

While noiseless image simulations are sufficient to show that metacalibration suffers from aliasing bias when adaptive moments are used, the mitigation strategies discussed in this section need to be validated on more realistic images with pixel noise in order to demonstrate their usefulness in practice. In particular, the use of wide weight functions could lead to noisier ellipticities and reweighting. In the presence of noise, the shape measurement failures increased from none to up to 5 per cent, thereby somewhat changing the selection function. The change in the bias then has (at least) two new components - a selection bias component and a noise bias component. We also found that the measurement failures had weak dependence on galaxy size. We quantify the change in bias due to a change in the selection function by selecting those galaxy pairs from the noiseless simulations that have valid shape measurements when the SNR is a given value \( \nu \). We see from Fig. 9 that as the level of noise increases, imposing the selection on noiseless simulations increases the bias in the positive direction.

The presence of a large residual selection bias is not an inherent limitation of metacalibration. Sheldon & Huff (2017) prescribe how to account for this selection bias, which we have omitted from our analysis for simplicity. A more detailed analysis of selection bias, including those arising from object detection, is presented in Hoekstra et al. (2021).

In Table 2, we report the shear multiplicative biases with various weighting schemes from images with finite SNRs. We find that the hybrid scheme (Section 5.1.3) is most robust in the presence of

\[
\nu := \frac{1}{N_{\text{side}} \sigma_{\text{pix}}} \sum_x I^2(x),
\]

where \( \sigma_{\text{pix}} \) is the root mean square (rms) of the noise per pixel and \( I(x) \) stand for the pixel values of an \( N_{\text{side}} \times N_{\text{side}} \) image. This definition assumes the use of a matched weight function, which does not hold true as we vary the weight function. Here, we use \( \nu \) as a proxy for the quality of the galaxy image. For an image consistent with noise, \( \nu \approx 1 \). We vary the noise rms for each galaxy (using \texttt{galSim.Image.addNoiseSNR} function) so that each galaxy in the sample has a given value of \( \nu \).

Since our bias estimate with the small sample is too noisy in the presence of pixel noise, we made the following changes to our simulation set-up described in Section 3. We use an input shear of larger magnitude (\( \gamma = -0.05 \) – \( 0.05 \)) in order to reduce the uncertainty in our estimate of the multiplicative bias. Based on the values in Table 1, we expect the higher-order shear contribution to increase the multiplicative bias by around 0.001, much smaller than the aliasing bias. We also augment our sample by rendering each galaxy pair thrice with different noise realizations, which gives us small enough uncertainties that are comparable to the requirements of the Stage IV surveys. We continue to use a stimulus shear of magnitude 0.01. To correct for noise correlations introduced during this shearing process, we follow Sheldon & Huff (2017) in applying a noise image sheared in opposite direction.

As in the noiseless case, we take a simple mean of ellipticities and responsivities, and hence, each galaxy contributes equally to the shear estimate. The justification is that we do not mix galaxies of different SNRs, because the galaxy weights within a sample are expected to be roughly the same\(^{13}\). Having identical weights has the added advantage that the rotated pairs have the same weights as well, thereby ensuring effective shape noise cancellation.

The uncertainties in the estimates of the shear, and hence in \( m_1 \), is calculated resampling the pairs of galaxies that are rotated by 90°, which also ensures shape noise cancellation within each bootstrap realization. We are unable to use the approach of Pujol et al. (2019) or the more common linear regression approach since we simulated images with only one value of true shear. The bootstrapped shear estimates are Gaussian distributed to a very good approximation. The 1σ error bars correspond to the 16 - 84 percentile confidence interval over the bootstrap samples.

In the presence of the noise, the shape measurement failures increased from none to up to 5 per cent, thereby somewhat changing the selection function. The change in the bias then has (at least) two new components - a selection bias component and a noise bias component. We also found that the measurement failures had weak dependence on galaxy size. We quantify the change in bias due to a change in the selection function by selecting those galaxy pairs from the noiseless simulations that have valid shape measurements when the SNR is a given value \( \nu \). We see from Fig. 9 that as the level of noise increases, imposing the selection on noiseless simulations increases the bias in the positive direction.

