







# Optical selection bias and projection effects in stacked galaxy cluster weak lensing

Hao-Yi Wu<sup>1</sup>    Matteo Costanzi<sup>2,3,4</sup>  Chun-Hao To<sup>5,6,7</sup>  Andrés N. Salcedo<sup>8</sup>  David H. Weinberg<sup>5,6,7</sup> James Annis,<sup>9</sup> Sebastian Bocquet,<sup>10</sup> Maria Elidaiana da Silva Pereira,<sup>11</sup> Joseph DeRose,<sup>12</sup> Johnny Esteves,<sup>13</sup> Arya Farahi,<sup>14</sup> Sebastian Grandis,<sup>10</sup> Eduardo Rozo,<sup>8</sup> Eli S. Rykoff,<sup>15,16</sup> Tamás N. Varga,<sup>17,18</sup> Risa H. Wechsler,<sup>15,16,19</sup> Chenxiao Zeng,<sup>5,6</sup> Yuanyuan Zhang,<sup>9</sup> and Zhuowen Zhang<sup>20</sup> (DES Collaboration)

*Affiliations are listed at the end of the paper*

Accepted 2022 July 18. Received 2022 June 15; in original form 2022 March 9

## ABSTRACT

Cosmological constraints from current and upcoming galaxy cluster surveys are limited by the accuracy of cluster mass calibration. In particular, optically identified galaxy clusters are prone to selection effects that can bias the weak lensing mass calibration. We investigate the selection bias of the stacked cluster lensing signal associated with optically selected clusters, using clusters identified by the redMaPPer algorithm in the Buzzard simulations as a case study. We find that at a given cluster halo mass, the residuals of redMaPPer richness and weak lensing signal are positively correlated. As a result, for a given richness selection, the stacked lensing signal is biased high compared with what we would expect from the underlying halo mass probability distribution. The cluster lensing selection bias can thus lead to overestimated mean cluster mass and biased cosmology results. We show that the lensing selection bias exhibits a strong scale dependence and is approximately 20–60 per cent for  $\Delta\Sigma$  at large scales. This selection bias largely originates from spurious member galaxies within  $\pm 20$ – $60 h^{-1}$  Mpc along the line of sight, highlighting the importance of quantifying projection effects associated with the broad redshift distribution of member galaxies in photometric cluster surveys. While our results qualitatively agree with those in the literature, accurate quantitative modelling of the selection bias is needed to achieve the goals of cluster lensing cosmology and will require synthetic catalogues covering a wide range of galaxy–halo connection models.

**Key words:** gravitational lensing: weak – galaxies: clusters: general – cosmology: theory.

## 1 INTRODUCTION

The number density of galaxy clusters as a function of mass and redshift is a sensitive probe for the growth rate of large-scale structure and the nature of cosmic acceleration (e.g. Vikhlinin et al. 2009; Mantz et al. 2010, 2014; Rozo et al. 2010; Bocquet et al. 2015, 2019; de Haan et al. 2016; Planck Collaboration XXIV 2016; Costanzi et al. 2019b, 2021; To et al. 2021a; also see e.g. Frieman, Turner & Huterer 2008; Allen, Evrard & Mantz 2011; Weinberg et al. 2013; Huterer et al. 2015 for reviews). Among different observational techniques, wide-field optical imaging surveys simultaneously provide large cluster samples and allow for weak gravitational lensing mass calibration. The upcoming Vera C. Rubin Observatory Legacy Survey of Space and Time (LSST), *Euclid*, and the *Nancy Grace Roman Space Telescope*’s High Latitude Survey have the potential to achieve unprecedented precision in cluster lensing and constraints on cosmological parameters (e.g. Sartoris et al. 2016; Eifler et al. 2021; Wu et al. 2021).

Precision cosmology from optical clusters relies on unbiased mass calibration for the underlying dark matter haloes. From survey data, one first identifies clusters as overdensities of galaxies in the sky and assigns each cluster a mass proxy, e.g. richness ( $\lambda$ ), the number of cluster member galaxies. Since the weak lensing signal for a single cluster usually has an insufficient signal-to-noise ratio for mass calibration, it is common to combine the lensing signal for clusters in a richness range and use this stacked signal to calibrate their mean mass (see e.g. Johnston et al. 2007; Rozo et al. 2010; Umetsu et al. 2014; Simet et al. 2017; Murata et al. 2018, 2019; McClintock et al. 2019). The number counts of galaxy clusters and their mean mass are used to constrain cosmology. In this process, biased mass calibration would lead to biased cosmological constraints.

This work focuses on the biased cluster lensing signal associated with optical selection. In the weak lensing mass calibration process, one usually assumes that the richness-selected cluster sample has an unbiased weak lensing signal for clusters of the same mass. However, if the richness selection preferentially includes clusters with higher lensing signals at a given mass, the stacked lensing signal and the derived mean mass would be biased high. We refer to this as the optical selection bias in cluster lensing and mass calibration.

\* E-mail: [hywu@boisestate.edu](mailto:hywu@boisestate.edu)

The recent cosmological results from the Dark Energy Survey Year 1 (DES Y1) cluster abundance and lensing (Abbott et al. 2020, hereafter **Y1CL**) suggest the presence of significant systematic bias in mass calibration associated with the optical cluster sample defined by the redMaPPer algorithm (Rykoff et al. 2014, 2016). In particular, **Y1CL** include an analysis that fixes the cosmology to the DES Y1  $3 \times 2$  point results (derived from galaxy clustering, galaxy–galaxy lensing, and cosmic shear; see Abbott et al. 2018) to derive the mass–observable relation using number counts. The mass–observable relation derived this way indicates that, without correcting for the optical selection bias, the lensing mass calibration is biased high  $\sim 10$  per cent for high-richness clusters ( $\lambda > 30$ ). On the other hand, the low-richness clusters ( $\lambda < 30$ ) have unexpectedly low lensing signals in **Y1CL**. In addition, To et al. (2021a) combine the DES Y1 cluster data (abundance, lensing, and clustering) with galaxy lensing and clustering and show that the impact of selection bias on the cluster lensing signal is approximately 15 per cent and is scale independent at  $\gtrsim 8 h^{-1}$  Mpc.

In this work, we show that the cluster lensing selection bias largely originates from projection effects, and to a smaller extent from the effect of halo concentration. We use the term ‘projection effect’ to collectively refer to changes in cluster lensing and richness due to matter and galaxies projected along the line of sight. The impact of projection effects on lensing can be further split into the non-spherical matter distribution inside the halo’s virial radius and the matter in filaments in large-scale structure. We will sometimes refer to the former as orientation bias, but we consider it to be a subset of projection effects, and the division between these two contributions is only approximate. Simulations have shown that dark matter haloes tend to have triaxial shapes, i.e. 3D ellipsoids described by three different axis lengths (see e.g. Jing & Suto 2002; Kasun & Evrard 2005; Bett 2012; Wu et al. 2013a; Bonamigo et al. 2015). If a cluster sample is dominated by clusters with major axes aligned along the line of sight, then the stacked lensing signal would be boosted relative to the prediction based on spherically symmetric haloes (e.g. Gavazzi 2005; Oguri et al. 2005; Corless & King 2007; Limousin et al. 2013; Dietrich et al. 2014; Osato et al. 2018; Park et al. 2021; Herbonnet et al. 2022; Zhang et al. 2022).

Projection effects change cluster richness by including galaxies outside the halo virial radius in the richness calculation (e.g. Erickson, Cunha & Evrard 2011; Köhlinger, Hoekstra & Eriksen 2015; Rozo et al. 2015; Farahi et al. 2016; Wojtak et al. 2018; Costanzi et al. 2019a; Myles et al. 2021). In imaging surveys, cluster members are identified by their colours, and galaxies matching the colour criterion but outside the cluster virial radius can be misidentified as members. These galaxies are associated with uncorrelated background or correlated large-scale structure, and the latter tends to increase the lensing signal and richness simultaneously.

The cluster lensing selection bias can be mathematically described as the correlation between richness residual and lensing residual at a fixed mass, i.e. after removing the mass dependence of richness and lensing, a cluster with a large positive richness residual tends to also have a large positive lensing residual. If our cluster sample includes all possible richness values at a given mass, then we expect the stacked lensing signal to be unbiased. However, it is necessary to select clusters above a certain richness limit because clusters with very low richness tend to be spurious. Therefore, in the presence of correlated residuals, a richness selection would lead to a biased stacked lensing signal.

Given that projection effects coherently impact both richness and lensing signal, it is imperative to study richness and lensing simultaneously using simulations that self-consistently model cluster

richness and lensing signal. In this work, we use the Buzzard simulations for DES (DeRose et al. 2019, 2022), with cluster samples defined by the redMaPPer cluster finding algorithm. We first mimic the cluster selection in **Y1CL** and quantify the lensing selection bias. We then investigate the root cause of the selection bias by examining the underlying dark matter and galaxy distributions.

Recently, Sunayama et al. (2020, hereafter **S20**) use a halo occupation distribution (HOD) model to populate galaxies in  $N$ -body simulations and show that a redMaPPer-like cylindrical member selection leads to a 20 per cent lensing bias for scales  $\gtrsim 10 h^{-1}$  Mpc. As we will discuss in detail in Section 7, the lensing selection bias they find is in general lower than ours, and we attribute this to the differences in the galaxy models.

This paper is organized as follows. We briefly review cluster weak lensing and introduce the selection bias formalism in Section 2 and describe our simulations in Section 3. Section 4 presents our main results on the cluster lensing selection bias. We investigate the correlation between richness residual and lensing residual in Section 5 and the link between projection effects and selection bias in Section 6. Our results are discussed in Section 7 and summarized in Section 8.

