Galaxy blending effects in deep imaging cosmic shear probes of cosmology

Erfan Nourbakhsh*, J. Anthony Tyson®, Samuel J. Schmidt®, and The LSST Dark Energy Science Collaboration

Department of Physics and Astronomy, University of California, One Shields Avenue, Davis, CA 95616

Accepted XXX. Received YYY; in original form ZZZ.

ABSTRACT

Upcoming deep imaging surveys such as the Vera C. Rubin Observatory Legacy Survey of Space and Time will be confronted with challenges that come with increased depth. One of the leading systematic errors in deep surveys is the blending of objects due to higher surface density in the more crowded images; a considerable fraction of the galaxies which we hope to use for cosmology analyses will overlap each other on the observed sky. In order to investigate these challenges, we emulate blending in a mock catalogue consisting of galaxies at a depth equivalent to 1.3 years of the full 10-year Rubin Observatory that includes effects due to weak lensing, ground-based seeing, and the uncertainties due to extraction of catalogues from imaging data. The emulated catalogue indicates that approximately 12% of the observed galaxies are “unrecognized” blends that contain two or more objects but are detected as one. Using the positions and shears of half a billion distant galaxies, we compute shear-shear correlation functions after selecting tomographic samples in terms of both spectroscopic and photometric redshift bins. We examine the sensitivity of the cosmological parameter estimation to unrecognized blending employing both jackknife and analytical Gaussian covariance estimators. A \( \Delta \sigma_8 \approx 0.2 \) decrease in the derived structure growth parameter \( \sigma_8 \) is seen due to unrecognized blending in both tomographies with a slight additional bias for the photo-z-based tomography. This bias is about 2\( \sigma \) statistical error in measuring \( \sigma_8 \).

Keywords: cosmological parameters – gravitational lensing: weak – dark matter – cosmology: observations – large-scale structure of Universe

1 INTRODUCTION

With the Legacy Survey of Space and Time¹ (LSST; Ivezić et al. 2019) to be undertaken over the next decade by the Vera C. Rubin Observatory, we will have samples of over 10 billion galaxies. No longer limited by statistical errors, systematic errors become dominant. For studies of the static sky, there is a class of systematic errors that emerge at low surface brightness. With over 800 images of each sky patch, the LSST will reach ultra-low surface brightness. At 30 mag arcsec\(^{-2} \) there are nontrivial challenges with instrument signature removal and sky determination. Most pipeline processing algorithms over-subtract sky at these levels (Ji et al. 2018). Even if a non-biased sky level can be determined, a related systematic emerges at ultra-low surface brightness: an increasing fraction of the field is occupied by galaxy faint halos.

In astronomical imaging surveys, galaxies are seen as twodimensional surface brightness objects. For galaxies that project very close to each other on the sky, a dip in surface brightness between them suffices to tag them as “recognized blends” and software can attempt to de-blend the pair or group (Melchior et al. 2018). Depending on depth, a certain fraction of galaxies are so blended as to be unrecognized as a blend. Observation of “unrecognized blends” yields common measurements for their component galaxies, affecting all quantities that are based on photometry, including most importantly for this study, shape and photometric redshift. The measurement of galaxy shapes can suffer systematic effects since the galaxies are not individually measured (Melchior et al. 2021; MacCrann et al. 2022). In ultra-deep surveys, a higher fraction of galaxies become unrecognized blends (Dawson et al. 2016). The observed galaxies in recent deep photometric surveys such as Dark Energy Survey² (DES; DeRose et al. 2021) and the Deep Lens Survey³ (DLS; Jee et al. 2016) are extracted from crowded images with a considerable fraction of them overlapping each other. This is also evident in the Subaru’s Hyper Suprime-Cam (HSC; Miyazaki et al. 2018) deep imaging. The HSC Subaru Strategic Program⁴ (HSC-
SSP; Aihara et al. 2019) is regarded as a precursor to LSST with a similar seeing.

At fainter magnitudes, the unrecognized blend fraction grows larger, and the effects of blending become more pronounced as the projected surface number density of the galaxies in a field increases. This prompted our investigation of the effects of galaxy blending on estimates of cosmology via a “realistic” synthetic catalogue, in the sense that a full N-body simulation, galaxy numbers and magnitudes, and lensing effects are included. To make this study feasible, such a catalogue must also include the forward-modelled effects of blending through a process we refer to as *emulation*. One can selectively turn on each of the components in the emulated catalogue: seeing (Point Spread Function; PSF) Full Width Half Maximum (FWHM), photometric redshift (redshift estimated via multi-band photometry), lensing magnification, and various angular and magnitude filtering schemes. However, the current work is devoted to the coupled effects of blending and photometric redshift.

To constrain cosmology with weak lensing (WL) measurements, robust observational results with per cent level precision over a wide range of scales and cosmological parameters are needed (Mandelbaum et al. 2018). Exploring shear-shear correlations across tomographic redshift bins is a powerful tool for such studies (e.g. Heymans et al. 2021; DeRose et al. 2021). Because these multiple correlations over a large volume jointly have much higher dimensionality than the 2D blending, incorporating this extra information is hoped to minimize biases in cosmological parameter estimation. The full 3D N-body simulation used in this study enables us to investigate the impact of unrecognized blending on the cosmological parameter estimation in tomographic analyses of large surveys such as LSST. We include the effect of unrecognized blending on the photometric redshifts and also study the effect of excluding from the sample the “catastrophic” recognized blends where the blend is recognized but its surface brightness profile is severely affected by the nearby galaxies beyond a threshold. This threshold set by “purity” is described in Section 7.3.

We have developed a pipeline, BlendSim, to perform the analysis described here. The paper is organized as follows. We first describe the synthetic sky catalogue and our enhancements in Section 2. We then introduce our blending emulation process in Section 3. Section 4 explains the method we use to estimate our photometric redshifts and Section 5 discusses “cosmic shear” as the utilized cosmological effects.

The simulation must be based on a credible model for large-scale structure, have a well-understood mechanism for tagging the resulting galaxies, and span a sufficient redshift range to support the cosmological analysis. Moreover, it is preferred that the simulation has more than one realization that helps with reducing the statistical noise in our analysis. Having more than one realization is also a check on reproducibility. As described in the subsection below, a catalogue of billions of faint galaxies called Buzzard meets our requirements.

A full image-based approach is prohibitively time-consuming for billions of galaxies. We bypass this procedure by simulating images for only a subset of the sky where the galaxy number density is high enough to potentially allow blending to take place. Hence, these galaxies are sufficiently close on the sky to be either recognized or unrecognized blend candidates. The catalogue used to generate these images consists of a table of galaxy properties such as position, size, ellipticity, colours, redshift, and gravitational lens shear produced by a dark matter power spectrum that evolves with redshift.

For our study, it is important to have a synthetic sky catalogue that is photometrically complete for all galaxy types to a limiting magnitude faint enough to enable realistic blending emulation and analysis. That is, the simulation must include a sufficiently high number density of galaxies so that when convolved with a seeing PSF, the surface density of galaxies in a fiducial “gold sample” galaxy catalogue (where the minimum flux $S/N$ for point sources is 20) results in noticeable blending with galaxies up to ~2 magnitudes fainter. This is expected for the “LSST gold sample” according to Dawson et al. (2016). The 2 magnitude buffer is a result of the product of the galaxy number-flux relation (for some spectral band) and the corresponding flux. The number-flux relation for galaxies in red passbands is much shallower than euclidean (stars) due to cosmological effects. Including even fainter galaxies does not have a significant impact on the apparent shape or overall flux of the blends.

2.1 Buzzard mock catalogue

For this study, we use data from the Buzzard simulation suite (DeRose et al. 2019), a full 3D cosmological simulation, which we then augment in several ways to explore the effects of blending on the inferred cosmology from data. The mock catalogues are available in HEALPix5 (Górski et al. 2005) format through the standalone Python module, GCRCata10ga5, and described in detail in DeRose et al. (2019), but we briefly summarize the key aspects below.

Buzzard is a set of synthetic galaxy catalogues tuned for the Dark Energy Survey (DES) but applicable to a variety of large-area photometric and spectroscopic surveys. The simulation employs a set of dark matter-only simulations and then utilizes the ADDGALS algorithm (Wechsler et al. 2021) to paint baryonic galaxies on these dark matter simulations. This simulation is constructed from a set of three dark matter N-body simulations run using L-GADGET2, a modified version of GADGET2 with box lengths ranging from 1.4–20 Gpc from which light-cones were constructed. The positions of mass halos are decorated with galaxies whose type, size, brightness, and redshift reflect those in the DES data. Currently, there are

---

5 Acronym for Hierarchical Equal-Area isoLatitude Pixelisation of a 2-sphere; http://healpix.sourceforge.net
6 http://github.com/LSSTDESC/gcr-catalogs
7 http://wwmpa.mpa-garching.mpg.de/gadget
two quarter-of-the-sky realizations available in total through GCR-Catalogs. Weak lensing shears are measured at galaxy positions using an algorithm (CaLCLen; Becker 2013) appropriate for full sky arc second ray-tracing. Galaxy magnitudes and shapes include the impact of shear and magnification of apparent magnitudes in many filter sets including LSST bands. Each Buzzard realization has about 3 billion galaxies before any sample cuts spanning an area of 10,313 square degrees. A number of validation tests for internal consistency of Buzzard_v2.0 have been carried out (DeRose et al. 2021). This is the version of Buzzard used in this study.

The Buzzard mock catalogue is populated with galaxies that have magnitudes and shapes designed to approximately match those observed in the DES imaging out to \( z = 2.35 \) and to a depth of the DES-like magnitude \( r \sim 26 \). To avoid potential biases associated with using deconvolved sizes from the actual DES data, the true half-light radii in Buzzard are drawn from empirical observations of galaxies imaged with the Hubble Space Telescope (HST) F125W band (J. DeRose, private communication). The spectral energy distributions (SEDs) are drawn from SDSS Data Release 6 (DR6) spectroscopic sample consisting of approximately \( 5 \times 10^5 \) galaxies. The observed distribution of SEDs at a given luminosity and spatial galaxy density is modelled in SDSS and then used to assign SEDs to simulated galaxies with similar properties.

Fig. 1 shows the true \( i \) band magnitude distribution of the Buzzard galaxies used in this work. The solid red line indicates our Buzzard cosmology sample cut which ensures \( S/N > 20 \) for point sources brighter than \( i = 24 \). The dashed orange line indicates a cut two magnitudes deeper that we impose in order to remain complete for the faint sample.

![Blending effects on cosmic shear](image)

**Figure 1.** The true \( i \) band magnitude distribution of the Buzzard galaxies used in this work. The solid red line indicates our Buzzard cosmology sample cut which ensures \( S/N > 20 \) for point sources brighter than \( i = 24 \). The dashed orange line indicates a cut two magnitudes deeper that we impose in order to remain complete for the faint sample.