\(^{13}\) This is not strictly true, as \( \nu \) is detection significance, and galaxies have different shape errors based on their intrinsic properties.
and marginally so for the multiplicative biases are estimated using equation (14). The schemes that mitigated aliasing bias sufficiently in noiseless images. The absence of explicit PSF-correction. et al. (2021) find the use of constant weight functions robust in the absence of PSF-correction steps within the KSB weight function (Section 5.1.2). We do not report the values from selection bias on markers. The dashed line shows the selection bias on \( m_2 \). The differences seen at low values of \( \sigma_{\text{min}} \) are due to aliasing.

Figure 9. Multiplicative bias due to selection from noisy images applied to noiseless measurements. The selection bias increases in the positive direction as the noise increases, fairly independent of minimum size of the weight function, \( \sigma_{\text{min}} \). The uncertainty in selection bias is smaller than the size of markers. The dashed line shows the selection bias on \( m_2 \). The differences seen at low values of \( \sigma_{\text{min}} \) are due to aliasing.

Table 2. Multiplicative biases in the presence of noise for a selected set of schemes that mitigated aliasing bias sufficiently in noiseless images. The multiplicative biases are estimated using equation (14).

| KSB weight | \( \nu \) | \( m_1 \) \([\times 10^2]\) | \( m_2 \) \([\times 10^2]\) |
|------------|---------|-----------------|-----------------|
| \( \sigma_w = \sigma_{\text{adap}} \) | 50     | \(-0.82 \pm 0.08\) | \(-0.16 \pm 0.08\) |
| \( \sigma_w = 1.4\sigma_{\text{adap}} \) | 50     | \(-0.22 \pm 0.09\) | \(-0.38 \pm 0.09\) |
| \( \sigma_{\text{min}} = 0^\prime.15 \) | 50     | \(-0.01 \pm 0.08\) | \(-0.1 \pm 0.08\) |
| \( \sigma_w = \sigma_{\text{adap}} \) | 40     | \(-0.86 \pm 0.10\) | \(0.02 \pm 0.10\) |
| \( \sigma_w = 1.4\sigma_{\text{adap}} \) | 40     | \(-0.13 \pm 0.11\) | \(-0.08 \pm 0.12\) |
| \( \sigma_{\text{min}} = 0^\prime.15 \) | 40     | \(-0.03 \pm 0.10\) | \(0.08 \pm 0.10\) |
| \( \sigma_w = \sigma_{\text{adap}} \) | 30     | \(-1.03 \pm 0.13\) | \(-0.22 \pm 0.14\) |
| \( \sigma_w = 1.4\sigma_{\text{adap}} \) | 30     | \(-0.45 \pm 0.16\) | \(-0.35 \pm 0.15\) |
| \( \sigma_{\text{min}} = 0^\prime.15 \) | 30     | \(-0.19 \pm 0.14\) | \(-0.21 \pm 0.14\) |
| \( \sigma_w = \sigma_{\text{adap}} \) | 20     | \(-1.47 \pm 0.21\) | \(-0.36 \pm 0.27\) |
| \( \sigma_w = 1.4\sigma_{\text{adap}} \) | 20     | \(-2.41 \pm 0.86\) | \(-5.20 \pm 2.02\) |
| \( \sigma_{\text{min}} = 0^\prime.15 \) | 20     | \(-0.83 \pm 0.21\) | \(-0.34 \pm 0.27\) |