We put most technical details and robustness tests in extensive appendices. Appendix A compares different versions of the Buzzard simulations, and Appendix B compares different diagnoses for calculating the selection bias. Appendix C shows that our results are robust against member galaxy selection criteria. Appendices D and E address the impact of halo triaxiality and concentration.

Throughout the paper we use the flat  $\Lambda$  cold dark matter ( $\Lambda$ CDM) cosmology implemented in the Buzzard simulations:  $\Omega_M = 0.286$ ,  $h = 0.7$ ,  $\sigma_8 = 0.82$ ,  $n_s = 0.96$ ,  $\Omega_B = 0.046$ , and  $N_{\text{eff}} = 3.046$ . All projected distances are in *physical* Mpc without  $h$  (denoted as pMpc),<sup>1</sup> while all line-of-sight distances are in comoving  $h^{-1}$  Mpc. All halo masses are virial mass  $M_{\text{vir}}$  (in the unit of  $h^{-1} M_\odot$ ) based on the redshift-dependent spherical overdensity provided in Bryan & Norman (1998).

## 2 FORMALISM OF CLUSTER COUNTS, WEAK LENSING, AND SELECTION BIAS

In this section, we briefly describe the formalism for modelling cluster number counts and stacked weak lensing, and we extend the formalism to model the selection bias. For comprehensive reviews for gravitational lensing, we refer readers to Bartelmann & Schneider (2001), Kilbinger (2015), and Umetsu (2020).

The weak lensing signal of a galaxy cluster is related to its excess surface mass density  $\Delta\Sigma$ , the surface density contrast at a projected distance  $r_p$ ,

$$\Delta\Sigma(r_p) = \bar{\Sigma}(< r_p) - \Sigma(r_p), \quad (1)$$

where  $\bar{\Sigma}(< r_p)$  is the cumulative mean surface mass density *within*  $r_p$ , and  $\Sigma(r_p)$  is the differential mean surface mass density at  $r_p$ ;  $\Sigma(r_p)$  can be calculated by integrating the line-of-sight 3D mass density distribution  $\rho(r)$ . Below we show the expressions for  $\Sigma$  and note that the expressions for  $\Delta\Sigma$  can be derived analogously.

<sup>1</sup>Using physical instead of comoving distances is somewhat unusual for a simulation-based cluster study. We have made this choice because the redMaPPer cluster boundaries are defined in physical units. We have checked that for our redshift bins of  $\Delta z = 0.15$ , stacking clusters using physical distances and comoving distances leads to negligible differences.

We focus on the number density and the mean lensing signal for a cluster sample defined by the richness range  $(\lambda_{\min}, \lambda_{\max})$  at a given redshift. The comoving number density of this sample is given by

$$n(\lambda_{\min}, \lambda_{\max}) = \int_{\lambda_{\min}}^{\lambda_{\max}} d\lambda \int dM \frac{dn}{dM} P(\lambda|M), \quad (2)$$

where  $P(\lambda|M)$  denotes the probability distribution function (PDF) of  $\lambda$  at a given mass, and  $dn/dM$  denotes the halo mass function and is determined by cosmological parameters. Because of the strong degeneracy between the parameters determining  $dn/dM$  and  $P(\lambda|M)$ , any bias in the latter will lead to biased cosmological parameters.

To calibrate  $P(\lambda|M)$ , we use the stacked lensing signal of the same cluster sample, which is given by

$$\Sigma(\lambda_{\min}, \lambda_{\max}) = \frac{\int_{\lambda_{\min}}^{\lambda_{\max}} d\lambda \int dM \frac{dn}{dM} P(\lambda|M) \langle \Sigma|M \rangle}{\int_{\lambda_{\min}}^{\lambda_{\max}} d\lambda \int dM \frac{dn}{dM} P(\lambda|M)}, \quad (3)$$

where  $\langle \Sigma|M \rangle$  is the mean lensing signal at a given mass,

$$\langle \Sigma|M \rangle = \int d\Sigma \Sigma P(\Sigma|M). \quad (4)$$

The equations above assume that  $\lambda$  and  $\Sigma$  are uncorrelated. However, if  $\Sigma$  is correlated with  $\lambda$  at a given mass,  $P(\lambda|M)$  and  $P(\Sigma|M)$  need to be replaced by a joint probability distribution  $P(\lambda, \Sigma|M)$ . The stacked lensing signal is then given by

$$\Sigma(\lambda_{\min}, \lambda_{\max}) \propto \int_{\lambda_{\min}}^{\lambda_{\max}} d\lambda \int dM \frac{dn}{dM} \int d\Sigma \Sigma P(\lambda, \Sigma|M). \quad (5)$$

To proceed, we need to model  $P(\lambda, \Sigma|M)$ . It has been shown that at fixed mass,  $\lambda$  is well described by a lognormal distribution (e.g. Anbajagane et al. 2020; To et al. 2021a). Similarly, we have verified from our simulations that  $\Sigma$  at fixed mass can be described by a lognormal distribution. To account for the correlation between the two observables, we consider a bivariate lognormal distribution. We note that deviations from this assumption could exist and will require further numerical modelling.

Below we analytically model the selection bias by assuming that  $\ln \lambda$  and  $\ln \Sigma$  follow a bivariate Gaussian distribution. Let us assume this joint PDF has a mean  $(\langle \ln \lambda|M \rangle, \langle \ln \Sigma|M \rangle)$  and a covariance matrix

$$\mathcal{C} = \begin{pmatrix} \sigma_{\ln \lambda}^2 & r \sigma_{\ln \lambda} \sigma_{\ln \Sigma} \\ r \sigma_{\ln \lambda} \sigma_{\ln \Sigma} & \sigma_{\ln \Sigma}^2 \end{pmatrix}, \quad (6)$$

where  $\sigma_{\ln \lambda}$  and  $\sigma_{\ln \Sigma}$  are the standard deviations of the two observables at a given mass, and  $r$  is the correlation coefficient between the residuals,  $\text{corr}[\ln \lambda - \langle \ln \lambda|M \rangle, \ln \Sigma - \langle \ln \Sigma|M \rangle]$ . The PDF of the lensing signal associated with the richness selection is given by

$$p(\ln \Sigma | \ln \lambda, M) = \frac{P(\ln \Sigma, \ln \lambda | M)}{P(\ln \lambda | M)} \\ \propto \frac{\text{bivariate Gaussian}(\langle \ln \lambda|M \rangle, \langle \ln \Sigma|M \rangle, \mathcal{C})}{\text{Gaussian}(\langle \ln \lambda|M \rangle, \sigma_{\ln \lambda}^2)}. \quad (7)$$

Under this assumption, the conditional probability distribution of  $\ln \Sigma$  given  $\ln \lambda$  corresponds to a Gaussian distribution with the mean

$$\langle \ln \Sigma | \ln \lambda, M \rangle = \langle \ln \Sigma | M \rangle + r \sigma_{\ln \Sigma} \frac{(\ln \lambda - \langle \ln \lambda|M \rangle)}{\sigma_{\ln \lambda}}. \quad (8)$$

The last term describes the selection bias associated with cluster lensing (also see White, Cohn & Smit 2010; Evrard et al. 2014

for similar derivations). This formalism is also mathematically equivalent to the impact of halo assembly bias on cluster lensing and clustering (see e.g. Wu, Rozo & Wechsler 2008).

The correlation between the  $\lambda$  residual and the  $\Sigma$  residual at a fixed mass is the essence of the cluster weak lensing selection bias. In Section 5, we will quantify the correlated residuals in the mock redMaPPer cluster sample.

### 3 GALAXY CLUSTERS IN THE BUZZARD SIMULATIONS

In this section, we describe our simulated redMaPPer clusters from the Buzzard simulations.

#### 3.1 The Buzzard simulations

The Buzzard simulations (DeRose et al. 2019, 2022) include a suite of synthetic catalogues based on the ADDGALS algorithm (Wechsler et al. 2022) and are designed for supporting the DES data analyses. Below we describe the simulation framework, and in Appendix A, we describe the particular versions of Buzzard we use.

The first step of the simulation creates a galaxy catalogue with  $r$ -band magnitudes. The algorithm performs subhalo abundance matching between an  $N$ -body simulation with well-resolved subhaloes and the observed luminosity function in the  $r$ -band. Using the resulting galaxy catalogue, the algorithm calibrates (1) the relation between mass and  $r$ -band magnitude for central galaxies, and (2) the relation between local density and  $r$ -band magnitude for satellite galaxies. These relations are then used to assign galaxies to resolved haloes or dark matter particles in a large light-cone simulation with a lower resolution.

The second step of the simulation assigns colour to each galaxy. At a given  $r$ -band magnitude, galaxies from the observed catalogue are ranked by their  $g - r$  colour, and galaxies from the simulated catalogue are ranked by an environmental proxy, such as the distance to the nearest neighbour. The algorithm then performs abundance matching between these two ranked lists with some scatter, and the observed galaxy spectral energy distributions (SEDs) are assigned to the corresponding galaxies in the simulated catalogue.

In this work, we use 12 realizations of the DES Y1 data (1120 deg<sup>2</sup> each) based on Buzzard version 1.9.2+1 and one realization of the DES Y3 data (4143 deg<sup>2</sup>) based on Buzzard version 1.9.8 (presented in DeRose et al. 2019, 2022, respectively). In Appendix A, we compare the two versions. The key difference relevant for this work is that the Y1 realizations have a narrower red sequence than the Y3 realization. We find that the lensing selection bias results from these two versions are statistically consistent, and therefore in the main text we combine the results from both versions.