### 2.2 Modified galaxy properties

Due to the fact that HEALPix pixels are assigned before lensing was applied, we truncate the sample by removing galaxies within 0.2 degrees of the edge of Buzzard’s footprint\(^8\) to avoid issues associated with overdense and underdense regions near the edge. This leaves us with \( \sim 10,254 \) square degrees of the simulation for this study. We modify the truth-catalogue properties of all faint sample \((i < 26)\) galaxies to more accurately reflect LSST-like samples by convolving the galaxy sizes and shapes with a \( \theta_{\text{eff}} \sim 0.6 \) circular Gaussian PSF – the “effective” PSF FWHM value expected at the median airmass of \( X = 1.2 \) for LSST where \( \theta_{\text{eff}} = 0.8'' \) is the “effective” zenith seeing FWHM for the \( i \) band taken from Table 2 of Ivezić et al. (2019). The reason we use these “effective” values is to account for the deviation of the LSST fiducial atmospheric PSF (a von Kármán type PSF) from the Gaussian shape (Angeli et al. 2016) assumed here.

The PSF convolution is done analytically in moment space. It is straightforward with known profiles, and also enables us to avoid memory issues that arise from the need for pixel-level convolution in large images. We model galaxies with Gaussian surface brightness profiles described by

\[
G(x) = A \exp \left[ -\frac{1}{2} (x - \mu)^T \Sigma^{-1} (x - \mu) \right],
\]

with \( \mu \) being the position of the galaxy center, \( A \) the normalized flux of the galaxy, and \( \Sigma \) the shape covariance matrix. Our analytical method of PSF convolution requires the shape covariance matrix of the galaxies (a.k.a. the second moments’ matrix) in order to work in the Fourier domain. This covariance is a 2 by 2 matrix that encapsulates the ellipticity and the position angle of each galaxy as seen on the 2D sky. Assuming a simple Gaussian model of equation (1) for both the galaxy and the PSF, we add the second moments of the galaxy \( \{Q_{ij}\} \) and the PSF \( \{P_{ij}\} \) to get the combined moments of

\[
Q_{ij} = Q'_{ij} + P_{ij},
\]

where \( i, j \in \{1, 2\} \), and convert them back to ellipticity according

\(8\) The region near the outer boundary of the catalogue footprint is affected by both the light cone and the sky footprint. The lack of galaxies outside the lightcone used to define the simulation sample may bias the modelling of the lensing effects and therefore likely to be discrepant from the rest of the catalogue.
It shows good agreement between our analysis (Fig. 3c) and compared with the pixel-level convolution of involved with the PSF that we adopt are shown in Fig. 3, and is boost that occur when the true galaxy image is analytically con-

It shows good agreement between our analysis (Fig. 3c) and compared with the pixel-level convolution of involved with the PSF that we adopt are shown in Fig. 3, and is boost that occur when the true galaxy image is analytically con-

Moreover, Fig. 4 shows how PSF convolution affects the observed ellipticity distribution of galaxies by reducing their ellipticity magnitudes.

We generate the simulated LSST-like magnitudes in 6 optical bands (ugrizy, covering the wavelength range 320–1050 nm) from Buzzard and estimate their errors utilizing and modifying the Dark Energy Science Collaboration (DESC) Photo-z Working Group Simulations Generator. This implements the simple magnitude error model described in Ivezic et al. (2019). We assign magnitude errors in each of the six bands assuming the equivalent of 1.3 years of LSST observations, consistent with the depth of our sample. This depth translates into S/N > 20 for all the point sources brighter than i = 24. We also take into account the presence of blends by re-measuring observed magnitudes and their errors after the blending emulation (outlined in Section 3.2), i.e. by using the combined magnitude and shape of each blend.

Fig. 5 shows the coverage of our galaxy sample on the sky. Using the spectroscopic and photometric redshift data, we split the sample into four tomographic redshift bins (0.3,0.5), (0.5,0.7), (0.7,0.9), and (0.9,1.2). We exclude low redshift galaxies where the lensing signal is low. We also chose a larger tomographic bin width (Δz_tomo = 0.3 as opposed to 0.2 used in the other bins) for the highest redshift bin to improve the number count for that bin. The maximum redshift of z = 1.2 we adopted here is the same as the analysis setup of LSST DESC Science Requirement Document (Mandelbaum et al. 2018) for the large-scale structure. However, they do not have an upper redshift cutoff for their “source” sample. See Section 4 for information regarding our upper redshift cutoff and photometric redshift estimation.

To model more realistic measured ellipticities rather than adopting the noise-free (only lensed and PSF-convolved) ellipticities, we include an analytic formulation of the effects of pixel noise. We use the analytical model proposed by Melchior & Viola (2012) which captures the entire process from noisy images to shear estimates. One advantage of this model is that it only requires the signal-to-noise ratio of each observed galaxy in order to predict the observed ellipticities from noise-free ellipticities. We use this model to add simulated pixel noise to the PSF-convolved shapes based on the expected i-band flux signal-to-noise (S/N) level derived from the magnitude errors, dm_i:

$$\frac{\Delta m_i}{N} = \frac{1}{10^{\text{dm}_i/2.5}} - 1$$

We show the resulting ellipticity distribution in Fig. 4 (green histogram). In the same figure, the histogram of the noise-free lensed shapes along with their PSF-convolved version are plotted for comparison. Due to the effect of noise on moment-based shapes, a minuscule fraction of the ellipticity magnitudes got pushed to values above unity and will not be used in our analysis. To model measured galaxy sizes rather than adopting the true lensed-only catalogue sizes, we include an analytic formulation of the observed galaxy sizes due to signal-to-noise ratio and the magnitude of the ellipticity. We employ an analytical model (Refregier et al. 2012, Table 1) to add Gaussian noise of standard deviation

$$\sigma_{\epsilon} \approx \frac{\epsilon}{1 + \epsilon^2} \frac{1}{2(S/N)^2} \quad \epsilon = \frac{2|\epsilon|}{1 + \epsilon^2}$$

Figure 2. The redshift distribution of the non-blended sources with (a) spectroscopic tomography (b) photometric tomography using point estimates (c) photometric tomography using full PDF, p(z) for the Buzzard source sample galaxies used in this work for cosmological inference. The redshift distribution of each tomographic bin is normalized such that its integral over redshift is unity. Panel (c) illustrates that the bins defined via photo-z have “tails” due to uncertainties in the redshift estimates. The vertical dashed lines mark the redshift bin edges at z = 0.3, 0.5, 0.7, 0.9, and 1.2 used for tomography. These plots are averaged over two realizations of Buzzard. See Table 2 for the number counts. We use a different colour palette to show tomographic bins for plot (a) to emphasize that galaxies selected for photo-z tomography (b and c) are selected based on photo-z point estimates and are not the same as those from the spec-z tomographic bins (a) that are selected based on spectroscopic redshifts.

Figure 2. The redshift distribution of the non-blended sources with (a) spectroscopic tomography (b) photometric tomography using point estimates (c) photometric tomography using full PDF, p(z) for the Buzzard source sample galaxies used in this work for cosmological inference. The redshift distribution of each tomographic bin is normalized such that its integral over redshift is unity. Panel (c) illustrates that the bins defined via photo-z have “tails” due to uncertainties in the redshift estimates. The vertical dashed lines mark the redshift bin edges at z = 0.3, 0.5, 0.7, 0.9, and 1.2 used for tomography. These plots are averaged over two realizations of Buzzard. See Table 2 for the number counts. We use a different colour palette to show tomographic bins for plot (a) to emphasize that galaxies selected for photo-z tomography (b and c) are selected based on photo-z point estimates and are not the same as those from the spec-z tomographic bins (a) that are selected based on spectroscopic redshifts.

GalSim (Rowe et al. 2015). It shows good agreement between our analysis (Fig. 3c) and the GalSim pixel-based output (Fig. 3b). Convolution with the PSF is assumed to preserve the flux thereby not changing the magnitudes. Moreover, Fig. 4 shows how PSF convolution affects the observed ellipticity distribution of galaxies by reducing their ellipticity magnitudes.

GalSim (Rowe et al. 2015). It shows good agreement between our analysis (Fig. 3c) and the GalSim pixel-based output (Fig. 3b). Convolution with the PSF is assumed to preserve the flux thereby not changing the magnitudes. Moreover, Fig. 4 shows how PSF convolution affects the observed ellipticity distribution of galaxies by reducing their ellipticity magnitudes.

GalSim (Rowe et al. 2015). It shows good agreement between our analysis (Fig. 3c) and the GalSim pixel-based output (Fig. 3b). Convolution with the PSF is assumed to preserve the flux thereby not changing the magnitudes. Moreover, Fig. 4 shows how PSF convolution affects the observed ellipticity distribution of galaxies by reducing their ellipticity magnitudes.
Blending effects on cosmic shear

3 BLENDING

3.1 Importance of blending in weak lensing

Blending of galaxy surface brightness profiles characteristically can be divided into two categories: (1) recognized blending, in which multiple surface brightness peaks can be recognized in a blended image, and (2) unrecognized or ambiguous blending, where the deblending algorithm is unable to deduce more than one surface brightness peak (see Fig. 6). We will focus here more on this latter aspect.

As noted by Mandelbaum (2018), the two prevalent systematic errors in weak lensing are thought to be those of blending on shear measurement and on photometric redshift. In an unrecognized blend, the single object can often have observed shapes and fluxes that are quite different in value from those of any of the two or more constituent objects that comprise the blend. We observe in our dataset that most of the unrecognized blends occur between objects at different redshifts, which agrees with the findings of Kirkby et al. (in preparation). Therefore, for most blends, it is not clear what to make of the collective shear and photometric redshift estimates. A recent study by MacCram et al. (2022) looked at the joint effects of blending in both shear and photometric redshift calibration using image simulations in the context of DES Year 3.

The fact that a significant number of objects will be part of unrecognized blends upon observation in the LSST survey means that simply removing (unrecognizably) blended objects from the data set is not an option, as they cannot be isolated and removed if they are unrecognized, by definition. Dawson et al. (2016) estimates the unrecognized blend fraction to be ~14% for the LSST gold sample. Sánchez et al. (2021) revisit the statistics of the blends.
The PSF FWHM of magnitudes of the two galaxies are 22.60 and 23.35, as viewed from left to the flux of galaxies and a sky level of els are noise-free, while the cause issues with galaxy detection in blended profiles. The two top ness profiles across the blend (els), simulated using GalSim Figure 6.

E. Nourbakhsh et al.

3.2 Emulating blending phenomenon in Buzzard

Our goal is to emulate blending effects in a sample analogous to what will be used in shear measurements in the upcoming deep imaging surveys such as LSST, sometimes referred to as the “gold sample” and often defined by an apparent magnitude cut, as we previously discussed throughout Section 2. However, we cannot simply limit ourselves to galaxies brighter than some limiting magnitude, since the steep slope of the apparent magnitude number counts observed in survey data indicate that galaxies with fluxes up to two magnitudes fainter will contribute enough flux to bias shape and photo-z measurements if blended with an object from our weak lensing sample. Therefore, we must include galaxies fainter than our fixed flux limit when studying blends. In the same vein as Dawson et al. (2016), we include and examine galaxies up to 2 magnitudes fainter (to $i < 26$) that will contribute to unrecognized blends in our cosmology sample ($i < 24$) during emulation. Note, however, that additional cuts are made to define the final source sample. These cuts are described in Section 5.