Figure 10 shows the bias in the shear estimated from our meta-calibration set-up with the hybrid scheme. Note that the data points are highly correlated, since the measurements are made from the same set of noisy images. The characteristic shape of aliasing bias is seen in \( m_1 \) values when \( \sigma_{\text{min}} \lesssim 0^\prime.125 \). The signature of aliasing bias is absent for \( m_2 \), as expected from the noiseless case, and is largely independent of the choice of \( \sigma_{\text{min}} \). The values of \( m_1 \) and \( m_2 \) are also consistent with each other for \( \sigma_{\text{min}} \gtrsim 0^\prime.125 \), and marginally so for the \( \nu = 20 \) sample. The uncertainty in the \( m \) values are smaller than the tolerance for \( \nu \gtrsim 30 \) galaxy samples. The values of \( m_2 \) (and \( m_1 \) for large enough \( \sigma_{\text{min}} \)) are also consistent with 0 for high SNR (\( \nu \gtrsim 30 \)) samples, only barely consistent with 0 for \( \nu = 30 \) sample and inconsistent with 0 for the \( \nu = 20 \) sample.

Though the performance of KSB with metacalibration falls short of the requirement for the \( \nu = 20 \) sample, the mitigation strategy reduces the bias by about 0.008 for all SNR samples. The residual bias for the low SNR sample may be due to selection effects, correlations in the noise etc. With further work, we believe it should be possible to reduce the bias further for \( \nu = 20 \), but we leave it outside the scope of this paper.

Despite the simplified nature of the simulations used in this work, Fig. 10 makes evident the presence of aliasing bias and the effectiveness of using a wider weight function in mitigating it. One may suspect that in regions where galaxies are clustered, the ringing effects from one galaxy could affect the shape measurements of neighbouring galaxies, which our simulations do not capture. However, we do not find any evidence of significant aliasing bias in Hoekstra et al. (2021) who simulate a scene at native pixel scale with multiple galaxies and use a wide weight function for shear measurement. Thus, we conclude that undersampling is not a limiting factor to use metacalibration for measuring shear using individual exposures from space telescopes if appropriate mitigation measures are taken.

6 DISCUSSION

The problem of the pixel scale being smaller than the Nyquist scale of the PSF is a fundamental one in estimating the weak lensing signal, as it renders all galaxy images undersampled in principle. In the previous section, we demonstrated a set of related strategies that can sufficiently suppress the effects of aliasing with metacalibration. In this section, we layout a few other ways of tackling this problem that may be explored further in the future.

First, we revisit the self-imposed restriction that the metacalibration procedures are to be performed on individual exposures, and not on the coadded image (see Section 3.2). Our analysis in Fig. 4 shows that the metacalibration procedure works, at least in principle, if performed on a coadded image. This is due to a smaller effective pixel scale compared to the native pixel scale. For this study, we considered uniformly dithered exposures with identical PSFs, which are effective at cancelling aliasing effects due to special symmetries. It is not obvious in the more generic case of variable PSFs and random subpixels dither if the coadded image would be free of aliasing. This is made worse when the exposures have to undergo generic geometric transformations, such as rotation and corrections for field distortions in addition to translations before being projected on to a common grid. A commonly used method to produce a coadded image from the individual HST exposures is the MultiDrizzle method (Fruchter & Hook 2002; Koekemoer et al. 2003), which generates a high-resolution image by means of interpolation on the individual (non-Nyquist sampled) exposures, and therefore shapes measured on them with metacalibration would exhibit some aliasing bias.

Since the coadded image is a linear combination of the individual exposures, one could construct a corresponding coadded PSF using the same linear combination and use that for deconvolution. Metacalibration has been shown to work on such coadds that are linear combination of individual exposures with PSFs that have smooth spatial-variation at the level expected for the LSST survey (Sheldon et al. 2020, Armstrong et al. in preparation). The caveat however is
that, with hundreds of exposures as in LSST, the issue of PSF discontinuity could be overcome by discarding the few exposures (around 3 per cent of exposures) that contribute to the discontinuity for a given galaxy with very little loss of SNR. Such an approach may not be feasible with Euclid with five (or fewer) exposures per field. But fewer exposures also means that regions with PSF discontinuities due to chip gaps are also much rarer. Additionally, pixels that are masked or affected by cosmic rays also lead to PSF discontinuities. However, with variable PSFs, each PSF image would be undersampled, and hence the coadded PSF image would also be aliased. Alternatively, it is possible to construct a coadded image with a desired target PSF, as in the IMage COMbination algorithm (mcCOM; Rowe et al. 2011). This introduces additional complications to the noise properties, which have to be dealt with carefully. Evaluating the performance of metacalibration on coadded images where the individual PSFs are varying and undersampled is therefore still an open question.