Wechsler et al. (2022) show that the massive haloes in Buzzard tend to have fewer member galaxies than observed, and this deficit is attributed to the artificial disruption of subhaloes in dense environments in the  $N$ -body simulation used for subhalo abundance matching. As a result, clusters in Buzzard tend to have lower richness values compared with observed clusters of similar mass, and above a richness threshold Buzzard has fewer clusters than observed. In addition, To et al. (2021a) show that the selection bias in Buzzard is higher than that indicated by the DES Y1 data at large scales, which can also be attributed to the low galaxy density in clusters in Buzzard. Because of this discrepancy, we do not directly use the selection bias derived from Buzzard to correct the observed lensing signal. Instead, we use Buzzard to study the nature of selection bias

and leave to future works a full calibration using suites of galaxy–halo connection models.

### 3.2 The redMaPPer cluster finding algorithm

The redMaPPer algorithm (Rykoff et al. 2014, 2016) identifies galaxy clusters in a photometric galaxy catalogue based on the red sequence, i.e. the tight colour–magnitude relation for galaxies in clusters. The algorithm first calibrates a red-sequence template using a sample with both photometric and spectroscopic information. This template is then used to select red galaxies as possible central galaxies.

For a central galaxy, the algorithm finds its candidate member galaxies and assigns each member a membership probability, the probability that a galaxy is a true cluster member calculated based on its magnitude, colour, and distance to the central galaxy. The member galaxies are selected in a projected aperture  $R_\lambda$ , which is iteratively calculated to match the relation  $R_\lambda = 1(\lambda/100)^{0.2} h^{-1}$  Mpc (physical). This relation has been calibrated to minimize the scatter of  $L_X$  given  $\lambda$  (Rykoff et al. 2012).

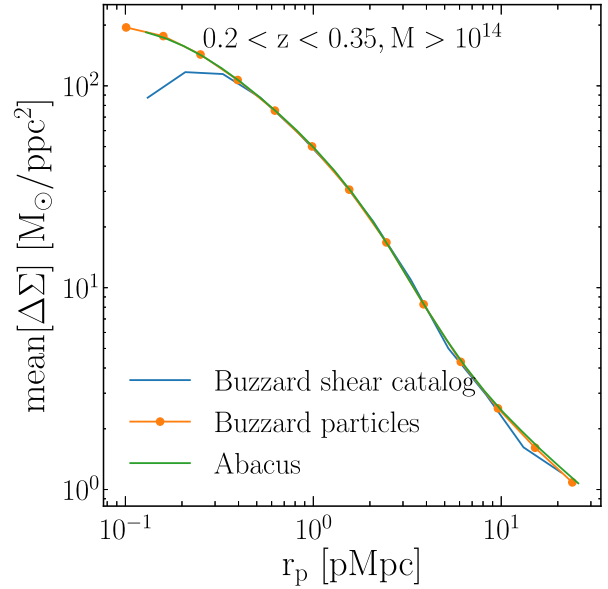
In a process called percolation, all possible central galaxies are ranked by a preliminary richness, and a higher ranked central galaxy is prioritized in obtaining its members, i.e. for a candidate galaxy member in the vicinity of two possible central galaxies, it will be assigned a higher membership probability to the higher ranked central.

In this work, we use the redMaPPer version 6.4.22, which is essentially the same version as used in DES Y1CL (see McClintock et al. 2019). We use the halo centres as cluster centres to calculate the richness to avoid mismatched halo–cluster pairs and misidentified central galaxies (e.g. Zhang et al. 2019).

### 3.3 Measuring stacked cluster lensing in simulations

We use dark matter particles to calculate the surface mass density  $\Sigma$  and excess surface density  $\Delta\Sigma$  of clusters, using cylinders of depth comoving  $\pm 100 h^{-1}$  Mpc. This projection depth is sufficient to account for the correlated structure along the line of sight (also see Osato et al. 2018), and we have tested that using cylinders of  $\pm 200 h^{-1}$  Mpc or using particles in cones leads to negligible differences. Because Buzzard joins two different boxes at  $z = 0.32$ , we discard haloes within  $\pm 100 h^{-1}$  Mpc of this discontinuity boundary. We calculate the profiles for all haloes in the parent  $N$ -body simulations with  $M_{\text{vir}} \geq 10^{13} h^{-1} M_\odot$ , regardless of whether a halo is in the mock cluster catalogue or not. As we will show below, we use all these haloes to form a control sample for lensing signals. We additionally calculate the triaxial shape of each halo using the dark matter particles within  $R_{\text{vir}}$  (see Appendix D).

Fig. 1 compares the  $\Delta\Sigma$  profiles for haloes with  $0.2 < z < 0.35$  and  $M_{\text{vir}} \geq 10^{14} h^{-1} M_\odot$ , calculated from particles and from the ray-tracing shear catalogue. We use the true redshift and shear in the shear catalogue, ignoring photometric errors and intrinsic galaxy ellipticities. The ray tracing and particle calculations agree with each other at large scales, while the former does not have sufficient resolution below 0.4 pMpc. To test the resolution limit of Buzzard we use the Abacus Cosmos simulations (Garrison et al. 2018); we use the dark matter particles in the  $720 h^{-1}$  Mpc boxes at  $z = 0.3$ , which have three times better mass resolution than Buzzard.<sup>2</sup> The comparison with Abacus shows that the spatial resolution of Buzzard



**Figure 1.** Mean  $\Delta\Sigma$  profiles from the Buzzard simulations, calculated using dark matter particles and the ray-tracing shear catalogue, compared with the Abacus Cosmos simulations with three times higher mass resolution. The Buzzard particle calculations have adequate resolution down to 0.1 pMpc. The pMpc and ppc in the axis labels refer to physical megaparsec and parsec.

is adequate down to 0.1 pMpc. Scales below 0.1 pMpc are not usually included in weak lensing cluster mass calibration.

## 4 QUANTIFYING THE SELECTION BIAS IN THE STACKED REDMAPPER CLUSTER LENSING SIGNAL

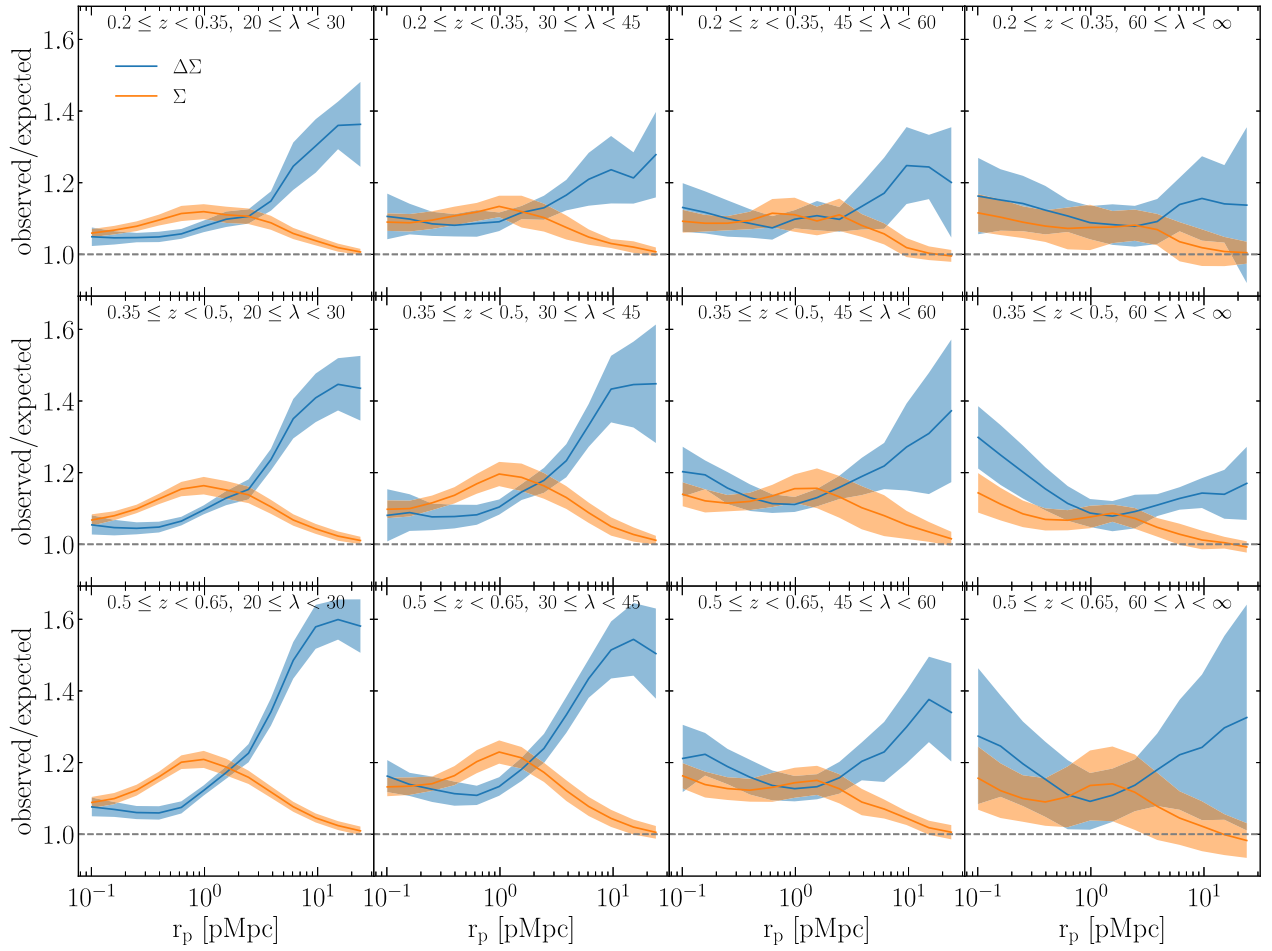
With the simulated cluster lensing signal described above, we calculate the stacked lensing signal in redshift and richness bins. We use the same binning as in DES Y1CL: three redshift bins bounded by (0.2, 0.35, 0.5, 0.65), and four richness bins bounded by (20, 30, 45, 60,  $\infty$ ).