Assuming a simple Gaussian galaxy model of equation (1), the quadrupole moments of the combined shape of $n$ galaxies that are blended together become (cf. Dawson et al., 2016, equation B3)

$$Q_{ij}^{\text{blend}} = \frac{1}{\sum_{k=1}^{n} A^{(k)}} \sum_{k=1}^{n} A^{(k)} [(\mu_i^{(k)} - c_i)(\mu_j^{(k)} - c_j) + \Sigma_{ij}^{(k)}].$$ (6)

where $k$ loops over the galaxies in the blend, $c_i$ and $c_j$ are the coordinates of the luminosity centroid of the blend, and $\Sigma_{ij}^{(k)}$ is the element $ij$ of the shape covariance matrix of the $k$th galaxy in the blend. The ellipticity of the blend can then be determined by using equation (3).

To emulate the effects of unrecognized blending on the truth catalogue, we first identify potential blend candidates in crowded regions (a.k.a. blended scenes), defined as two or more galaxies closer than the spatial separation $\theta_9 = 4.0''$ measured from the centroids of galaxies already being PSF-convolved prior to image simulation, as detailed in Section 2.2. All simulated images are based on 1.3 years’ worth of LSST observations. The source detection is done by taking advantage of the software package SEP (Barbary 2016), an in-memory python implementation of the Source Extractor or Sextractor software (Bertin & Arnouts 1996). In cases where detection and truth do not correspond, we treat all non-blended galaxies in each detected segment as a single combined galaxy (i.e. an unrecognized blend) and create a so-called “blended” catalogue using their collective values. We sum non-blended lensed-only galaxy fluxes within each unrecognized blend to get the total flux, convert it back to the total magnitude, and add noise to the magnitude, as done in Section 2.2 for galaxies in the non-blended catalogue. We calculate combined second moments, $Q_{ij}^{\text{blend}}$, using equation (6). The elements of this matrix give us the half-light radius of each unrecognized blend. They can also be plugged into equation (3) to

(2013) presents the impact of rejecting blend systems that exhibit major blending when determining expected galaxy source number density and redshift distribution from LSST. The rejection of these recognized blends and their impact on cosmology is the focus of Section 7.3.

In weak lensing measurements, unrecognized blending poses both a noise problem and a bias problem: the random nature of overlapping galaxies can add spurious ellipticity noise, and the corresponding random ellipticities dilute the shear signal causing a shear polarizability bias. It has been shown by Dawson et al. (2016) that as a result, shear variance will increase both through a reduction in counts and a higher contribution to shape noise than non-blended galaxies. However, multiplicative shear bias is of greater significance here.

For efficiency, this process is performed by HEALPix, and therefore does not include the pairs that cross the HEALPix borders. We believe that disregarding these pairs does not change our overall conclusion, especially since the HEALPix pixels are relatively large (~53.7 square degrees).
yield the collective shape for each blended system. We then add noise to the half-light radii and shapes of the unrecognized blends in accordance with the methods described in Section 2.2. If a galaxy in the non-blended catalogue goes undetected by SEP: (1) we keep it in our blended catalogue if the observed (noisy) version of its magnitude is brighter than our sample cut at $i = 24$, or (2) remove it from our blended catalogue if it is fainter than $i = 24$. However, such non-detections tend to be quite faint in general and thus are more likely to fall under the latter category.

The emulation has resulted in approximately 12% of galaxies in the blended catalogue being unrecognized blends. See Figs 7 and 8 for a simplified toy example and real production example of the image simulation, respectively. The detection and deblending parameters that we specified for SExtractor through SEP are listed in Table 1. Later in Section 5, we calculate galaxy angular correlations and check for any systematic changes that occur due to blending by comparing measurements made using our blended catalogue to measurements made on the equivalent Buzzard catalogue where blending effects were not emulated.

Fig. 9 shows the distribution of the angular separation between the confirmed unrecognized blend components for pairs (orange histogram) and for all unrecognized blends, including multiples (blue histogram). For the multiples, we use the average angular distance of the blend components from the luminosity centroid of the blend. There is still a very small tail at 4'' in the orange histogram (the equivalent of 2'' for the blue histogram). Thus, our conservative choice of the separation angle threshold ($\theta_b = 4''$) we use to find the blended scenes is appropriate.

A measure of purity – described in detail in Section 7.3 – can be used to quantify the degree of blending with known surface brightness profiles. Therefore, during the emulation process, we calculate the purity value for every unrecognized blend, recognized blend, or otherwise isolated galaxy in the blended scenes (all brighter than $i = 24$) to be used later in Section 7.3 for excluding severe cases of unrecognized blends. A relevant point to mention is that unrecognized blends themselves can be (recognizably) blended with other galaxies in which case they will be regarded as recognized blends by the observer who has no insight into the truth. This is exemplified by the unrecognized blend in Fig. 7.

| Parameter            | Value         |
|----------------------|---------------|
| DETECT_NTHRESH       | 2.0           |
| DETECT_MINAREA       | 5             |
| DEBLEND_NTHRESH      | 32            |
| DEBLEND_MINCONT      | 0.0001        |
| FILTER               | gauss_5.0_7x7.conv |
| CLEAN_PARAM          | 1.1           |

Table 1. Specified SExtractor detection and deblending parameters through the Python library for Source Extraction and Photometry (SEP).

4 PHOTOMETRIC REDSHIFT ESTIMATION

We use Trees for Photo-Z (TPZ; Carrasco Kind & Brunner 2013), to estimate photometric redshifts for our sample. TPZ is a supervised machine learning algorithm in Python that uses prediction trees and random forest techniques to produce both robust photometric redshift PDFs and ancillary information for a galaxy sample. Bernstein & Huterer (2010) and Newman et al. (2015) showed that \( \sim 10^4 \) spectra are necessary to calibrate photo-z’s to Stage IV requirements of the LSST Science Requirements Document (SRD)\(^{12}\). With these findings in mind, we train our estimator consisting of 100 trees using a training set of \( 10^5 \) galaxies with spec-z’s randomly drawn from our sample. Fig. 10 provides an illustration of the algorithm’s performance for point estimates for the blended and non-blended validation samples. The spectroscopic redshift and photometric redshift point estimates are denoted by \( z_s \) and \( z_p \), respectively. We have restricted our redshift tomography to \( z < 1.2 \) beyond which the reliability of photo-z estimation is expected to decline significantly. There is a slight tail due to the outliers at low \( z_s \) in the blended case because \( z_s \) is taken from the redshift of the brightest galaxy in the i band, which tends to be at lower redshift than other galaxies in the blend.

Figs 2a, 2b and 2c show our tomographic redshift binning with spec-z and photo-z for the non-blended sample. While one can see a good agreement between the two-point estimates (Figs 2a and 2b), the photo-z PDFs (Fig. 2c) have a smooth and long tail in the distribution due to relying on the photometric data to estimate the redshift. To continue our analysis for the spec-z tomography in the next section, we make the same assumption as in Fig. 10, by considering the spectroscopic redshift of an unrecognized blend to be

\[^{11}\text{https://github.com/mgckind/MLZ}\]
\[^{12}\text{Obtainable from https://docushare.lsstcorp.org/docushare/dsweb/Get/LPM-17}\]
Figure 8. An example of the performance of the detection module in BlendSim. All the galaxies from the non-blended catalogue are shown by the red pluses. We use different markers for galaxies that belong to different detections according to the segmentation map. Two unrecognized blends each consisting of two objects are marked by cyan triangles and orchid hollow X shapes, while two isolated galaxies are marked by a crimson triangle and a yellow pentagon. The values in the parentheses are the $i$-band magnitude and purity (Section 7.3), respectively (in red for individual galaxies from the non-blended catalogue, and in green for blends collectively). A galaxy at the bottom which is undetected by SExtractor is marked by a faint red square, and, in this case, is removed from the sample since it is fainter than $i = 24$. The segmentation maps corresponding to each detection are overlaid using the same colour as the markers. The left panel, which is used for detection, has Poisson noise due to the flux of the galaxies and a sky level of $i = 20 \pm 3$ mag arcsec$^2$, while the right panel is noiseless. The green dashed ellipses with green pluses at their centres are the shapes measured by SExtractor, and the red dashed ellipses on the unrecognized blends with cyan crosses at their centres (marking the analytical luminosity centre of the blend) are the shapes computed analytically from the combined moments of individual galaxies in the blend without pixel noise. We consider the noisy version (method described in Section 2.2) of the latter as the unrecognized blends’ shapes in calculating correlations for the blended sample. The colour bars depict the co-added pixel values over 1.3 years of LSST.

Figure 9. The distribution of the average distance from the unrecognized blend’s centroid (blue) and the separation angle of two-member unrecognized blends (orange), obtained from one HEALPix cell of the blended sample. Note that the orange histogram peaks at about the PSF FWHM of our sample marked by the red dashed line at $\sim 0.9''$. The blue histogram uses the luminosity centroid as a reference for blends with two or more members and peaks at $\sim 0.45''$ which is half of the PSF FWHM. The very small tail at $\theta = 4''$ indicates that using this rather conservative value as a separation angle threshold when defining blend candidate groups in Section 3.2 is appropriate.

The plots for the redshift distribution of the blended sample (not shown) do not show any marked differences when compared to the non-blended sample. This depends on what we are assuming about the impact of blending on each of the three measures of redshift tomography. Although a sizable proportion ($\sim 12\%$) of the objects are blended and re-simulated, most blends have multiple faint objects blended with one bright source whose magnitude and colour dominates the blend thereby introducing a slight change in the observed photometry, and thus photo-$z$. To get more insight into the population of the faint neighbours, one useful quantity to look at could be the magnitude difference, $d_{m_{12}}$, between the brightest and the second brightest galaxy inside each blend. We find that about 60% of blends have $d_{m_{12}} > 1.0$ and about 30% of them have $d_{m_{12}} > 2.0$. Hence, most galaxies inside unrecognized blends incorporate blended galaxies that are so faint that they do not affect the photometric redshift significantly.
components of shear in relation to the line connecting each pair of underlying dark matter. The correlation is computed as the sum of the squares of the shape noise and shape error (values are per-component): 

\[ \xi^i_j(\theta) = \int_0^\infty \frac{d\ell}{2\pi} C^i_j(\ell) g_{ij}(\ell). \]

under the flat-sky approximation with \( J_m(x) \) being the \( m \)-th order Bessel function of the first kind and \( \theta \) the angular separation. The superscripts \( i \) and \( j \) label different tomographic redshift bins and \( C^i_j \) represent the power spectra that contain information about the underlying dark matter. The correlation is computed as the sum \( \xi^i(\theta) \) and difference \( \xi^i(\theta) \) of the product of tangential and cross components of shear in relation to the line connecting each pair of galaxies. Here we add that \( \xi^i_j(\theta) = \xi^j_i(\theta) \) for all tomographic bin pairs.