Second, the requirement for the PSF to be Nyquist sampled arises only if we wish to reconstruct the continuous image from the pixel values without any assumption about the galaxy profile. This is different from the requirement of having sufficient knowledge of PSF, which can be obtained from forward modelling of the telescope optics in addition to multiple star images. It is possible to overcome the Nyquist theorem by assuming a moderately realistic model for galaxies, as the model-fitting methods do. Such methods assume a galaxy model with a few parameters and attempt to find the best-fitting model parameters by maximizing the likelihood that the observed image is given by a convolution of the model and the PSF. A similar approach should be possible even when the PSF is undersampled. Given an aliased image after the metacalibration procedures, by fitting to each exposure an aliased model convolved with an aliased PSF model, the best-fitting parameters could still be obtained, at least in principle, by likelihood-based methods. Additionally, likelihood-based methods also enjoy adopting an accurate description of correlated noise introduced in the image processing steps. However, a plausible difficulty while trying to implement such an approach is that the continuous image becomes periodic, with a periodicity comparable to pixel width. This could hinder likelihood methods from achieving convergence in the presence of noise. Tight priors on the centroid of the galaxy would be required to prevent diverging solutions. In this paper, we did not use any likelihood-based model-fitting methods and are hence defer testing this hypothesis to the future.

Third, if one were to employ inverse variance weights, or weights that are smooth monotonically increasing functions of galaxy sizes and use equation (12), the amplitude of the aliasing bias may be reduced further. This could help achieve a compromise between the systematic and statistical errors. It may even be possible to conceive a somewhat sophisticated colour-dependent weighting scheme for galaxies so that blue galaxies are downweighted more compared to red galaxies (of the same SNR and size) due to their small PSF size. Such weighting schemes are not without caveats. A colour-dependent weight couples errors in photometric colours (and redshifts) with errors in shear which could be difficult to decouple later on. Also, in the presence of noise, the estimated size of the galaxies will themselves be noisy and a selection based on these noisy quantities could introduce biases comparable to aliasing bias. Moreover, such weighting schemes might preferentially downweight galaxies at high redshifts, thereby potentially reducing the figure of merit for the dark energy equation of state parameters significantly. A thorough investigation of optimal galaxy weights to minimise the overall bias requires far more realistic image simulations and is beyond the scope of this work.

Figure 10. Plot of change in multiplicative bias with change in $\sigma_{\text{min}}$ parameter for different values of the SNR $\nu$. The shaded regions represent the tolerance of $\pm 2 \times 10^{-3}$. The change in bias on noisy images closely follows the trend in the noiseless images for $\sigma_{\text{min}} \lesssim 0.15$, suggesting the validity of the mitigation strategy even in the presence of pixel noise.

7 SUMMARY

The shear field due to the gravitational lensing of the large-scale structure of the Universe contains rich information about the distribution of dark matter and the rate at which it forms the cosmic web. Metacalibration is a promising way to calibrate the shear from observed images, without having to rely on external image simulations for calibration. In the case of space-based surveys such as Euclid and the Roman Space Telescope, this procedure is compli-
cated by the fact that one has to interpolate a possibly undersampled image. Galaxies that are small and barely resolved lead to a significant aliasing bias as a result of this interpolation. Nevertheless, by using moments of those galaxies measured with a weight function wider than their adaptive weight function for estimating shear, it is possible to reduce the amplitude of this bias to a tolerable level, at the mild expense of reducing the SNR. Thus, assuming that the PSF for the individual exposures are well known, we conclude that METACALIBRATION, even when applied to individual exposures, is not limited by undersampling. We understand this conclusion as a result of suppression of high-frequency modes which would otherwise masquerade as low-frequency modes when interpolated.