To quantify the cluster lensing selection bias, we compare the stacked lensing signal of clusters selected by richness with the signal expected from the underlying dark matter halo mass PDF. This ‘expected’ signal accounts for the scatter between richness and mass but assumes uncorrelated residuals between richness and lensing at a given mass. In Appendix B, we detail three methods for calculating the expected lensing signal from this mass PDF and show that they give consistent results. In the main text, we present the ‘weighting’ method.

Fig. 2 shows the selection bias for  $\Sigma$  and  $\Delta\Sigma$ . We show the mean and standard deviation calculated from the 12 realizations of the DES Y1 data and the one realization of the DES Y3 data, weighted by the area. For  $\Sigma$ , all richness bins exhibit biases of approximately 10–20 per cent, with strong scale dependence. For the low-richness bins, the selection bias peaks at approximately 1 pMpc and is weak at small and large scales. For the high-richness bins, the selection bias is substantial at small scales. In all cases, the selection bias for  $\Sigma$  vanishes at scales  $\gtrsim 20$  pMpc. In contrast to  $\Sigma$ , the bias of  $\Delta\Sigma$  is non-vanishing at large scales because  $\Delta\Sigma$  at each  $r_p$  contains the information of  $\Sigma$  from  $r < r_p$  (equation 1). The bias in  $\Delta\Sigma$  can be as high as 20–60 per cent at large  $r_p$ .

We would like to understand to what extent the biased lensing profile presented in Fig. 2 is caused by a biased 3D density profile. To

<sup>2</sup>The Abacus Cosmos simulations are based on  $M_{200m}$ , a different cosmology, and a different redshift, and thus we only use them for resolution comparison.

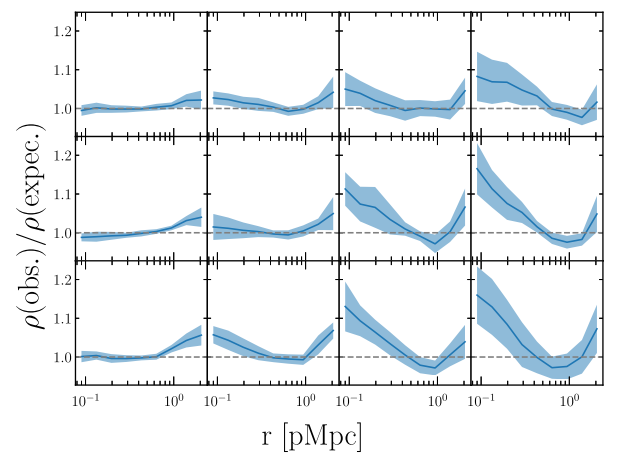


**Figure 2.** Cluster lensing selection bias of  $\Sigma$  and  $\Delta\Sigma$ , quantified by the ratio between the signal from a richness-selected sample (observed) and the signal expected from the underlying halo mass PDF (expected). We present different redshift bins (rows) and richness bins (columns). For  $\Sigma$ , the selection bias peaks at approximately 1 pMpc and vanishes at large scales. For  $\Delta\Sigma$ , the selection bias includes the small-scale effects and can be as high as 20–60 per cent at large scales.

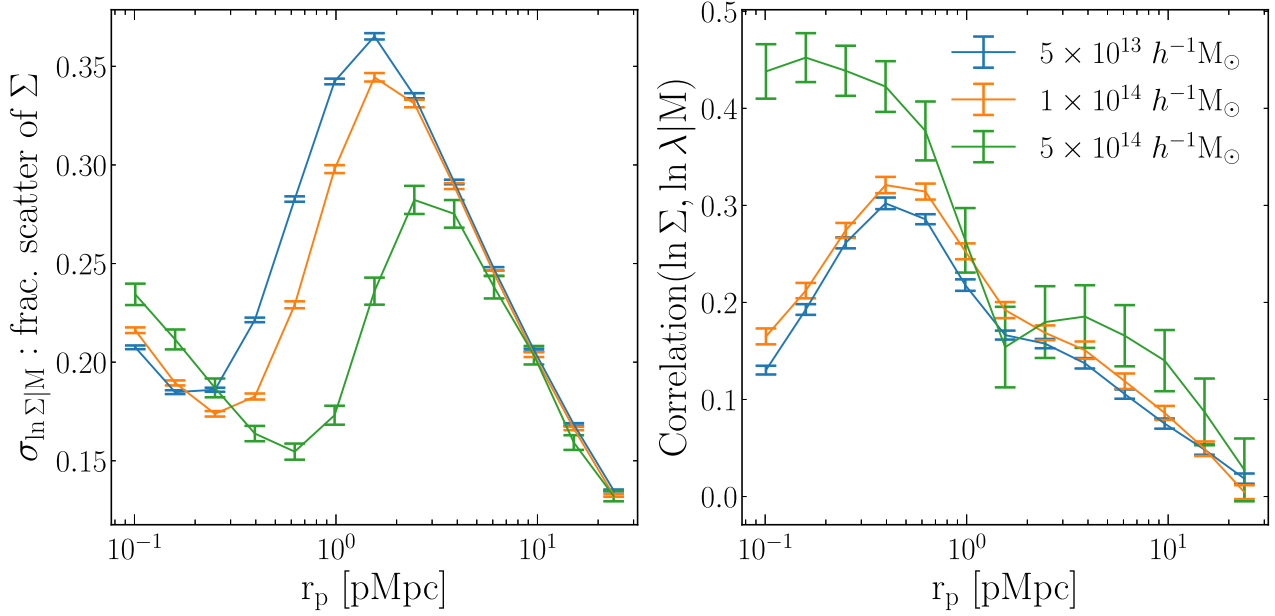
answer this, we repeat the selection bias calculation for the 3D density profile  $\rho(r)$ . For each halo, we calculate the spherically averaged  $\rho(r)$  by counting dark matter particles in spherical shells around the halo centre. We then calculate the mean  $\rho$  in a richness bin and the  $\rho$  expected from haloes with the same mass PDF.

Fig. 3 shows the selection bias of 3D density profiles  $\rho$  out to 3 pMpc. For low-richness clusters, the selection bias of  $\rho$  is negligible, while for high-richness clusters, the small scales exhibit a  $\sim 10$  per cent selection bias in  $\rho$ . This difference in small-scale behaviour explains the difference between high- and low-richness clusters shown in Fig. 2. For low-richness clusters, the small-scale lensing selection bias is associated with the 2D projection, while for high-richness clusters, part of the small-scale selection bias is due to the biased 3D density profiles.

Fig. 3 implies that our high-richness sample preferentially selects haloes with higher 3D density at small scales at a given mass. We expect that these haloes have higher concentrations. We investigate the influence of halo concentration in Appendix E. Fig. E1 compares the concentration distribution for a richness-selected sample and for a sample with the same mass PDF. As expected from Fig. 3, the high-richness clusters tend to have higher concentrations than haloes of the same mass PDF, while the low-richness clusters do not show such a bias. Fig. E2 shows how concentration affects the lensing



**Figure 3.** Selection bias in 3D density profiles  $\rho(r)$ , analogous to Fig. 2 with matching panels. While low-richness clusters exhibit negligible selection bias in  $\rho$ , high-richness clusters show  $\sim 10$ – $15$  per cent selection bias in  $\rho$  at small scales. This explains the different scale dependence of  $\Sigma$  bias for low- and high-richness clusters seen in Fig. 2.



**Figure 4.** Left: scatter of the lensing signal at a given mass as a function of projected radius, for  $0.2 \leq z < 0.35$  and three halo masses. Right: correlation between the lensing residual and the richness residual at a given mass, as a function of projected radius. The selection bias is proportional to the product of the lensing scatter and the correlation coefficient (equation 8), and their scale dependence explains the scale dependence of the selection bias seen in Fig. 2.

profile. The dependence of  $\Sigma$  on halo concentration is quite different from what we see in Fig. 2, and therefore the halo concentration has limited predictive power for the selection bias.

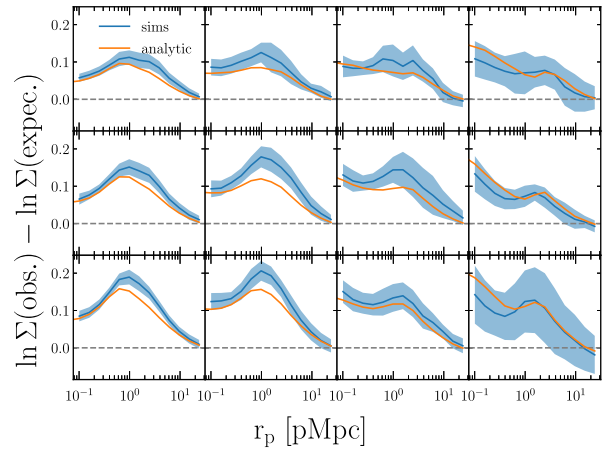
We also examine the selection bias associated with the triaxial halo orientation in Appendix D. Similar to concentration, our cluster sample preferentially selects haloes with major axes parallel to the line of sight. However, the orientation and concentration selection cannot fully account for the selection bias we find (Fig. E3).

In the next sections, we will investigate the origin of the selection bias. In Section 5, we will calculate the correlation between the richness residual and the lensing residual at a given halo mass. In Section 6, we will examine the relation between projection effects and selection bias.