To quantify the cosmic shear information using two-point statistics, we use equation (7) to compute the shear-shear correlation between the shapes of source galaxies after analytically deconvolving from them the \( i \)-band PSF in the moment space by solving equation 2 for \( Q^i_j \) - this time with noisy galaxy moments - while assuming the flux is preserved. Recall that the PSF is circular and Gaussian with FWHM \( \sim 0.9" \), in keeping with what we had assumed for the convolution process. The measurement is done across spectroscopic and photometric redshift bins, both averaged over the entire sample.

In order to account for the uncertainty in the ellipticity measurements, we weight our correlations according to the shape noise per tomographic bin\(^\text{13} \) (\( \sigma_{\text{shape}} \)) and shape error (\( \delta e \)) of our galaxies. The galaxy weight for each galaxy is defined as the inverse of the shape noise: 

\[ w_i = \frac{1}{\sigma^2_{\text{shape}} + (\delta e)^2}. \]

where the subscript \( i \) refers to the \( i \)-th galaxy. We exclude from our

\(^{13}\) The variation in shape noise for the different tomographic bins is not large for our lensing samples. For the non-blended sample with spec-z tomography the mean shape noise is on average 0.21 per component with a standard deviation of 0.003 among four bins.

Figure 10. Photometric redshift (\( z_p \)) point estimate versus the spectroscopic redshift (\( z_s \)) for \( 20 \), 000 randomly selected galaxies in each of the non-blended (left panel) and blended (right panel) samples with \( 22 < t < 24 \) and lensing quality cuts described in Section 3 as the validation sets drawn from Buzzard. We employed a hexagonal binning scheme where the value of the hexagon is determined by the number of points in the hexagon mapped onto the colour bars shown at the top of each panel. The central dashed line represents a one-to-one match between \( z_p \) and \( z_s \) and the two dashed lines around the central line draw the borders for the catastrophic outliers where \(|(z_s - z_p)/(1 + z_s)| > 0.15 \). The plots are annotated by the values for the interquartile range \( \sigma_{\text{shape}} \) in \( e_s = (z_p - z_s)/(1 + z_s) \) and the percentage for the catastrophic outliers. The blended set shows a larger interquartile range and outlier rate. Note that the \( z_s \) values for the blended set are taken from the redshift of the brightest galaxy in the \( i \)-band among the galaxies in the blend.

5 WEAK LENSSING COSMOLOGICAL PROBE

Weak lensing enables the measurement of the mass power spectrum directly through cosmic shear, which refers to the distortion of the shapes of distant galaxies due to the gravitational effects of the intervening foreground large-scale mass structure. As a result, there are weak but statistically meaningful correlations in the way galaxies are aligned across the sky and in redshift due to the total mass distribution that their light rays encounter. In view of the small amplitude of the weak lensing effect, it is possible to identify these correlations between the shapes of galaxies by averaging over a sufficiently large sample. This can be done for different distances (i.e. tomographic redshift bins) that gives us a three-dimensional map showing how dark matter is distributed in the universe, independent of assumptions about the relationship (i.e. galaxy bias) between visible/baryonic matter and dark matter. The angular shear-shear two-point correlation function can be formed as

\[ \xi^i_j(\theta) = \frac{\int_0^\infty \frac{d\ell}{2\pi} C^i_j(\ell) J^2_{i+j}(\ell)}{\sigma^2_{\text{shape}} + (\delta e)^2}. \]

for all tomographic bin pairs.

To quantify the cosmic shear information using two-point statistics, we use equation (7) to compute the shear-shear correlation between the shapes of source galaxies after analytically deconvolving from them the \( i \)-band PSF in the moment space by solving equation 2 for \( Q^i_j \) - this time with noisy galaxy moments - while assuming the flux is preserved. Recall that the PSF is circular and Gaussian with FWHM \( \sim 0.9" \), in keeping with what we had assumed for the convolution process. The measurement is done across spectroscopic and photometric redshift bins, both averaged over the entire sample.

In order to account for the uncertainty in the ellipticity measurements, we weight our correlations according to the shape noise per tomographic bin\(^{\text{13}} \) (\( \sigma_{\text{shape}} \)) and shape error (\( \delta e \)) of our galaxies. The galaxy weight for each galaxy is defined as the inverse of the shape noise: 

\[ w_i = \frac{1}{\sigma^2_{\text{shape}} + (\delta e)^2}. \]

where the subscript \( i \) refers to the \( i \)-th galaxy. We exclude from our

\(^{13}\) The variation in shape noise for the different tomographic bins is not large for our lensing samples. For the non-blended sample with spec-z tomography the mean shape noise is on average 0.21 per component with a standard deviation of 0.003 among four bins.
Figure 11. Shear-shear cross-correlation (a) $\xi_{\parallel}$ and (b) $\xi_{\perp}$ between spec-$z$ bins measured in Buzzard, for the no-blended and blended samples. The diagonal panels show the autocorrelation. Error bars are obtained from 2,000 jackknife resamples. Overlaid are the fiducial theory lines (not fitted to data) with no shear calibration parameter applied where we use the redshift distribution of the non-blended sample. A suppression of correlation due to the effects of blending is noticeable.
Figure 12. Shear-shear cross-correlation (a) $\xi_{\gamma\gamma}$ and (b) $\xi_{\gamma}$, between photo-$z$ bins measured in Buzzard, for the no-blended and blended samples. The diagonal panels show autocorrelation. Error bars are obtained from 2,000 jackknife resamplings. Overlaid are the fiducial theory lines (not fitted to data) with no photo-$z$ bias or shear calibration parameter applied where we use the redshift distribution of the non-blended sample. As in Fig. 11, a suppression of correlation due to the effects of blending is noticeable.
analysis galaxies with the deconvolved ellipticity \( e > 0.8 \), \( \delta e > 0.25 \), i-band flux \( S/N < 10 \) and the trace of the second moments' matrix \( (T = Q_{11} + Q_{22}) \) of the observed shape (not deconvolved) less than 1.5 times that of the PSF. The error in shape measurement \( \delta e \) is estimated as the absolute magnitude of the difference between the observed (noisy) ellipticity and the noise-free (lensed only) ellipticity. Our observed shapes are calculated in terms of the i-band \( S/N \) (Section 2.2), so are the associated errors \( \delta e \).

In addition, we restrict ourselves to galaxies detected in at least four LSST bands, including the i-band with \( i > 22 \) from our cosmology sample already with the \( i < 24 \) magnitude cut. We have thus come to the Buzzard source sample, which we will make separately for different cases (depending on blending and tomography) and use for cosmological inference. Using these quality cuts ensures reliable shape measurements and photometry while maintaining a high shear signal for the lensing sample. Including brighter galaxies \( (i < 22) \) would only add 28% more galaxies to the source sample given the steep apparent magnitude vs number counts relation. We are aiming at the background galaxies and therefore it is more appropriate to select fainter galaxies for that purpose. A recent example of such a cut at the 22nd magnitude is Hasan et al. (2021). The fraction of galaxies in the Buzzard cosmology sample that satisfy all the criteria (i.e., will be kept in the source sample) are 63% for both the non-blended and blended catalogues.

We use TreeCorr \(^{14} \) (Jarvis et al. 2016) to compute the correlation functions in the source sample for all redshift bin combinations (made from four redshift bins shown in Fig. 2) in \( N_{\text{th}} = 20 \) angular bins, logarithmically spaced between 6 and 150 arcmin while setting \( \text{bin}_\text{slop} = 0 \) for a high level of accuracy to brute-force calculation\(^ {15} \). Correlations are averaged over two realizations. Figs 11 and 12 show the calculated shear-shear correlations, \( \xi_\ell(\theta) \) and \( \xi_r(\theta) \), for several spec-z and photo-z tomographic bin combinations in our analysis with and without the effects of blending. Error bars are obtained from 2,000 jackknife resamplings (1,000 per mock realization) described in Section 6. The theory prediction lines (not fitted to data) from our fiducial \( \Lambda \) CDM model are produced using The Core Cosmology Library\(^ {16} \) (CCL, Chisari et al. 2019). One expects non-zero correlations in off-diagonal panels due to foreground large scale structure dark matter lensing galaxies across redshift bins. Overall, the correlations suggest that there is suppression from blending across all panels.

Galaxy-galaxy blending systematically changes several different types of measurements over a large volume – data of dimension much higher than the one dimension of blending. There are \( 2 \times N_{\text{th}} \times (N_{\text{th}} - 1) \) independent two-point shear-shear correlation functions, defined in \( N_{\text{th}} \) redshift bins (the factor of two accounts for the fact that shear is a spin-2 field requiring two kinds of correlations to describe it in real space, i.e. \( \xi_\ell \) and \( \xi_r \)). Given \( N_{\text{th}} = 4 \) and \( N_{\text{th}} = 20 \), the total number of measurements for our cosmic shear data vector, \( \vec{v} \), comes to \( N_v = 400 \). Although there is information in the combined analysis of these many correlations vs redshift, it may still be insufficient to fully suppress the noise arising from unrecognized blending of galaxies that propagates to cosmology. We investigate this in Section 7 by analyzing systematic changes in cosmological inference considering blending effects in the cosmological volume of the Buzzard simulations.

### 6 Covariance

In cosmological parameter estimation using the clustering and weak lensing probes, it is crucial to construct a reliable covariance matrix that includes the effects of both systematic errors and sample variance. Taylor, Joachimi, & Kitching (2013) show that the accuracy of parameter estimation is limited by the accuracy of the inverse data covariance matrix – the precision matrix. In addition to calculating the correlation functions, a hybrid of jackknife+mock method\(^ {17} \) is employed to internally estimate the covariance between the correlations obtained from \( N_r \) mock realizations. This is achieved by replacing every \( i \)th jackknife pseudo data vector (calculated from each jackknife pseudo-sample) from the \( j \)th realization, \( v_i^j \), by the weighted average of \( v_i^j \) and the full data vectors from \( N_r - 1 \) other realizations, that is

\[
\gamma_i^j = \frac{(N_r^j - 1)v_i^j + \sum_{k=1}^{N_r} N_r^j N_r^k \vec{v}_k^j}{(N_r^j - 1) + \sum_{k=1}^{N_r} N_r^k N_r^k}, \quad i \in \{1, ..., N_r^j\}
\]

and then calculating the delete-one jackknife covariance matrix using a reconstructed “design matrix” made by concatenating these weighted pseudo data vectors taken from all realizations as

\[
\mathcal{V} = \left[ \begin{array}{c} v_1^1 \\ \vdots \\ v_{N_r}^1 \end{array} \right] \quad \text{and} \quad \mathcal{V} = \left[ \begin{array}{c} v_1^N_r \\ \vdots \\ v_{N_r}^{N_r} \end{array} \right].
\]

Here \( N_r^i \) is the number of patches used for jackknifing in the \( j \)th realization and \( \vec{v}^j \) is the full data vector of the \( k \)th realization. With the same number of jackknife patches, \( N_r \), used for each realization, equation (9) can be simplified as

\[
\gamma_i^j = \frac{N_r^j(v_i^j + \sum_{k=1}^{N_r} N_r^k \vec{v}_k^j) - v_i}{N_r N_r^j - 1}, \quad i \in \{1, ..., N_r^j\}.
\]

The size of jackknife patches has traditionally been restricted to larger than the scale of the two-point correlation function of concern. Nevertheless, according to a recent study by Favole et al. (2021), when using jackknife as the covariance estimator for cosmological analyses, it is more beneficial to beat down the noise by building many resamplings than to select a jackknife size greater than e.g. the maximum galaxy clustering scale. We chose 1,000 equal-area jackknife patches over 10,254 square degrees in our sample using the kmeans algorithm via TreeCorr that leads to an effective jackknife

\(^{14} \) https://github.com/rmjarvis/TreeCorr

\(^{15} \) According to TreeCorr's documentation, bin_slope specifies the maximum possible error any pair of points can have, given as a fraction of the bin size. For more information see https://rmjarvis.github.io/TreeCorr_/build/html/binning.html.