Colour-gradient bias, an effect arising due to wavelength dependent PSF and spatial variation of SED within a galaxy, is also mitigated by using similarly wide weight functions for a different reason (Er et al. 2018). This chromatic effect is not unique to space-based surveys, and is also relevant to LSST (Meyers & Burchat 2015). Thus, sampling is the main fundamental difference between space-based and ground-based telescopes. With the strategy proposed in this paper to mitigate the bias due to sampling, the success of METACALIBRATION for Euclid must follow the success for ground-based surveys as well. Our results therefore encourage the possibility of METACALIBRATION as an independent shear measurement pipeline, which can validate the lensing shear signal measured using methods that are already existing or under development for Euclid.

ACKNOWLEDGEMENTS
The authors are grateful to Erin Sheldon and the developers of GALSIM for making their software packages publicly available. The authors thank Patricia Liebing and Lance Miller for helpful conversations about Euclid PSFs and Peter Melchior and Erin Sheldon for other technical discussions, and Jason Rhodes and the referee Michael Troxel for carefully reading the manuscript and suggesting several improvements. ER thanks the LEAPS programme14 and its organizers. This work was supported by the Netherlands Organisation for Scientific Research (NWO) through grant 639.043.512, and by the National Science Foundation under Cooperative Agreement 1258333 managed by the Association of Universities for Research in Astronomy (AURA), and the Department of Energy under Contract No. DE-AC02-76SF00515 with the SLAC National Accelerator Laboratory. Additional funding for Rubin Observatory comes from private donations, grants to universities, and in-kind support from LSSTC Institutional Members.

DATA AVAILABILITY
The Python scripts to reproduce the simulations and analyses may be requested from the corresponding author. The external libraries and data products used in this article are publicly available.