## 5 CORRELATION BETWEEN THE RICHNESS RESIDUAL AND THE LENSING RESIDUAL AT A GIVEN HALO MASS

We compare the simple model presented in equation (8) with the selection bias results from simulations. We first calculate the standard deviations ( $\sigma_{\ln \lambda}$ ,  $\sigma_{\ln \Sigma}$ ) and correlation coefficient ( $r$ ) from simulations. To capture their mass dependence, we use the Kernel Localized Linear Regression method (Farahi, Anbajagane & Evrard 2022). In this algorithm, each halo is assigned a Gaussian kernel centred on its  $\ln M$ ; for a given mass, the linear regression is performed with each halo weighted by this kernel. We combine all 13 Buzzard realizations, split haloes above  $10^{13} h^{-1} M_\odot$  into 20 log-mass bins, and use redshift bins  $\Delta z = 0.15$ . We choose a Gaussian kernel width of 0.2, perform the regression independently for each  $r_p$ , and use 100 bootstrap samples to estimate the error bars.

The left-hand panel of Fig. 4 shows  $\sigma_{\ln \Sigma}$  as a function of projected radius, for three halo masses. The scale dependence is non-monotonic and exhibits a peak at  $\approx 2$  pMpc. The right-hand panel shows the correlation between lensing residual and richness residual at a given mass,  $r = \text{corr}[\ln \lambda - \langle \ln \lambda | M \rangle, \ln \Sigma - \langle \ln \Sigma | M \rangle]$ . For  $5 \times 10^{13}$  and  $10^{14} h^{-1} M_\odot$ , the correlation peaks at  $\approx 0.4$  pMpc, while for



**Figure 5.** Lensing selection bias estimated from equation (8), shown in orange, based on the correlated residuals between richness and lensing at a given mass. The blue bands show the simulation results (same as Fig. 2).

$5 \times 10^{14} h^{-1} M_\odot$ , the correlation is the largest at small radii. From equation (8), we can see that the selection bias is proportional to the product of  $\sigma_{\ln \Sigma}$  and  $r$ , and the scale dependence we see in Fig. 2 can be explained by the scale dependence shown here.

Fig. 5 shows that the prediction from equation (8) agrees well with the selection bias shown in Fig. 2. In this calculation, we first apply equation (8) to each halo and then average over all haloes in a given richness–redshift bin. The small discrepancy is due to the deviations from the Gaussian assumption associated with equation (8).

To improve our understanding of selection bias, it is essential to calibrate each component in equation (8) using simulations and observations. The scatter of the lensing signal at a given mass can be estimated from simulations. However, in the absence of an accurate galaxy–halo connection model, the correlated residuals between lensing and richness need to be modelled empirically. One option is to

use simulations to motivate a functional form for the scale-dependent correlation between the residuals and use observed lensing profiles to constrain this correlation. In addition, one may use multiwavelength observations to calibrate such a correlation; for example, by studying lensing and richness for a sample selected based on the Sunyaev–Zeldovich (SZ) effect.

## 6 PROJECTION EFFECTS AND SELECTION BIAS

We investigate the origin of selection bias by examining the impact of line-of-sight projection on the redMaPPer richness. Since redMaPPer is a rather complex algorithm, we adopt a simplified approach: we use a basic colour–magnitude cut on galaxies and count the number of galaxies in a cylinder along the line of sight. Below we show that our simplified calculation can largely account for the lensing selection bias seen in Buzzard. In this calculation, we focus on the single DES Y3 realization because its red-sequence width is closer to that observed by DES.

### 6.1 Properties of the redMaPPer member galaxies

We first quantify the redMaPPer member galaxy properties in our simulation. We focus on galaxies with membership probability greater than 0.9, as they form a tight red sequence and contribute to most of the richness of a cluster.

#### 6.1.1 Magnitude

We have verified that the magnitude selection of our simulated redMaPPer member galaxies is consistent with the  $0.2L_*$  threshold in  $i$  band calibrated from the Sloan Digital Sky Survey (SDSS), as presented in equation (9) in Rykoff et al. (2014). We use this magnitude threshold for our fiducial calculation.

#### 6.1.2 Colour

The left-hand panel of Fig. 6 shows the medians and 68 per cent intervals of the member colours ( $g - r$ ,  $r - i$ ,  $i - z$ ), as a function of cluster redshift. At a given redshift, we find that the colour of red-sequence galaxies is approximately independent of the magnitude, and we use the median colour  $\langle c \rangle$  and 68 per cent interval  $\sigma_c$  as a simplified red-sequence template for selecting member galaxies.

For the calculations in this section, we select galaxies with

$$\chi_{\text{colour}}^2 = \sum_c (c - \langle c \rangle)^2 / \sigma_c^2 < 9, \quad (9)$$

assuming no correlation between colours. Fig. C1 shows that changing this  $\chi_{\text{colour}}^2$  threshold has little effect on the resulting lensing selection bias.

#### 6.1.3 Line-of-sight distances

We calculate the line-of-sight comoving distances between member galaxies and their host haloes. The right-hand panel of Fig. 6 shows the histogram for these distances for clusters in the range  $0.2 < z < 0.65$ . The prominent peak near 0 corresponds to the galaxies physically associated with the cluster, while the broad tails on both sides correspond to the projected members. The histogram shows a transition from cluster galaxies to background galaxies at approximately  $30\text{--}50 h^{-1}$  Mpc. A  $50 h^{-1}$  Mpc line-of-sight distance

includes approximately 90 per cent of the members. We will test the impact of projection depth in the calculations below.

### 6.2 Selecting member galaxies in a cylinder: impact of projection depth

We calculate the number of galaxies within a cylinder around a halo centre as our mock richness. We use galaxies from the GOLD catalogue derived from Buzzard, which represents the parent galaxy sample for the DES cosmology analyses (Drlica-Wagner et al. 2018). We use galaxies' 3D positions, magnitudes, and colours. We adopt the following fiducial choice: an aperture  $R_\lambda$  (from the redMaPPer output), a magnitude threshold of  $0.2L_*$  in  $i$  band, and a colour cut of  $\chi_{\text{colour}}^2 < 9$ . In addition, we mimic the percolation process in redMaPPer to avoid double counting member galaxies: if a galaxy falls in the cylinder of multiple haloes, it is only counted as the member of the most massive halo. We have found that including this process only slightly changes the results in the lowest richness bin.

In this section, we vary the depth of the cylinder from  $\pm 1$  to  $60 h^{-1}$  Mpc. In Appendix C, we compare different magnitude thresholds, colour thresholds, and apertures. For each of our cylinder-richness definitions, we use the same redshift bins as the redMaPPer calculation and split the cylinder richness into four bins, each of which has the same number of clusters as the redMaPPer  $\lambda$  bins.

Fig. 7 shows the selection bias in  $\Sigma$  associated with different projection depths. For all panels, a projection depth of  $20\text{--}60 h^{-1}$  Mpc gives a selection bias comparable to that in the redMaPPer sample (black curves). For a projection depth of  $1 h^{-1}$  Mpc, we see no selection bias for low-richness clusters but a significant selection bias for high-richness clusters for  $r_p < 1$  pMpc. This selection bias exists even when we eliminate the colour and magnitude selection. This is consistent with the  $\rho(r)$  bias shown in Fig. 3 and with the concentration bias shown in Fig. E1. In Buzzard, for haloes above  $\approx 5 \times 10^{14} h^{-1} M_\odot$ , high-concentration haloes tend to have higher richness. This is opposite to the theoretical expectation that high-concentration haloes tend to form earlier and have fewer surviving satellite galaxies (e.g. Wu et al. 2013b; Mao, Williamson & Wechsler 2015). This correlation between concentration and richness could be spurious and could lead to overestimated selection bias for high-richness clusters in Buzzard.

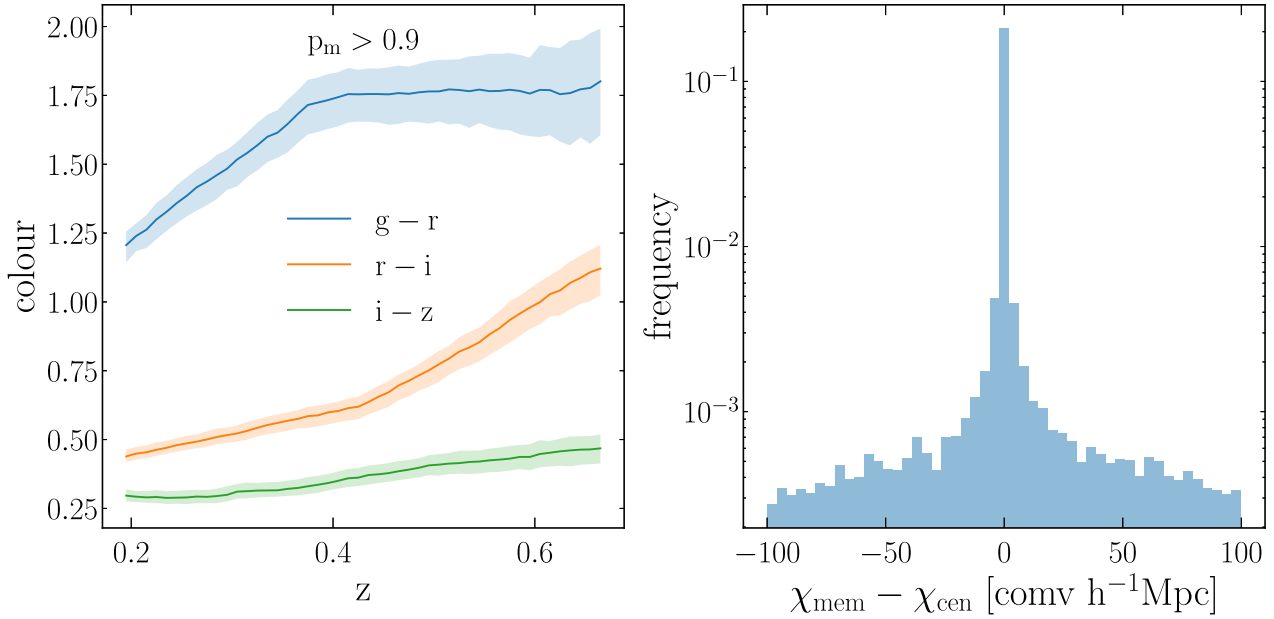
Given that the selection bias of redMaPPer is well approximated by selecting member galaxies in a cylinder, we can use this approach to systematically study cluster selection bias in a wide range of simulations. Since redMaPPer self-calibrates the red sequence, it can only be applied to mock catalogues with realistic galaxy colour; therefore, redMaPPer has only been applied to a limited number of mock catalogues. Our cylinder selection, on the other hand, can be readily applied to any mock galaxy catalogue for red galaxies and is computationally inexpensive. In our upcoming work, we plan to use this tool to study a wide range of mock galaxy catalogues, including those generated from hydrodynamic simulations and HOD models.