\(^{16} \) https://github.com/LSSTDESC/CCL.

\(^{17} \) The delete-one jackknife pseudo-samples are supposed to be resulting from the entire set minus one patch at a time. Since multiple mock realizations are involved, the approach we present here does not exactly replicate what jackknife does, but it is closer to the expected answer than simply averaging individual jackknife covariance matrices for each realization (e.g. Escoffier et al. 2016).
scale of 3.2 degrees. Hence, there is an adequate number of jackknife patches while their effective diameter (3.2 degrees) is greater than the maximum angular scale of our correlation measurements (2.5 degrees). In this work, \( N_t = 2 \) and \( N_k = 1,000 \), so the number of jackknife replicates totals \( n' = N_t N_k = 2,000 \) for two realizations. As it turns out, this is desirably five times larger than the length of the data vector, 400, which makes it less likely for the resulting covariance matrix to be unreasonably noisy or singular. The three-dimensional design matrix, \( V \), is then fed to TreeCorr functions to generate the two-dimensional jackknife covariance matrix, whose \( i \)th row and \( j \)th column reads

\[
C_{ij} = \frac{n'}{n'} \sum_{k=1}^{N_k} \sum_{l=1}^{N_l} (V_{ij}^k - \bar{V}_i)(V_{lj}^k - \bar{V}_j),
\]

where \( \bar{V}_i \) is the \( i \)th correlation measurement averaged over all weighted jackknife pseudo data vectors from all realizations. Fig. 5 shows the jackknife regions for one realization of our sample.

The covariance matrices used in cosmic shear studies are typically large since they incorporate many tomographic bin combinations and many angular bins per combination. For the internal covariance estimators such as jackknife, this requires a sample of billions of galaxies to maintain a high S/N. With this in mind, we will also use an analytic approach to construct a Gaussian covariance matrix using the TJPCCov\(^{18}\) code for comparison. It utilizes the techniques outlined in Singh (2021) and Skylen\(^{19}\), and ignores the non-Gaussian features whose effects on final parameter constraints are very small. Since the evaluation of the Legendre polynomials is extremely slow at low angles (high \( \ell \)’s), we use a flat-sky approximation that transforms them into Bessel functions. An example of the covariance matrices used in this study is shown in Fig. 13. The jackknife and the Gaussian covariance matrices are shown respectively in the lower and upper triangles of the same plot for comparison. Here we demonstrate that our internally estimated covariance matrix is comparable, at least pattern-wise, with the analytical covariance matrix.

7 BLENDING AND COSMODOLOGICAL INFERENCE

7.1 Analysis setup

This goal of the study is to perform a cosmic shear analysis using all the correlation functions across the redshift bins to evaluate the impact of blending on cosmological parameter estimation. A joint analysis of all the LSST probes is expected to increase the precision (Zhan & Tyson 2018), though this analysis is postponed to a later study. We find the likelihoods comparing the theoretical and experimental correlation functions using the Markov chain Monte Carlo (MCMC) technique for which we provide a covariance matrix from the jackknife resampling on the data and the analytical covariance matrix explained in the previous section.

This analysis is carried out using CosmoSIS\(^{20}\), a modular pipeline developed by Zuntz et al. (2015) that performs the Bayesian inference and provides constraints on the cosmological parameters. In order to generate theory predictions, the matter power spectra are computed using the CAMB\(^{21}\) (Lewis, Challinor, & Lasenby 2000; Lewis, Challinor, & Lasenby 2000; Howlett et al. 2012) Boltzmann module with the HALOFIT (Smith et al. 2003; Takahashi et al. 2012) non-linear correction. We employ the MultiNest sampler (Feroz, Hobson, & Bridges 2009) with 500 livepoints, efficiency set to 0.1, tolerance set to 0.1 and constant efficiency set to False. The reader is referred to Feroz, Hobson, & Bridges (2009) for a detailed description of these parameters. The contours and confidence intervals in the parameter space are then extracted and plotted using ChainConsumer (Hinton 2016) for both the jackknife covariance matrix and the Gaussian covariance matrix. The contour plots for jackknife covariance matrix are displayed in the main text while the ones for Gaussian covariance matrix can be found in Appendix B.

The priors used for parameter constraints are shown in Table 3. This table consists of non-informative priors for cosmological parameters of interest to ensure that they will not bias our inferences. However, for the systematic parameters associated with measurements of shear (the shear calibration bias, \( m_{1-4} \)) and photometric redshift (the photo-z shift bias, \( \Delta z^{1-4} \)), we use our \textit{a priori} knowledge to impose informative priors. For the shear calibration parameter, we take the ensemble average of the observed (sheared, convolved, noisy, and then deconvolved) ellipticity \( \epsilon_{\text{obs}} \) and shear-only ellipticity \( \epsilon_{\text{true}} \) of the Buzzard sample:

\[
\langle \epsilon_{\text{obs}} \rangle = (1 + m)\langle \epsilon_{\text{true}} \rangle
\]

This results in a value for \( m \) of \( \sim 0.02 \). It is not required to maintain this value across all tomographic bins and for various samples since it depends on multiple factors including galaxy flux, redshift, and size distribution, and is statistically correlated between tomographic bins. We, therefore, assign a conservatively wide Gaussian prior of \( \sigma_{m} = 0.03 \) for the shear calibration parameter in all tomographic bins. In effect, for each pair of tomographic bins, the theoretical predictions of shear-shear correlations are re-scaled as

\[
\xi_{\pm}^{ij}(\theta) \rightarrow (1 + m_i)(1 + m_j)\xi_{\pm}^{ij}(\theta).
\]
where \( i \) and \( j \) denote different tomographic bins. For the photo-z shift bias, we cross-correlate our observed sample with the validation sample of ~200,000 galaxies with corresponding spec-z values to determine the bias in the mean of the redshift distribution, \( \Delta z \), along with its standard deviation derived as

\[
S_{\Delta z} = \sqrt{S_{\Delta}^2 + S_{\rho}^2 - \text{Cov}(S_{\Delta}, S_{\rho})},
\]

where \( S_{\Delta} = \sigma_{\Delta}/n \) and \( S_{\rho} = \sigma_{\rho}/n \) with \( \sigma_{\Delta} \) and \( \sigma_{\rho} \) being the weighted standard deviations of the spectroscopic and photometric samples, respectively. Here \( n \) is the number of galaxies cross-correlated and \( \text{Cov}(S_{\Delta}, S_{\rho}) = 2rS_{\Delta}S_{\rho} \) with \( r \) being the Pearson correlation coefficient of spec-z and photo-z point estimates. This is necessary because the samples are correlated. This calculation is conducted for each tomographic bin resulting in \( \Delta z^i \) values where \( i \in \{ 1, 2, 3, 4 \} \) as seen in Table 3. The stacked photo-z PDFs of galaxies (\( \rho(z) \) in Fig. 2c) are fed to CosmoSIS, in place of the redshift number density, \( n(z) \), for the cases with photo-z tomography. The uncertainties in the redshift distribution mean values transform the redshift distribution of the \( i \)th tomographic photo-z bin as

\[
n'(z) \rightarrow n'(z - \Delta z^i).
\]

It should be noted that all priors are adjusted based on the non-blended sample with spec-z tomography (we will not change them at any point in our analysis) to serve as a reference for how the inference changes due to blending and with the use of photometric redshifts for tomography.

We generate runs with and without blending using all the data points available in our two-point correlation functions. Since baryonic effects are not present in Buzzard, no additional angular scale cuts are imposed on our data vectors to remove these effects. Moreover, the lower bound of our angular scale (6 arcmin) is large enough to minimize the impact of finite simulation resolution on the cosmic shear signal discussed in DeRose et al. (2019). Our results may be optimistic in terms of constraining power, as baryonic effects will always be present in the real universe, but this is less relevant to the primary purpose of this study, which is to compare cosmological findings of blended vs. non-blended universe.

In the course of our analysis, we look only at constraints on parameters most crucial to the cosmic shear signal: the mean matter density in the universe, \( \Omega_m \), the amplitude of matter density fluctuations on scales of 8 h^{-1}Mpc, \( \sigma_8 \), and a combination of these two parameters defined as \( S_{\Delta} = \sigma_8(\Omega_m/0.3)^{0.5} \) to break the degeneracy between \( \Omega_m \) and \( \sigma_8 \).

### 7.2 Unrecognized blending bias in cosmological inference

The estimated cosmological parameters for both the non-blended and blended samples are shown in Figs 14 and 15 using spec-z and photo-z tomography, respectively, where we employ jackknife covariance matrices, and probe cosmology only with shear-shear correlations. Clear bias is visible in the \( S_{\Delta} - \Omega_m \) plane. While blending affects the covariance of cosmological parameters, in both the spec-z tomography and photo-z tomography the difference between the blended case and non-blended case is within 1\( \sigma \) error for \( \sigma_8 \) and \( \Omega_m \) individually. However, it extends beyond a 1\( \sigma \) error for the derived parameter \( S_{\Delta} \). As far as statistics are concerned, this is still a relatively small change, and gives some hope that one could model and correct for the effects of blending for unrecognized blends in a well defined sample.

Our attention is drawn to the fact that the number of galaxies lost due to blending is partially offset by the unrecognized blended objects that collectively exceed the magnitude limit, i.e. by becoming brighter than our upper magnitude cut at \( i = 24 \) for the source sample. Table 2 tabulates the final number counts after applying all the quality cuts mentioned in Section 5. This shows an overall decrease in the number of galaxies due to blending regardless of the method used to assign galaxies to tomographic redshift bins. While a decrease in \( \sigma_8 \) is seen due to blending in Figs 14 and 15, \( \Omega_m \) is not as affected by it, and neither is by the loss of galaxies in the sample. By emulating blends in the catalogue we are smoothing out shapes, i.e. shears in the statistical sense, over the ensemble. In other words, blending introduces altered (collective) shapes for the unrecognized blends in ways that are unrelated to lensing. This tends to dilute the lensing signal because the random nature of the blends does not have a preferred direction. Consequently, this lowers the amplitude of the correlations between the shapes meaning the amplitude of matter fluctuations, \( \sigma_8 \), is lower when blending is present, and so is the normalized version of it, \( S_{\Delta} \). The analysis above is repeated with analytical Gaussian covariance matrices and shows similar biases. For more details see Figs B1 and B2.