REFERENCES
Aihara H., et al., 2018, PASJ, 70, S4
Akeson R., et al., 2019, arXiv e-prints, p. arXiv:1902.05569
Albrecht A., et al., 2006, arXiv e-prints, pp astro-ph/0609591
Antilogus P., Astier P., Doherty P., Guyonnet A., Regnault N., 2014, Journal of Instrumentation, 9, C03048
Asgari M., et al., 2021, A&A, 645, A104
Bartelmann M., Schneider P., 2001, Physics Reports, 340, 291
Bernstein G., 2010, MNRAS, 406, 2793
Bernstein G., Jarvis M., 2002, AJ, 123, 583
Bosch J., et al., 2018, PASJ, 70, 55
Bridle S., et al., 2009, Annals of Applied Statistics, 3, 6
Carlsten S. G., Strauss M. A., Lupton R. H., Meyers J. E., Miyazaki S., 2018, MNRAS, 479, 1491
Coulton W. R., Armstrong R., Smith K. M., Lupton R. H., Spergel D. N., 2018, AJ, 155, 258
Dawson W. A., Schneider M. D., Tyson J. A., Jee M. J., 2016, ApJ, 816, 11
Dore O., et al., 2018, arXiv e-prints, p. arXiv:1804.03628
Eckert K., et al., 2020, MNRAS, 497, 2529
Er X., et al., 2018, MNRAS, 476, 5645
Euclid Collaboration et al., 2019, A&A, 627, A59
Fenech Conti I., Herbonnet R., Hoekstra H., Merten J., Miller L., Viola M., 2017, MNRAS, 467, 1627
Fruchter A. S., Hook R. N., 2002, PASP, 114, 144
Gurvich A., Mandelbaum R., 2016, MNRAS, 457, 3522
Heymans C., et al., 2006, MNRAS, 368, 1323
Hikage C., et al., 2019, PASJ, 71, 43
Hildebrandt H., et al., 2020, A&A, 633, A69
Hirata C., Seljak U., 2003, MNRAS, 343, 459
Hoekstra H., Jain B., 2008, Annual Review of Nuclear and Particle Science, 58, 99
Hoekstra H., Franx M., Kuijken K., Squires G., 1998, ApJ, 504, 636
Hoekstra H., Herbonnet R., Muzzin A., Babul A., Mahdavi A., Viola M., Cacciato M., 2015, MNRAS, 449, 685
Hoekstra H., Viola M., Herbonnet R., 2017, MNRAS, 468, 3295
Hirata C., Seljak U., 2003, MNRAS, 343, 459
Kacprzak T., Zuntz J., Rowe B., Bridle S., Refregier A., Amara A., Vogt L., Hirsch M., 2012, MNRAS, 427, 2711
Kaiser N., 2000, ApJ, 537, 555
Kaiser N., Squires G., Broadhurst T., 1995, ApJ, 449, 449
Kamath S., Meyers J. E., Burchat P. R., (LSST Dark Energy Science Collaboration), 2020, ApJ, 888, 23
Kannawadi A., Mandelbaum R., Lackner C., 2015, MNRAS, 449, 3597
Kannawadi A., Shapiro C. A., Mandelbaum R., Hirata C. M., Krak J. W., Rhodes J. D., 2016, PASP, 128, 095001
Kannawadi A., et al., 2019, A&A, 624, A92
Kilbinger M., 2015, Reports on Progress in Physics, 78, 86901
Kitching T., et al., 2011, Ann. Appl. Stat., 5, 2231
Koeckemoer A. M., Fruchter A. S., Hook R. N., Hack W., 2003, in HST Calibration Workshop: Hubble after the Installation of the ACS and the NICMOS Cooling System. p. 337
Koeckemoer A., et al., 2007, ApJS, 172, 196
LSSST Science Collaboration et al., 2009, arXiv e-prints, p. arXiv:0912.0201
Lackner C., Gunn J., 2012, MNRAS, 421, 2277
Lauer T. R., 1999, PASP, 111, 227
Laureijs R., et al., 2011, arXiv e-prints, p. arXiv:1110.3193
Leauthaud A., et al., 2007, ApJS, 172, 219
Luppino G. A., Lauer T. R., 1999, PASP, 111, 229
Mandelbaum R., Hirata C., 2008, Annual Review of Nuclear and Particle Science, 58, 99
http://leaps.strw.leidenuniv.nl/
APPENDIX A: IMPACT DUE TO THE CHOICE OF THE INTERPOLATION KERNEL

In principle, when a discretized signal is convolved with a sinc kernel, it reproduces the original signal in the continuous domain, provided that the signal was Nyquist-sampled. However, such perfect interpolation is computationally too expensive and other approximate interpolation kernels may need to be used in practice. Moreover, due to finite size of the postage stamp, the signal is never perfectly band-limited and hence is not ideal to begin with. Table A1 shows a significant increase (about 100 fold) in run times when sinc interpolation is used instead of Lanczos kernels.

For a representative subsample of 500 noiseless galaxy pairs, we show the effect of choosing interpolation kernel on the measured shear in Fig. A1. The Quintic interpolation scheme, which is the default option in GaLM, could potentially introduce biases up to a few percent in both components. However, with the sinc kernel, we find that $m_2$ is fully consistent with zero but $m_1$ shows biases about $-0.01$. We obtain very similar values with Lanczos kernels for very high orders (greater than 75). For lower order Lanczos kernels, the value of $m_1$ becomes somewhat smaller but still outside the requirements. The decrease in the bias may be attributed to the lower order kernels having a support that is much smaller than typical postage stamp sizes. However, for a metadetection-

| Interpolation kernel | Average time in seconds per galaxy per exposure |
|----------------------|-----------------------------------------------|
| Quintic              | 6.8 ± 2.0                                     |
| Lanczos3             | 7.0 ± 1.8                                     |
| Lanczos5             | 7.3 ± 1.8                                     |
| Lanczos10            | 6.8 ± 1.9                                     |
| Lanczos15            | 7.3 ± 2.5                                     |
| Lanczos50            | 6.4 ± 1.5                                     |
| Lanczos75            | 9.3 ± 2.6                                     |
| Lanczos100           | 7.3 ± 1.9                                     |
| sinc                 | 808.7 ± 165.4                                 |

Figure A1. The multiplicative bias $m$ introduced in the shear measured from undersampled images with metadcalibration for Quintic, Lanczos, and sinc interpolation kernels.

like set-up as in Hoekstra et al. (2021), we expect the higher order kernels do perform better. For this reason, we chose to use the 50th order Lanczos kernel, which is closer to sinc despite it being suboptimal.