## 7 DISCUSSION

In this section, we compare our results with previous work and discuss mitigation strategies for selection bias.

### 7.1 Comparison with previous studies

S20 quantify the impact of projection effects on cluster lensing and clustering using HOD-based galaxy catalogues. They simulate the redMaPPer richness using cylinders of depth  $\pm 60 h^{-1}$  Mpc. We have



**Figure 6.** Left: the colour of the redMaPPer member galaxies in Buzzard as a function of cluster redshift. The lines and bands show the medians and the 68 per cent intervals. Right: the line-of-sight comoving distances between member galaxies and their host haloes. Here we consider galaxies with a membership probability greater than 0.9.

shown that the richness calculated from  $\pm 20$  to  $60 h^{-1}$  Mpc cylinders gives results broadly consistent with redMaPPer; therefore, their calculation procedures and ours are comparable. Overall, they find negligible selection bias at small scales, and their selection bias amplitude is lower than ours at large scales by approximately a factor of 2 (their fig. 6).

We attribute the difference between their results and ours to how galaxies are populated in  $N$ -body simulations. S20 assign galaxies to resolved haloes, while Buzzard assigns galaxies to both halo centres and dark matter particles. Therefore, Buzzard has a larger field galaxy population. In addition, S20 use an observationally constrained HOD, while Buzzard has lower galaxy number densities in cluster-size haloes compared with observations, resulting in lower cluster number counts above a given richness (see figs 4 and 8 in Wechsler et al. 2022 and figs 12 and 13 in DeRose et al. 2019). The combination of a lower cluster galaxy content and a larger background population compared with observations leads to stronger projection effects in Buzzard.

Several analyses also suggest that the lensing selection bias in Buzzard is higher than that in observations. In Y1CL, the lensing selection bias needed to reconcile cluster counts with DES  $3 \times 2$  point analysis is weaker than what we have calculated from Buzzard (their fig. 12). Combining cluster abundances, cluster lensing and clustering, and galaxy lensing and clustering, To et al. (2021a) find that the best-fitting selection bias in DES Y1 clusters is  $\approx 15$  per cent for scales  $\gtrsim 8 h^{-1}$  Mpc (comoving), which is smaller than that in Buzzard.

Because of the uncertainty associated with galaxy modelling, care must be taken when applying lensing selection bias derived from simulations to cosmological analyses. A conservative approach is to adopt a functional form for the selection bias motivated by simulations and let the observational data self-calibrate the model parameters. This is similar to the approach in To et al. (2021a), who model the large-scale selection bias with a power law in mass and fit for the model parameters (also see Park et al. 2021). However, in an analysis combining cluster counts and weak lensing such as

Y1CL, fully uninformative priors on the selection bias parameters would completely wash out the cosmological constraining power of the catalogue. It is thus necessary to develop informative priors using a wide range of galaxy models, which could be achieved by exploring HOD parameters.

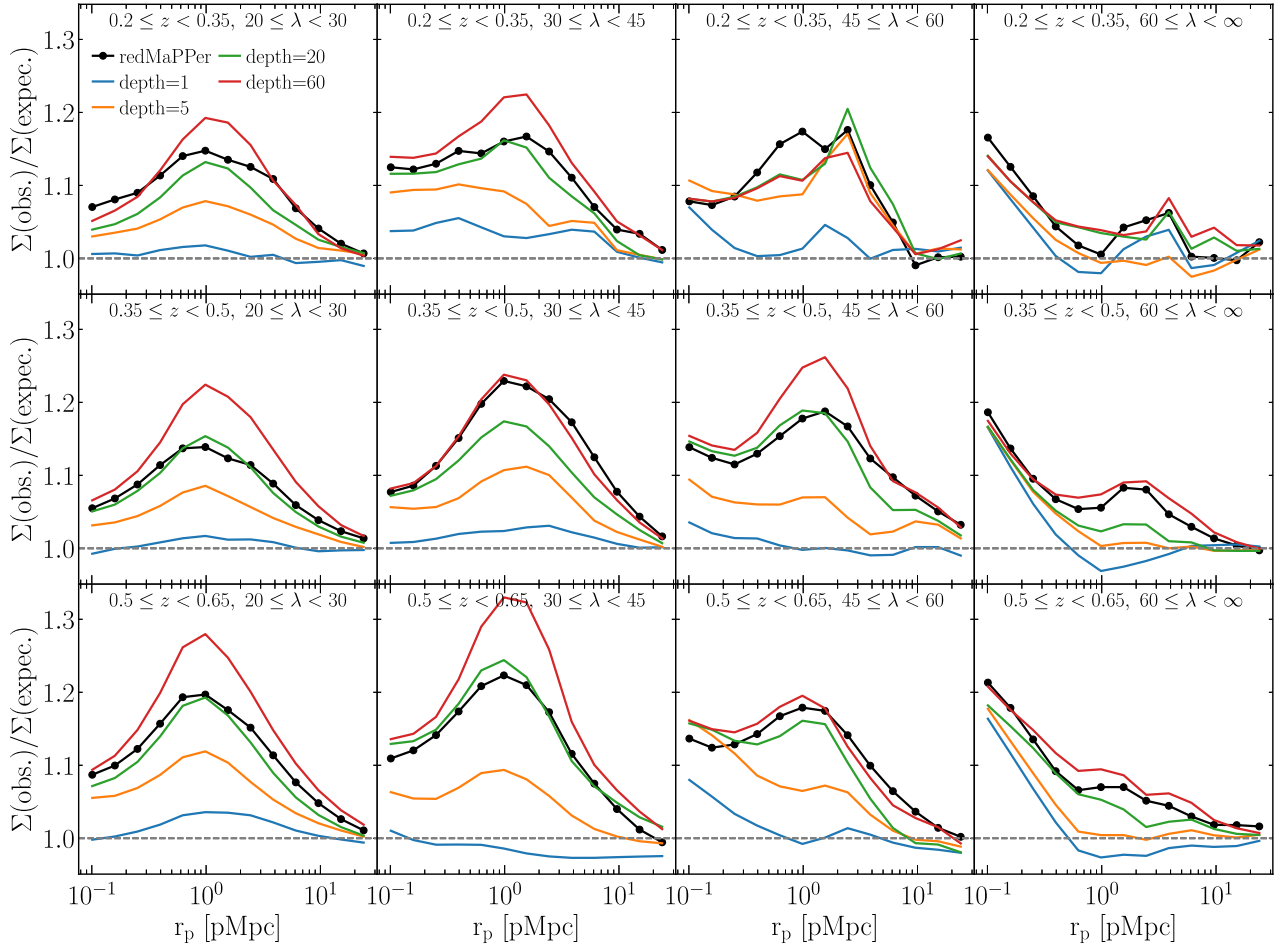
An alternative approach is to model the correlated residuals between observables and fit for the correlation coefficients (e.g. Grandis et al. 2021a; Chiu et al. 2022). This approach extends the analytic model described in Section 2 and can be part of the forward-modelling procedure. However, this approach could lead to many weakly constrained nuisance parameters, and informative priors on the correlation coefficients would also be necessary.

It would be valuable to develop emulators for calculating the selection bias for a range of galaxy models and cluster selection methods. Recent studies have used  $N$ -body simulation-based emulators to predict cluster lensing on non-linear scales (e.g. Nishimichi et al. 2019; Salcedo et al. 2020; Cromer et al. 2021), but they have so far used halo-based cluster selection. With cylinder selection as a proxy for the full redMaPPer selection, it may be feasible to construct emulators that directly model the cluster selection procedure.

Our results indicate that the redMaPPer cluster projection effects are dominated by galaxies within approximately  $\pm 60 h^{-1}$  Mpc along the line of sight, which corresponds to a redshift difference of 0.02. This is slightly smaller than the observational results in Myles et al. (2021). They fit the spectroscopic redshift distribution of galaxies associated with redMaPPer clusters using a double Gaussian mixture model to account for true and spurious members, finding that the latter has a standard deviation of approximately 0.03. Their result is consistent with Costanzi et al. (2019a), who develop a projection effect proxy  $\sigma_z$  to describe how widely cluster members are spread along the line of sight. We plan to use these observations to constrain the projection depth and to improve the modelling of projection effects in simulations.

Baryonic effects have been shown to alter the cluster lensing signal and bias the lensing-derived mass low by 5–10 per cent (e.g.