### 7.3 Effect of omitting recognized blends

Since the separation angle threshold (\( \theta_{\text{sep}} = 4^{\prime} \)) is larger than the sum of the half-light radii of an average galaxy pair in the Buzzard cosmology sample (~ 1.6\( ^{\prime} \)), the surface brightness shows a detectable dip between the centroid positions of some overlapping galaxies in the blended scenes. Therefore, blended galaxies in these instances...
are recognized blends where detection algorithms are able to separate objects that have been blended together and attempt to deblend them (Melchior et al. 2018). In very deep imaging such as planned for LSST, the low surface brightness tails of galaxies may extend to the mean galaxy-galaxy separation on the sky out to several arc seconds. Bosch et al. (2018) find that nearly 60% of galaxies in HSC Wide survey images are extracted from blended groups in above-threshold regions of the sky. In general, deblending closely-spaced galaxies can be trusted more for pure photometry than for shape measurement. It can cause systematic biases in determining galaxy shapes along the direction connecting the (recognizably) blended galaxies. Therefore, at least in the case of shear-based precision cosmology, it may be prudent to eliminate severe cases of recognized blends from the source sample.

Table 2. Table of the average non-blended and blended number of galaxies used for the cosmic shear analysis after sample cuts per realization of Buzzard simulation in tomographic bins. We use two total realizations in this study. Tomographic binning is based on both the spectroscopic and photometric redshift. These galaxies are used to compute the correlation functions for this study.

| Tomographic bin number | Redshift range | # of non-blended spec-z source galaxies | # of non-blended photometric photo-z source galaxies | # of blended spec-z source galaxies | # of blended photometric photo-z source galaxies |
|------------------------|---------------|-----------------------------------------|-----------------------------------------------|-----------------------------------|-----------------------------------------------|
| 1                      | 0.3 – 0.5     | 73,855,956.5                            | 85,873,670.5                                  | 69,800,775                       | 80,357,357.5                                  |
| 2                      | 0.5 – 0.7     | 76,273,201                               | 70,510,758                                   | 73,946,380                       | 69,636,192                                    |
| 3                      | 0.7 – 0.9     | 55,982,626                               | 60,930,782.5                                | 55,945,993.5                     | 60,773,952                                    |
| 4                      | 0.9 – 1.2     | 30,693,222.5                             | 29,936,703                                  | 32,550,779.5                     | 32,294,214.5                                  |
| Total                  | 0.3 – 1.2     | 236,535,006                             | 247,251,914                                 | 232,663,485.5                    | 243,433,117                                   |

We define the purity, $\rho$, of a galaxy as a measure of the severity of its blending with overlapping galaxies as (Sánchez et al. 2021),

$$\rho_i \equiv \frac{\sum_{p} \rho_{ip} \cdot s_{ip}}{\sum_{j} s_{jp}} ,$$

where $\rho$ sums over all pixels that belong to the galaxy of interest $i$ above a limiting surface brightness where the signal is at least 5% of the sky noise, and $j$ sums over all galaxies that overlap $i$, including $i$ itself. In other words, purity is a ratio of weighted pixel sums and includes a range $0 < \rho_i \leq 1$ with $\rho_i = 0$ for completely blended and $\rho_i = 1$ for completely isolated galaxies. This quantity tells us how much of a galaxy’s surface brightness profile overlaps with other sources’ surface brightness profiles weighted by flux. In this respect, it is not symmetric between overlapping pairs.

As a proxy for identifying catastrophic recognized blends, we assume a minimum purity threshold for galaxies that we keep in the source sample, i.e. we discard the less pure ones. In practice, the purity threshold is set by the limiting surface brightness of the survey and the method used to detect and deblend the sources. If we flag galaxies in crowded regions with $\rho < 0.95$ as catastrophic recognized blends, we retain $\approx 95\%$ of the total sample. Note that galaxies that are not in the crowded regions (for which we made image cutouts) do not have any neighbours within 4 arcsec. Therefore, it is safe to assume that they all have the purity of 1, thereby being isolated according to the adopted scheme.

We shall assume that the majority of recognized blends are properly deblended and do not need to be thrown out of the sample. For typical galaxies in our Buzzard source sample, the pre-convolution half-light radii range between 0.5-1 arcsec, and convolved with our ~0.9′ effective PSF FWHM, a pair of such galaxies is likely to show large systematic distortions in their second intensity moments for pair distances out to a few arc seconds. However, to reiterate what we pointed out in the previous paragraph, at larger separation than our adopted threshold for the blended scenes ($\theta_b = 4''$), a typical pair of galaxies may be deblended accurately for purposes of shape measurement at the sub per cent shear systematic level.

It is therefore useful to ask what the effect on the bias in cosmological inference would be of omitting the catastrophic recognized blends with low purities from the blended sample. The LSST at full depth is expected to have over 30% recognized blends with still a significant dip in surface brightness between the galaxies, as discussed in Melchior et al. (2018). We repeat the cosmological analysis that we undertook above for the unrecognized blends, but omit galaxies in the blended scenes by setting a purity cutoff to the mean galaxy-galaxy separation on the sky out to several arc seconds. Bosch et al. (2018) find that nearly 60% of galaxies in HSC Wide survey images are extracted from blended groups in above-threshold regions of the sky. In general, deblending closely-spaced galaxies can be trusted more for pure photometry than for shape measurement. It can cause systematic biases in determining galaxy shapes along the direction connecting the (recognizably) blended galaxies. Therefore, at least in the case of shear-based precision cosmology, it may be prudent to eliminate severe cases of recognized blends from the source sample.

---

$^{22}$ Purity relates to “blendedness” as defined in the LSST Data Management (DM) software stack (Jurić et al. 2017), by the relation $\rho = 1$—blendedness.
that flags a fraction of the total galaxies in the blended scenes as catastrophic recognized blends. The resulting systematic shift in the error ellipsoid for cosmological parameters is shown in Fig. 16 for spec-z tomography and in Fig. 17 for photo-z tomography. Comparing this with the shift for unrecognized blends in Figs 14 and 15, we see that the shifts in $S_8$-$\Omega_m$ plane are orthogonal. Eliminating recognized blends in two cases where we keep only galaxies with $\rho > 0.96$ and $\rho > 0.98$ (which removes $\sim 9\%$ and $\sim 9\%$ of the galaxies from the full sample, respectively) produces no significant change in $S_8$ for spec-z tomography. However, the extreme case of $\rho > 0.98$ results in a $\sim 0.005$ increase in $S_8$ for photo-z tomography. Interestingly enough, these contour shifts are statistically insignificant when compared with the precision of our analysis. The development of the shifts depend in a non-linear way on the purity threshold for recognized blends. Section 8 discusses the advocacy of these results.

The above analysis is repeated using the Gaussian covariance matrix and shows biases to even a lesser extent. As you can see in Figs B3 and B4, based on the Gaussian covariance matrix as opposed to the jackknife covariance matrix, the results show statistically indistinguishable biases attributed to the omission of low-purity recognized blends.

### Summary of the constraints on $S_8$

Fig. 18 compares the estimated values of the derived parameter $S_8 = \sigma_8 (\Omega_m / 0.3)^{0.5}$ and corresponding uncertainties inferred from the jackknife (left panel) and Gaussian (right panel) covariance matrices. Utilizing two independent types of the covariance matrix, we

\[ \Omega_m = \mathcal{U}(0.1, 0.6) \]

\[ \sigma_8 = \mathcal{U}(0.2, 1.1) \]

\[ h = \mathcal{U}(0.5, 1) \]

\[ \Omega_b = \mathcal{U}(0.01, 0.09) \]

\[ n_s = \mathcal{U}(0.8, 1.07) \]

\[ m_1 = \mathcal{G}(0.02, 0.03) \]

\[ \Delta\chi^2 \sim \mathcal{G}(-0.0064, 0.0003) \]

\[ \Delta\chi^2 \sim \mathcal{G}(-0.0050, 0.0004) \]

\[ \Delta\chi^2 \sim \mathcal{G}(-0.0060, 0.0004) \]

\[ \Delta\chi^2 \sim \mathcal{G}(-0.0037, 0.0005) \]

### Table 3

The priors used in the analysis for cosmological parameter constraints. We used $\mathcal{U}$ to represent non-informative flat priors in the given range and $\mathcal{G}$ for informative Gaussian priors with the first argument as the mean and the second argument as the dispersion. For the purpose of this study, we fix $w = -1.0$, $\Omega_b = 0.0$, $\tau = 0.08$ and consider three massless neutrino species and $N_{\text{eff}} = 3.046$ from the input cosmology (DeRose et al. 2019). The $m_1$ parameters correspond to the shear calibration and $\Delta\chi^2$ to the source photo-z bias, with $i = 1...4$ for four tomographic redshift bins. The systematic parameters are set based on the non-blended sample so that they can serve as our reference to measure potential biases in the blended sample. Future analyses will likely need to expand these prior ranges further in order to account for the effects of blending.

---

**Figure 16.** Cosmological parameter estimation using cosmic shear with spec-z tomography, before and after the exclusion of recognized blends based on their purity values ($\rho$ in equation 17) in the blended scenes. The covariance matrix is produced by the jackknife estimator separately for each sample. The contours of the blended, blended with $\rho > 0.96$, and blended with $\rho > 0.98$ cut samples are displayed in filled blue, unfilled red (solid line) and unfilled black (dashed line), respectively. In each case, the contours drawn correspond to $1\sigma$ (68%) and $2\sigma$ (95%) confidence levels. The green star marks the fiducial parameters.

**Figure 17.** Cosmological parameter estimation using cosmic shear only, before and after the exclusion of recognized blends based on their purity values ($\rho$ in equation 17) in the blended scenes with photo-z tomography. The covariance matrix is produced by the jackknife estimator separately for each sample. The contours of the blended, blended with $\rho > 0.96$ cut and blended with $\rho > 0.98$ cut samples are displayed in filled blue, unfilled red (solid line) and unfilled black (dashed line), respectively. In each case, the contours drawn correspond to $1\sigma$ (68%) and $2\sigma$ (95%) confidence levels. The green star marks the fiducial parameters.
Blending effects on cosmic shear

8 DISCUSSION

The driver for the Dawson et al. (2016) blend fraction estimate was that LSST will have $\sim 0.7''$ FWHM PSF, which at the limit of the main 18,000 square degree survey to its full depth, will result in $\sim 14\%$ of galaxies in the LSST gold sample ($i < 25.3$ where $S/N = 20$ for point sources) blended with gold-sample galaxies as well as galaxies up to $\sim 2$ magnitudes fainter. Naturally, the blend fraction depends on the magnitude limit. For example, the fraction of gold-sample galaxies blended with gold-sample galaxies is far smaller. The science impact can only be assessed via a simulation, which is the subject of this study. The ground-based PSF effect on galaxy apparent shape, noise, and systematic errors due to resulting blends and its science implication is three-fold:

1. Distant blended pairs of galaxies have apparent shape that is dominated by the blend (and not the true shape), resulting in noisier shear statistics in a large random sample;
2. Blended galaxies can have incorrectly estimated photo-$z$, generating false correlations across tomographic redshift bins, impacting all cosmology probes including cosmic shear;
3. The ground-based PSF circularizes all galaxy images, yielding systematically biased measured shear for a given shear. This is the shear calibration term (Massey et al. 2013; Jarvis et al. 2016).