APPENDIX B: QUADRUPOLE MOMENTS IN THE FOURIER DOMAIN

The quadrupole moments defined in equation (1) may be expressed as

$$Q_{ij} = -\frac{\partial}{\partial k_i} \frac{\partial}{\partial k_j} \int d^2 x W(x) I(x) e^{-i k \cdot x} \bigg|_{k=0}. \quad (B1)$$

The integral is the Fourier transform of $W(x) I(x)$ which is given by $(\hat{W} \otimes \hat{I})(k) := \int d^2 k' \hat{W}(k - k') \hat{I}(k')$, where $\hat{W}$ is the Fourier transform of the weight function and $\hat{I}$ is the Fourier transform of the image. In terms of the Fourier transforms of the weight functions and images, we can express the quadrupole moments as

$$Q_{ij} = -\frac{\partial}{\partial k_i} \frac{\partial}{\partial k_j} \int d^2 k' \hat{W}(k - k') \hat{I}(k') \bigg|_{k=0} = -\int d^2 k' \hat{W}_{ij}(-k') \hat{I}(k') \bigg|_{k=0} = \hat{W}_{ij}(k) \hat{I}(k).$$

where $\hat{W}_{ij}(k) := \frac{\partial}{\partial k_i} \frac{\partial}{\partial k_j} \hat{W}(k)$. We deliberately make $\hat{I}$ a function of only the variable of integration so that the derivatives are taken not on the image $\hat{I}$ but on the weight function which has an analytic
form typically. Using the fact that the weight function is real-valued, we can express $\tilde{W}(-k^\prime) \equiv \tilde{W}^*(k^\prime)$, where $*$ denotes complex conjugation. Furthermore, if the weight function has even parity (in real space), $\tilde{W}^*(k^\prime) \equiv \tilde{W}(k^\prime)$. This completes our recasting of the quadrupole moments as integrals in Fourier space.

For Gaussian weight functions, the quadrupole moments may also be expressed as an integral over Fourier modes as

$$Q_{ij} = -\int_{|k| < k_{\text{max}}} d^2k \, q_{ij}(k) \tilde{W}(k) \tilde{I}(k),$$  \hspace{1cm} \text{(B3)}$$

where $q_{ij}(k)$ is an appropriate quadratic expression, whose details we need not concern ourselves with. The integral is restricted to the disc of radius $k_{\text{max}}$ as $\tilde{I}(k) \equiv 0$ outside the disc by definition. We find that the form of the quadrupole moments expressed in equation (B3) to be more convenient for this work. Note that this is different from Fourier moments proposed by Zhang (2008).

The weighted moments measured post-metacalibration are

$$Q_{ij}^{\text{interp}} = -\int_{|k| < k_{\text{max}}} d^2k \, q_{ij}(k) \tilde{W}(k) \tilde{f}^{\text{interp}}(k)$$

$$= -\sum_i \int_{|k| < k_{\text{max}}} d^2k \, q_{ij}(k) \tilde{W}(k) \tilde{I}(k + n\Delta k),$$ \hspace{1cm} \text{(B4)}

$$= -\int_{|k| < k_{\text{max}}} d^2k \, q_{ij}(k) \tilde{W}(k) \tilde{I}(k + n\Delta k)$$ \hspace{1cm} \text{(B5)}