**Figure 7.** Lensing selection bias from our simplified richness proxies calculated with cylinders of various projection depths. A projection depth of  $\pm 20$  to  $60 h^{-1}$  Mpc reproduces the selection bias seen in the full redMaPPer calculation. For high-richness clusters, part of the selection bias is contributed by galaxies within  $\pm 1 h^{-1}$  Mpc.

Bahé, McCarthy & King 2012; Henson et al. 2017; Lee et al. 2018; Debackere, Schaye & Hoekstra 2021; Grandis et al. 2021b); however, selection bias is comparable or larger in magnitude, extends to larger scales, and is potentially more difficult to model because it depends on the uncertain relation between galaxies and haloes. Therefore, we expect that projection effects would be the most significant systematic uncertainty in optical cluster lensing. On the other hand, baryonic effects tend to make clusters more spherical (e.g. Bryan et al. 2013; Henson et al. 2017), which would reduce orientation bias. It would be valuable to use full hydrodynamic simulations with reliable galaxy populations to self-consistently study the projection-induced correlated residuals between lensing and richness.

Forecast studies show that the cosmological precision attainable from cluster weak lensing is competitive with that attainable from cosmic shear analyses of the same weak lensing data set, if the statistical limits can be achieved (e.g. Oguri & Takada 2011; Yoo & Seljak 2012; Weinberg et al. 2013; Salcedo et al. 2020; Wu et al. 2021). For example, Wu et al. (2021) find that a DES-like survey of cluster lensing could achieve 0.26 per cent precision on  $\sigma_8$  (with other cosmological parameters held fixed) if the mass–observable scatter is constrained independently, and Salcedo et al. (2020) forecast a  $\sigma_8$  precision of better than 1 per cent if the scatter is not known

independently but constrained by cluster–galaxy cross-correlations and galaxy autocorrelations. The challenge is to realize this statistical precision in the face of selection bias that affects  $\Sigma(r_p)$  at the 10–20 per cent level.

## 7.2 Mitigation strategies

Below we discuss strategies for mitigating the cluster lensing selection bias.

### 7.2.1 Simulating different galaxy models

As discussed earlier, the lensing selection bias depends on the underlying galaxy population. One way to reduce this modelling uncertainty is to quantify how projection effects depend on the HOD parameters of galaxies that contribute to the redMaPPer richness. Specific choices of HOD parameters have been studied (e.g. Costanzi et al. 2019a; S20), but quantifying their effects would require a systematic study of a wide range of HOD models. For example, HOD models with a larger satellite fraction or more galaxies in low-mass haloes would exhibit stronger projection effects.

### 7.2.2 Combining multiwavelength cluster observables

We can potentially use multiwavelength observations to quantify the optical selection bias. For example, one can use the SZ signal to select clusters and study the correlated residuals between richness and lensing, taking advantage of the small mass scatter of SZ-selected clusters. Such cluster samples exist for  $\lambda \gtrsim 50$  systems and have been used to calibrate the scatter of optical clusters (e.g. Rozo & Rykoff 2014; Saro et al. 2015; Farahi et al. 2019). In addition, the cross-comparison between optical and SZ clusters can be used to study projection effects (e.g. Grandis et al. 2021a); for example, a cluster heavily contaminated by galaxies along the line of sight would have a lower SZ signal than expected from its richness.

### 7.2.3 Combining cluster lensing and clustering

The multiwavelength approach described above is usually not applicable to low-richness clusters. To calibrate the selection bias for low-richness clusters, one can combine the clustering and lensing of galaxy clusters. At large scales, the selection bias manifests as the clustering bias, and we can calibrate it by combining cluster lensing, cluster–galaxy cross-correlation, and galaxy autocorrelation. This is similar to combining clustering and lensing to constrain the mass–observable relation (e.g. Chiu et al. 2020; Salcedo et al. 2020; To et al. 2021a).

Spectroscopic observations of member galaxies can distinguish true from spurious member galaxies (e.g. Rozo et al. 2015; Rines et al. 2018; Sohn et al. 2018; Myles et al. 2021; Wetzell et al. 2022). As we have shown in Section 6, galaxies with different line-of-sight distances lead to different amounts of selection bias. Therefore, quantifying the redshift distribution of the redMaPPer member galaxies associated with the line-of-sight structure would reduce the modelling uncertainties of selection bias. The Dark Energy Spectroscopic Instrument (DESI) and *Roman Space Telescope*’s grism spectroscopy will provide large cluster samples for such analyses.

Stellar mass can potentially serve as a low-scatter mass proxy, especially with optimally chosen member galaxies (e.g. Golden-Marx & Miller 2018; Anbajagane et al. 2020; Bradshaw et al. 2020; Huang et al. 2022). A cluster sample selected by stellar mass could also provide a useful sanity check for the commonly used richness selection (e.g. Pereira et al. 2018, 2020; Palmese et al. 2020). The stellar mass may have weaker projection effects than richness because it has a large contribution from the brightest cluster galaxies.

### 7.2.4 Defining a cluster sample by a threshold

One way to simplify the modelling of selection bias is to use a single richness threshold instead of multiple richness bins to define our cluster sample. Wu et al. (2021) have shown that the former requires fewer nuisance parameters for the mass–observable relation and can avoid diluting the cosmological information. In addition, the threshold approach does not require the power-law assumption of the mass–observable relation, an assumption that could be too restrictive. Similarly, the threshold approach only requires modelling the selection bias near the richness threshold and can significantly simplify the analyses.

Mitigating selection bias requires us to consider all aspects of the cosmological parameter inference, including the impact of projection on cluster number counts and on the determination of the richness–mass scatter, which is the critical nuisance parameter for analyses focusing on cluster number counts and weak lensing. For example,

in a conventional analysis such as **Y1CL**, which models number counts and  $\Delta\Sigma(r_p)$  in richness and redshift bins based on the halo mass function and a parametrized richness–mass relation, one can incorporate selection bias curves like those in Fig. 2 into the model prediction. It is essential to study a wider range of galaxy HODs and cosmologies to establish the appropriate priors for such corrections. Other approaches bring in cluster autocorrelations, cluster–galaxy cross-correlations, and galaxy autocorrelations as additional constraints (Salcedo et al. 2020; To et al. 2021b), and for these one must examine the impact of selection bias on these additional observables.

## 8 SUMMARY

We investigate the bias of the stacked weak lensing signals around optically selected clusters, using the redMaPPer cluster finder applied to the Buzzard simulations. We find that the large-scale excess surface mass density  $\Delta\Sigma(r_p)$  of richness-selected clusters in Buzzard is 20–60 per cent higher than that expected from the underlying halo mass PDF. Expressed in surface mass density  $\Sigma(r_p)$  rather than  $\Delta\Sigma(r_p)$ , the bias shows strong scale dependence and peaks at  $r_p \approx 1$  pMpc with an amplitude of 10–20 per cent. This scale dependence is well explained by an analytical model that accounts for the correlated residuals between the surface mass density and richness at a given halo mass (equation 8). The correlated residuals arise mainly from projection effects, the boosting of richness and surface mass density by galaxies and matter that lie along the line of sight but outside the halo virial radius. At high richness and small scales, the preferential selection of higher concentration haloes also makes a significant contribution.

We have shown that the complex redMaPPer cluster selection can be modelled by a cylinder member selection. We have found that galaxies within  $\pm 20$  to  $60 h^{-1}$  Mpc along the line of sight but outside the halo virial radius are the main cause of the selection bias. This simplified cylinder selection method can be efficiently applied to a wide range of simulations to study the impact of the galaxy model on selection bias. Our ultimate goal is to mitigate the impact of this bias on cosmological constraints derived from cluster weak lensing surveys.

The selection bias is currently one of the key systematic effects that limit the statistical power of optical cluster cosmology analyses. As discussed above, solving the selection bias would require a concerted effort of simulations, multiwavelength observations, and combined-probe analyses. Currently, DES provides an unprecedented data set for cluster weak lensing, and in the next decade *Euclid*, LSST, and the *Roman Space Telescope* will all provide data sets that are more powerful still. Exploiting the measurements from these data sets is a theoretical challenge, with a potentially critical payoff in unveiling the physics behind cosmic acceleration.

## ACKNOWLEDGEMENTS

H-YW is supported by DOE Grant DE-SC0021916 and NASA grant 15-WFIRST15-0008. MEdSP is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany’s Excellence Strategy – EXC 2121 ‘Quantum Universe’ – 390833306.

This research uses 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. H-YW acknowledges the high-performance computing support of the Borah computer cluster (DOI:

10.18122/oit/3/boisestate) provided by Boise State University's Research Computing Department.

This paper has gone through an internal review by the DES Collaboration.

Funding for the DES Projects has been provided by the U.S. Department of Energy, the U.S. National Science Foundation, the Ministry of Science and Education of Spain, the Science and Technology Facilities Council of the United Kingdom, the Higher Education Funding Council for England, the University of Illinois at Urbana-Champaign - National Center for Supercomputing Applications, the Kavli Institute for Cosmological Physics, University of Chicago, the Center for Cosmology and Astro-Particle Physics at the Ohio State University, the Mitchell Institute for Fundamental Physics and Astronomy at Texas A&M University, Financiadora de Estudos e Projetos, Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro, Conselho Nacional de Desenvolvimento Científico e Tecnológico and the Ministério da Ciência, Tecnologia e Inovação, the Deutsche Forschungsgemeinschaft, and the Collaborating Institutions in the Dark Energy Survey.