The latter is not the focus of this paper, although it is included and calibrated for in our study and is significant for cosmic shear and galaxy-mass probes. Surveys such as DES (DeRose et al. 2021), DLS (Jee et al. 2016), HSC (Mandelbaum et al. 2018), and KiDS (Heymans et al. 2021) have dealt with this via simulations or METACALIBRATION (Huff & Mandelbaum 2017; Sheldon & Huff 2017).

It is important that shear errors remain random averaged over all the visits to each sky patch. If not, the systematic error arising from such a situation will limit the cosmological precision achieved. For LSST we need to reduce the residual shear-shear power systematic error to well below the cosmological signal on the angular scales of ten arcminutes to several degrees, which sets a goal of $< 3 \times 10^{-7}$ for the full gold sample, that will be used for cosmological parameter estimation (Mandelbaum et al. 2018). The main problem arises when one has a mix of effects (1), (2), and (3). We need some understanding of the effects (1)+(2) separately from effect (3), since (3) will likely be calibrated to the needed accuracy for real observations. Effects (1) and (2) also create a shear multiplicative bias (corrected through the shear calibration parameter) on their own. That is a natural outcome of blending itself, and we need to understand it separately from the PSF rounding. Thus, our simulation is focused on the galaxy-galaxy blending effect itself.

A large survey like LSST will rely on photometric information to estimate redshifts and produce a three-dimensional map of galaxies. To isolate the effects of blending in a catalogue with distances determined by photometric redshifts, we also analyze the enhanced Buzzard catalogue data replacing spectroscopic redshifts with TPZ photometric redshifts.

We estimate that about 12% of the galaxies in our blended catalogue are unrecognized blends that have more than one counterpart in the truth catalogue. We then investigate how systematic errors due to blending (with spec-$z$ or photo-$z$ tomography) propagate to cosmology constraints. In each case (Figs 14 and 15 using the jackknife covariances or Figs B1 and B2 using the Gaussian covariances), our simulations show a small reproducible effect on our posteriors due to unrecognized blending, and in most cases are consistent with some known mechanism. This suggests that there is some hope of modelling the effects of blending in LSST in order to correct the cosmological parameters to first order, e.g. through an additional shear calibration parameter, $m_b$, and perhaps widening of the prior on the shear calibration parameter. This will be aided by cutting the sample in multiple ways in order to validate the models: a sample with the best seeing; samples with various quality recognized blends removed; a sample absent galaxies having strong colour gradients; and samples with different limiting apparent magnitudes. Finally,
...the deep drilling fields will enable most of these investigations, cutting ultra-deep data in all these ways early in the survey.

We notice that tightening the priors on the shear calibration parameter (i.e., adopting a smaller $\sigma_m$) results in stronger discrepancies between the non-blended and blended cosmology constraints. On the contrary, relaxing these priors by adopting a larger $\sigma_m$ decreases such discrepancies at the expense of lessening the constraining power.

In any case, as summarized in Fig. 18, including the effects of blending in our simulation induces a bias in the measured $S_8$ parameter of $\sim 0.02$. The amplitude of this shift is twice the width of the expected 68% confidence interval, and thus a source of uncertainty that needs to be mitigated in future cosmological analyses that utilize deep imaging data. Blending apparently has only a modest effect on photometric redshifts. We believe that this is due to the fact that most of the blends are with at least three times fainter objects that only moderately pollute the flux of the brighter object, and our photo-$z$ algorithm seems to be mostly robust to these flux variations.

While methods like Metacalibration (Huff & Mandelbaum 2017; Sheldon & Huff 2017) and “Bayesian Fourier Domain” (BFD; Bernstein & Armstrong 2014) are developed to deal with problems in the shape measurement of recognized blends, they do not handle unrecognized blends since, by definition, these blends are not recognized to be two or more galaxies and will not have their shapes separately measured in the first place. Dealing with the detection bias (e.g., due to unrecognized blends) was the motivation behind developing Metadetection (Sheldon et al. 2020), a technique that includes object detection in the Metacalibration process. It can mitigate the same-redshift unrecognized blending biases in shear measurement as a result of overlapping objects at the same redshift having the same shear. The cross-redshift blending – which is the most probable scenario – still remains impossible to handle when the blend is unrecognized.

The omission of severe cases of recognized blends (Section 7.3) is done by applying two lower cuts on the purity of galaxies. As pointed out in Section 3.2, all galaxies detected in the blended scenes – including unrecognized blends – are assigned purity values and undergo purity cuts. The cuts are applied to the blended catalogue and introduce slight shifts in the estimated cosmological parameters, more noticeably in the results of photo-$z$ tomography. Overall, the cuts trend toward the fiducial values as we use more stringent purity cuts, except for the estimates of $S_8$ using spec-$z$ tomography with jackknife covariance. By removing the low-purity galaxies with only one counterpart in the non-blended catalogue, we should not observe any visible changes toward the fiducial values in the estimation of our parameters. This is due to the fact that recognized blending bias is not considered when modelling the shape measurement in this study (i.e., purity is not taken into account when adding noise to the shapes), and therefore removing galaxies mentioned above does not remove any blending bias in the shape measurement. Nevertheless, eliminating low-purity galaxies more than likely leads to the elimination of some ($\sim 8\%$ in the extreme case of eliminating galaxies with $p < 0.98$) of the unrecognized blends in the blended scenes, thereby reducing the unrecognized blending bias in our cosmology inference. The removal of a fraction of unrecognized blends decreases the rate of outliers with catastrophically wrong photo-$z$ assignments. This may explain why the aforementioned reduction in the unrecognized blending bias (i.e., the tendency to approach the fiducial or non-blended values) is more noticeable for the inference with photo-$z$ tomography.

The future lies in increasing both the depth and area of sky surveys. The number of galaxies that we used for this simulation ($\sim 0.5$ billion as sources) is still smaller than the number expected in the LSST gold sample ($\sim 3$ billion), and we expect an increase in statistical accuracy by using a larger sample. This also enables the model validation tests via multiple cuts on the catalogue. Even so, there are sufficient galaxies present in our sample such that the systematic effects we seek are more significant than the statistical noise due to the limited sample size. We also have no artefacts such as bright stars, varying sky background, dust, or changing and non-circular PSF. Finally, we include no camera effects that could print through to the co-add and the catalogue. We are focused on blending – our goal is to isolate the effects of galaxy blends.

We analyzed a sample equivalent of 1.3 years of LSST operation. The ultimate 10-year depth of LSST exacerbates the blending impact, for which we will need additional techniques to model it out. This study yields an estimate of the effect of galaxy blending and prospects for improvement. Due to the shared 3D large-scale dark matter structure, there is hope that a joint analysis of multiple independent probes will suppress the 2D blending systematic errors in cosmology. This so-called $3\times2pt$ analysis – which includes galaxy clustering and galaxy-galaxy lensing (a.k.a. galaxy-mass correlation) in addition to the cosmic shear studied in this paper – will be carried out in a follow-on study.

ACKNOWLEDGMENTS

This paper has undergone internal review in the LSST Dark Energy Science Collaboration. We thank the internal reviewers, Bob Armstrong, Patricia Burchat, and Javier Sánchez for their helpful comments.

We acknowledge support from NSF/AURA/LSST grant N509810CC and DOE grant DE-SC0009999. This work used the Extreme Science and Engineering Discovery Environment (XSEDE; Towns et al. 2014) Comet and Expanse clusters at the San Diego Supercomputer Center (SDSC) through XSEDE Research Allocation (under award number AST200036) which is supported by National Science Foundation grant number ACI-1548562. This work also benefited from the HPC at UC project program DPD340 through the SDSC Comet cluster. This research used resources of the National Energy Research Scientific Computing Center (NERSC), a U.S. Department of Energy Office of Science User Facility located at Lawrence Berkeley National Laboratory, operated under Contract No. DE-AC02-05CH11231. We gratefully acknowledge support from the CNRS/IN2P3 Computing Center (Lyon - France) for providing computing and data-processing resources needed for this work.

We thank Alexie Leauthaud for providing the weak lensing catalogue for COSMOS which was used for testing the BlendSim pipeline. COSMOS data is based on data products from observations made with ESO Telescopes at the La Silla Paranal Observatory under ESO programme ID 179.A-2005 and on data products produced by TERAPIX and the Cambridge Astronomy Survey Unit on behalf of the UltraVISTA consortium. EN would like to thank Mike Jarvis for his advice on weak lensing analysis and using TreeCorr and GalSim, Joe DeRose for his advice on using the Buzzard simulation, Patricia Burchat and Sowmya Kamath on blending emulation, Sukdeep Singh on the Gaussian covariance matrices, Yao-Yuan Mao on using the GCRotaLogs package, and Javier Sánchez, James Jee, Imran Hasan, Joe Zuntz, Peter Melchior, David Kirkby, AndrewBradshaw, Mohsen Khorasani, AlanHeavens, David Alonso, AliMahzarnia, JoshMeyers, ErinSheldon, MatiasCarrasco Kind,
Will Dawson, Michael Schneider, Andrew Hearin, and Stephanie Escoffier for helpful discussions. EN also thanks Mahidhar Tatineni for his help with the Comet and Expanse clusters at SDSC.

The DESC acknowledges ongoing support from the Institut National de Physique Nucléaire et de Physique des Particules in France; the Science & Technology Facilities Council in the United Kingdom; and the Department of Energy, the National Science Foundation, and the LSST Corporation in the United States. DESC uses resources of the IN2P3 Computing Center (CC-IN2P3–Lyon/Villeurbanne - France) funded by the Centre National de la Recherche Scientifique; the National Energy Research Scientific Computing Center, a DOE Office of Science User Facility supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231; STFC DiRAC HPC Facilities, funded by UK BIS National E-infrastructure capital grants; and the UK particle physics grid, supported by the GridPP Collaboration. This work was performed in part under DOE Contract DE-AC02-76SF00515.

We made use of the following software packages in addition to those already cited in the text: NumPy and SciPy (Oliphant 2007), Astropy (Astropy Collaboration et al. 2013, 2018), APLpy (Robitaille & Bressert 2012), Matplotlib (Hunter 2007), Pandas (Reback et al. 2021), Healpy (Zonca et al. 2019), MP14Py (Dal-cin, Paz, & Storl 2005), and NetworkX (Hagberg, Schult, & Swart 2008).

DATA AVAILABILITY

The pipeline written for this project is available at https://github.com/enourbakhsh/BlendSim. The processed data underlying this paper will be available upon request. This article uses data provided by the Buzzard Flock simulation group (DeRose et al. 2019), which will also be available upon request, and with the permission of the simulation group.