where $\tilde{f}^{\text{interp}}(k)$ is given in equation (13). When the PSF is undersampled, i.e. $\Delta k < 2k_{\text{max}}$, there is an undesirable contribution to $Q_{ij}^{\text{interp}}$ from some of the $n \neq 0$ terms. In particular, for small galaxies, the contribution to $Q_{ij}^{\text{interp}}$ from $n \neq 0$ terms is significant not only due to large amplitude of $\tilde{I}(k)$ but also due to that of the matched weight function $\tilde{W}(k)$ for $|k| > k_{\text{max}}$. The amplitude of these contributions can be reduced by choosing a weight function that is narrow in Fourier space (equivalently, wide in real space), thereby reducing the amplitude of $q_{ij}(k) \tilde{W}(k)$. Thus, the deviation of $Q_{ij}^{\text{interp}}$ from its true value is small when the weight functions are wide.

**APPENDIX C: VALIDATION ON GALAXY IMAGES WITH REALISTIC MORPHOLOGIES**

Here, we show that the mitigating strategies discussed in Section 5 are effective even for galaxies with complex morphologies. Instead of using the best-fitting parameters to the HST COSMOS galaxies, we use the PSF-deconvolved real galaxy images by setting \texttt{real=True} for \texttt{COSMOSCatalog} in \texttt{GalSim}. We refer the reader to Mandelbaum et al. (2012) for details on how the images were created. We continue to use an Airy PSF with $\lambda_{\text{eff}} = 800$ nm. The remainder of our set-up is the same as described in Section 3.1.

We take a different approach in demonstrating that our mitigating strategies work even when the galaxies have realistic morphologies. The measured shears have a much larger fraction of outliers, which prevents us from robustly estimating the values of the shear calibration parameters. We therefore simply show the existence of aliasing bias from the distribution of estimated shear values rather than tabulating the values of the multiplicative bias as before.

We average the shear estimated from each galaxy with its rotated pair to cancel out the intrinsic shape noise and plot the distribution of shear per galaxy pair in Figure C1. From the upper panels in Figure C1, we see that for a large majority Sérsic galaxies, the shear is measured accurately. The one-sided tail, occupied by galaxies with half-light radii smaller than 0.7′, had introduced the bias in the ensemble shear estimates. In contrast, for realistic galaxies, while the peak is sharply centred at true value, the distribution is two-sided. However, the tail is fatter towards positive values (blue distribution in Figure C1), thereby biasing the ensemble shear low in amplitude.

To eliminate the possible contribution due to additive biases, we repeated the simulations with the sign of the input shear flipped (orange distribution in Figure C1). We now find that the tail is fatter towards negative values. This confirms that the ensemble shear is indeed smaller in amplitude relative to the true shear, signalling a negative multiplicative bias as before. The asymmetry between the distributions for positive and negative input shear values strongly suggests the presence of additive bias, which was absent for Sérsic galaxies. Thus, our single-shear estimator given in equation 14 cannot be used to evaluate the multiplicative bias.

To demonstrate the efficacy of our mitigation strategy, we repeat the above exercise after scaling the adaptive moments of all galaxies by 1.4. From the lower panels of Figure C1, it is evident that when we use a larger weight function to estimate the shear, the distribution is much more symmetrical about the true shear value. This is true for both positive and negative values of shear components. Thus, despite the presence of a few outlier measurements, we conclude that the multiplicative bias due to aliasing can be mitigated by adjusting the width of the weight function.

This paper has been typeset from a TEX/L\LaTeX\ file prepared by the author.
Figure C1. Histogram of shear estimated from pairs of galaxies using the constant scaling method described in Section 5.1.2. The distribution of shear values after subtracting the true shear are shown in blue for real galaxies and in green for parametric galaxies. The skewness towards positive values is due to aliasing bias. To show that the skewness is multiplicative, we include shear estimates for real galaxies with the sign of the true shear values reversed (in orange) and find that the skewness is reversed as well. The lack of parity relation between the blue and orange histograms also indicates the presence of additive bias introduced due to aliasing. We see in the lower panels that use of a wider weight function mitigates this skewness.