References

Aihara H., AIsayay Y., Ando M., Armstrong R., Bosch J., Egami E., Furusawa H., et al., 2019, PASI, 71, 114.
Angeli G. Z., Xin B., Claver C., Cho M., Dribusch C., Neill D., Peterson J., et al., 2016, SPIE, 9911, 991118.
Astropy Collaboration, Robitaille T. P., Tollerud E. J., Greenfield P., Droettboom M., Bray E., Aldcroft T., et al., 2013, A&A, 558, A33.
Astropy Collaboration, Price-Whelan A. M., Sipőcz B. M., Gunther H. M., Robitaille T., Bressert E., 2012, ascl.soft.1208.017.
Bertin E., Arnouts S., 1996, A&AS, 117, 393.
Bischof K., (2016), SEP: Source Extractor as a library, Journal of Open Source Software, 1(6), 58.
Becker M., 2013, MNRAS, 435, 115.
Bernstein G. M., Armstrong R., 2014, MNRAS, 438, 1860.
Bernstein G., Huterer D., 2010, MNRAS, 401, 1399.
Bertin E., Arnouts S., 1996, A&AS, 117, 393.
Bosch J., Armstrong R., Bickerton S., Furusawa H., Ikeda H., Koike M., Lupton R., et al., 2018, PASI, 70, S5.
Carrasco Kind M., Brunner R. J., 2013, MNRAS, 432, 1483.
Chang C., Jarvis M., Jian B., Kain S. M., Kirkby D., Connolly A., Krugoff S., et al., 2013, MNRAS, 434, 2121.
Chisari N. E., Alonso D., Krause E., Leonard C. D., Bull P., Neveu J., Villarreal A., et al., 2019, ApJS, 242, 2.
Dal Cin L., Paz R., Storl M., 2005, JPPDC, 65(9):1108-1115.
Dawson W. A., Schneider M. D., Tyson J. A., see M. J., 2016, ApJ, 816, 11.
DeRose J., Wechsler R. H., Becker M. R., Busha M. T., Rykoff E. S., MacCrann N., Erickson B., et al., 2019, preprint (arXiv:1901.02401).
DeRose J., Wechsler R. H., Becker M. R., Rykoff E. S., Pandey S., MacCrann N., Amon A., et al., 2021, preprint (arXiv:2105.13547).
Escoffier S., Coupinou M.-C., Tilquin A., Pisani A., Aguiriche A., de la Torre S., Eales A., et al., 2016, preprint (arXiv:1606.00233).
Favole G., Granett B. R., Silva Lafiura J., Sapone D., 2021, MNRAS, 505, 5833.
Feroz F., Hobson M. P., Bridges M., 2009, MNRAS, 398, 1601.
Górski K. M., Hivon E., Banday A. J., Wandelt B. D., Hansen F. K., Reinecke M., Bartelmann M., 2005, ApJ, 622, 759.
Hagberg A. A., Schult D. A., Swart P. J., 2008, “Exploring network structure, dynamics, and function using NetworkX”, in Proceedings of the 7th Python in Science Conference (SciPy2008), Gaiser Varoquaux, Travis Vaught, and Jarrod Millman (Eds), (Pasadena, CA USA), pp. 11–15, August 2008.
Hartlap J., Simon P., Schneider P., 2007, A&A, 464, 399.
Hasan L., Schmidt S. J., Schneider M. D., Tyson J. A., 2021, preprint (arXiv:2104.03407).
Heymans C., Tröster T., Asgari M., Blake C., Hildebrandt H., Joachimi B., Kuijken K., et al., 2021, A&A, 646, A140.
Hinton S. R., 2016, JOSS, 1, 00045.
Howlett C., Lewis A., Hall A., Challinor A., 2012, JCAP, 2012, 027.
Huff E., Mandelbaum R., 2017, preprint (arXiv:1702.02600).
Hunter J. D., 2007, CSE, 9, 90.
Ivezic Z., Kahn S. M., Tyson J. A., Abel B., Acosta E., Allsman R., Alonso D., et al., 2019, ApJ, 873, 111.
Jarvis M., Sheldon E., Zuntz J., Kacprzak T., Bridle S. L., Amara A., Armstrong R., et al., 2016, MNRAS, 460, 2245.
Jee M. J., Tyson J. A., Hilbert S., Schneider M. D., Schmidt M., Wittman D., 2016, ApJ, 824, 77.
Ji I., Hasan I., Schmidt S. J., Tyson J. A., 2018, PASP, 130, 084504.
Juric M., Kantor J., Lim K.-T., Lupton R. H., Dubois-Feldmann G., Jenness T., Axlrod T. S., et al., 2017, ASPC, 512, 279.
Lewis A., Challinor A., Lasenby A., 2000, ApJ, 538, 473.
MacCrann N., Becker M. R., McCullough J., Amon A., Gruen D., Jarvis M., Choi A., et al., 2022, MNRAS, 509, 3371.
Mandelbaum R., 2018, ARA&A, 56, 393.
Mandelbaum R., Lunasse F., Leauthaud A., Armstrong R., Simet M., Miyatake H., Meyers J. E., et al., 2018, MNRAS, 481, 3170.
Mandelbaum R., 2015, MNRAS, 453, 2757.
Melchior P., Viola M., 2012, MNRAS, 424, 2757.
Melchior P., Moolekamp F., Jerdee M., Armstrong R., Sun A.-L., Bosch J., Lupton R., 2018, A&C, 24, 129.
Melchior P., Joseph R., Sánchez J., MacCrann N., Gruen D., 2021, NatRP, 3, 712.
Miyazaki S., Komiyama Y., Kawanomoto S., Doi Y., Furusawa H., Hamana T., Hayashi Y., et al., 2018, PASI, 70, S1.
Newman J. A., Abate A., Abdalla F. B., Allam S., Allen S. W., Ansari R., Bailey S., et al., 2015, ApJS, 63, 81.
Oliphant T. E., 2007, CSE, 9, 90.
Reback J., Birkmelminder, MKinney W., Van Den Bosche J., Augspurger T., Cloud P., Hawkins S., et al., 2021, Zenodo, doi:10.5281/zenodo.5774815.
Regrefier A., Kacprzak T., Amara A., Bridle S., Rowe B., 2012, MNRAS, 425, 1951.
Robitaille T., Bressert E., 2012, ascl.soft.1208.017.
Rowe B. T., Jarvis M., Mandelbaum R., Bernstein G. M., Bosch J., Simet M., Meyers J. E., et al., 2015, A&C, 10, 121.
Sánchez J., Mendoza L., Kirkby D. P., Burchat P. R., LSST Dark Energy Science Collaboration, 2021, JCAP, 2021, 043.
Sellentin E., Heavens A. F., 2016, MNRAS, 456, L132.
Sheldon E. S., Becker M. R., MacCrann N., Jarvis M., 2020, ApJ, 902, 138.
Sheldon E. S., Huff E. M., 2017, ApJ, 841, 24.
Singh S., 2021, MNRAS, 508, 1632.
APPENDIX A: THE EFFECT OF NON-GAUSSIAN LIKELIHOOD

The inverse of the covariance matrix is used in the likelihood that infers cosmology from the data. Due to the large dimensionality of the covariance matrix in current cosmological studies and the fact that we have a limited number of pseudo-samples to internally estimate the covariance matrix, we often end up with a noisy covariance matrix e.g. from jackknife resampling. These noisy covariance matrices lead to biased estimates of the inverse covariance thereby biasing our estimates for cosmological parameters.

Efforts have been made to correct for this effect due to a finite number of samples from simulations of mock data. Hartlap, Simon, & Schneider (2007) tries to de-bias the inverse of the covariance matrix $\mathbf{C}^{-1}$ while keeping the likelihood Gaussian. The corrected inverse is estimated by

$$C_{\text{corrected}}^{-1} = \frac{N_{\text{sim}} - N_{\text{bin}} - 2}{N_{\text{sim}} - 1} C^{-1} \quad (A1)$$

where $N_{\text{bin}}$ is the number of angular bins and $N_{\text{sim}}$ is the number of samples the covariance is generated from. However, Sellentin & Heavens (2016) noted that the likelihood is more consistent with a $t$-distribution rather than a Gaussian due to marginalization over the noisy covariance matrix.

In order to investigate this effect in our data, we have calculated the correction both ways using CosmoSIS. Fig. A1 demonstrates that the differences between these two corrections (and compared to the case without a correction) are insignificant for our data and in fact, are mostly within the numerical noise. This is due to the fact that $N_{\text{sim}}$ is sufficiently large in our sample. Note that the cosmic shear probe does not place strong constraints on $h$, $\Omega_b$, and $n_s$, as seen from this figure; however, they will become more relevant in a full 3 \times 2pt analysis. The fiducial points (green stars) for the shear calibration parameters are not supplied by the Buzzard simulation, contrary to the cosmological parameters that are used to generate the simulation. They all display a fixed “non-tomographic” Gaussian prior mean of 0.02 which was the basis for the sampling. We, therefore, do not expect them to remain constant across different tomographic redshift bins.

We choose to use the Sellentin-Heavens correction throughout this paper with a caveat that the effective number of independent samples, which is assumed to be $N_{\text{sim}} = 2,000$, may not be known for jackknife covariances. However, it might be better to use a correction on a jackknife covariance than to not use any correction at all. The $t$-distribution will widen the likelihood tails, and that certainly has to be better (especially if one is investigating few-sigma tensions) than the Gaussian, which cuts off sharply. For that reason, the Hartlap correction (equation A1) would not be a viable idea.

APPENDIX B: RESULTS WITH THE GAUSSIAN COVARIANCE MATRIX

We repeat our analysis with the analytical Gaussian covariance matrix described in Section 6, instead of the jackknife covariance matrix. The results are shown in Figs B1–B4.
Blending effects on cosmic shear

Figure A1. Cosmological parameter estimation based on cosmic shear in the non-blended spectroscopic sample and two different corrections, Anderson-Hartlap and Sellentin-Heavens, on the covariance matrix produced by the jackknife estimator. The contours of the chains with not-corrected, Anderson-Hartlap corrected and Sellentin-Heavens corrected covariance matrices are displayed in filled blue, unfilled red (solid line) and unfilled black (dashed line), respectively. In each case, the contours drawn correspond to 1σ (68%) and 2σ (95%) confidence levels. The green star marks the fiducial parameters. The three sets of posteriors do not differ significantly except that the non-corrected case imposes tighter constraints for $\Omega_m$ and $\sigma_8$. 

---

MNRAS 000, 1–20 (2021)
Figure B1. Same as Fig. 14 but with a Gaussian covariance matrix for each of the two cases.

Figure B2. Same as Fig. 15 but with a Gaussian covariance matrix for each of the two cases.

Figure B3. Same as Fig. 16 but with a Gaussian covariance matrix for each of the three cases.

Figure B4. Same as Fig. 17 but with a Gaussian covariance matrix for each of the three cases.