The Collaborating Institutions are Argonne National Laboratory, the University of California at Santa Cruz, the University of Cambridge, Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas-Madrid, the University of Chicago, University College London, the DES-Brazil Consortium, the University of Edinburgh, the Eidgenössische Technische Hochschule (ETH) Zürich, Fermi National Accelerator Laboratory, the University of Illinois at Urbana-Champaign, the Institut de Ciències de l'Espai (IEEC/CSIC), the Institut de Física d'Altes Energies, Lawrence Berkeley National Laboratory, the Ludwig-Maximilians Universität München and the associated Excellence Cluster Universe, the University of Michigan, NSF's NOIRLab, the University of Nottingham, The Ohio State University, the University of Pennsylvania, the University of Portsmouth, SLAC National Accelerator Laboratory, Stanford University, the University of Sussex, Texas A&M University, and the OzDES Membership Consortium.

Based in part on observations at Cerro Tololo Inter-American Observatory at NSF's NOIRLab (NOIRLab Prop. ID 2012B-0001; PI: J. Frieman), which is managed by the Association of Universities for Research in Astronomy (AURA) under a cooperative agreement with the National Science Foundation.

The DES data management system is supported by the National Science Foundation under Grant Numbers AST-1138766 and AST-1536171. The DES participants from Spanish institutions are partially supported by MICINN under grants ESP2017-89838, PGC2018-094773, PGC2018-102021, SEV-2016-0588, SEV-2016-0597, and MDM-2015-0509, some of which include ERDF funds from the European Union. IFAE is partially funded by the CERCA program of the Generalitat de Catalunya. Research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Program (FP7/2007-2013) including ERC grant agreements 240672, 291329, and 306478. We acknowledge support from the Brazilian Instituto Nacional de Ciência e Tecnologia (INCT) do e-Universo (CNPq grant 465376/2014-2).

This paper has been authored by Fermi Research Alliance, LLC under Contract No. DE-AC02-07CH11359 with the U.S. Department of Energy, Office of Science, Office of High Energy Physics.

## DATA AVAILABILITY

The data used in this work are available upon request.

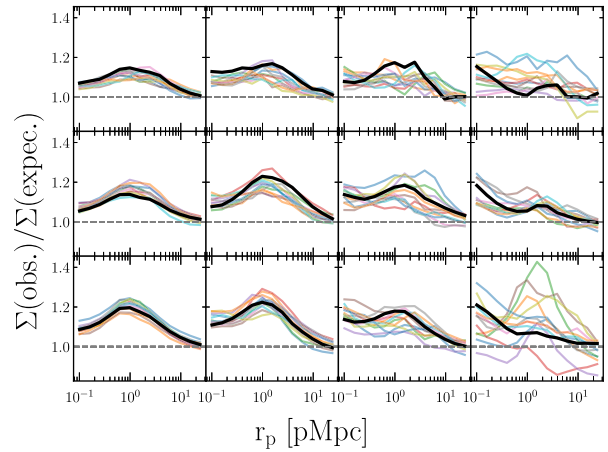
## REFERENCES

- Abbott T. M. C. et al., 2018, *Phys. Rev. D*, 98, 043526  
 Abbott T. M. C. et al., 2020, *Phys. Rev. D*, 102, 023509 (Y1CL)  
 Allen S. W., Evrard A. E., Mantz A. B., 2011, *ARA&A*, 49, 409  
 Anbajagane D., Evrard A. E., Farahi A., Barnes D. J., Dolag K., McCarthy I. G., Nelson D., Pillepich A., 2020, *MNRAS*, 495, 686  
 Bahé Y. M., McCarthy I. G., King L. J., 2012, *MNRAS*, 421, 1073  
 Bartelmann M., Schneider P., 2001, *Phys. Rep.*, 340, 291  
 Behroozi P. S., Wechsler R. H., Wu H.-Y., 2013, *ApJ*, 762, 109  
 Bett P., 2012, *MNRAS*, 420, 3303  
 Bocquet S. et al., 2015, *ApJ*, 799, 214  
 Bocquet S. et al., 2019, *ApJ*, 878, 55  
 Bonamigo M., Despali G., Limousin M., Angulo R., Giocoli C., Soucail G., 2015, *MNRAS*, 449, 3171  
 Bradshaw C., Leauthaud A., Hearin A., Huang S., Behroozi P., 2020, *MNRAS*, 493, 337  
 Bryan G. L., Norman M. L., 1998, *ApJ*, 495, 80  
 Bryan S. E., Kay S. T., Duffy A. R., Schaye J., Dalla Vecchia C., Booth C. M., 2013, *MNRAS*, 429, 3316  
 Chiu I. N., Okumura T., Oguri M., Agrawal A., Umetsu K., Lin Y.-T., 2020, *MNRAS*, 498, 2030  
 Chiu I.-N. et al., 2022, *A&A*, 661, A11  
 Corless V. L., King L. J., 2007, *MNRAS*, 380, 149  
 Costanzi M. et al., 2019a, *MNRAS*, 482, 490  
 Costanzi M. et al., 2019b, *MNRAS*, 488, 4779  
 Costanzi M. et al., 2021, *Phys. Rev. D*, 103, 043522  
 Cromer D., Battaglia N., Miyatake H., Simet M., 2021, preprint (arXiv:2104.06925)  
 Debackere S. N. B., Schaye J., Hoekstra H., 2021, *MNRAS*, 505, 593  
 de Haan T. et al., 2016, *ApJ*, 832, 95  
 DeRose J. et al., 2019, preprint (arXiv:1901.02401)  
 DeRose J. et al., 2022, *Phys. Rev. D*, 105, 123520  
 Dietrich J. P. et al., 2014, *MNRAS*, 443, 1713  
 Drlica-Wagner A. et al., 2018, *ApJS*, 235, 33  
 Eifer T. et al., 2021, *MNRAS*, 507, 1746  
 Erickson B. M. S., Cunha C. E., Evrard A. E., 2011, *Phys. Rev. D*, 84, 103506  
 Evrard A. E., Arnault P., Huterer D., Farahi A., 2014, *MNRAS*, 441, 3562  
 Farahi A., Evrard A. E., Rozo E., Rykoff E. S., Wechsler R. H., 2016, *MNRAS*, 460, 3900  
 Farahi A. et al., 2019, *MNRAS*, 490, 3341  
 Farahi A., Anbajagane D., Evrard A. E., 2022, *ApJ*, 931, 166  
 Frieman J. A., Turner M. S., Huterer D., 2008, *ARA&A*, 46, 385  
 Garrison L. H., Eisenstein D. J., Ferrer D., Tinker J. L., Pinto P. A., Weinberg D. H., 2018, *ApJS*, 236, 43  
 Gavazzi R., 2005, *A&A*, 443, 793  
 Golden-Marx J. B., Miller C. J., 2018, *ApJ*, 860, 2  
 Grandis S. et al., 2021a, *MNRAS*, 504, 1253  
 Grandis S., Bocquet S., Mohr J. J., Klein M., Dolag K., 2021b, *MNRAS*, 507, 5671  
 Henson M. A., Barnes D. J., Kay S. T., McCarthy I. G., Schaye J., 2017, *MNRAS*, 465, 3361  
 Herbonnet R. et al., 2022, *MNRAS*, 513, 2178  
 Huang S. et al., 2022, *MNRAS*, in press (arXiv:2109.02646)  
 Huterer D. et al., 2015, *Astropart. Phys.*, 63, 23  
 Jing Y. P., Suto Y., 2002, *ApJ*, 574, 538  
 Johnston D. E. et al., 2007, preprint (arXiv:0709.1159)  
 Kasun S. F., Evrard A. E., 2005, *ApJ*, 629, 781  
 Kilbinger M., 2015, *Rep. Progress Phys.*, 78, 086901  
 Köhlinger F., Hoekstra H., Eriksen M., 2015, *MNRAS*, 453, 3107  
 Lee B. E., Le Brun A. M. C., Haq M. E., Deering N. J., King L. J., Applegate D., McCarthy I. G., 2018, *MNRAS*, 479, 890  
 Limousin M., Morandi A., Sereno M., Meneghetti M., Ettori S., Bartelmann M., Verdugo T., 2013, *Space Sci. Rev.*, 177, 155  
 McClintock T. et al., 2019, *MNRAS*, 482, 1352  
 Mantz A., Allen S. W., Rapetti D., Ebeling H., 2010, *MNRAS*, 406, 1759  
 Mantz A. B., Allen S. W., Morris R. G., Rapetti D. A., Applegate D. E., Kelly P. L., von der Linden A., Schmidt R. W., 2014, *MNRAS*, 440, 2077

Mao Y.-Y., Williamson M., Wechsler R. H., 2015, *ApJ*, 810, 21  
Murata R., Nishimichi T., Takada M., Miyatake H., Shirasaki M., More S., Takahashi R., Osato K., 2018, *ApJ*, 854, 120  
Murata R. et al., 2019, *PASJ*, 71, 107  
Myles J. et al., 2021, *MNRAS*, 505, 33  
Nishimichi T. et al., 2019, *ApJ*, 884, 29  
Oguri M., Takada M., 2011, *Phys. Rev. D*, 83, 023008  
Oguri M., Takada M., Umetsu K., Broadhurst T., 2005, *ApJ*, 632, 841  
Osato K., Nishimichi T., Oguri M., Takada M., Okumura T., 2018, *MNRAS*, 477, 2141  
Palmese A. et al., 2020, *MNRAS*, 493, 4591  
Park Y., Sunayama T., Takada M., Kobayashi Y., Miyatake H., More S., Nishimichi T., Sugiyama S., 2021, preprint ([arXiv:2112.09059](https://arxiv.org/abs/2112.09059))  
Pereira M. E. S. et al., 2018, *MNRAS*, 474, 1361  
Pereira M. E. S. et al., 2020, *MNRAS*, 498, 5450  
Planck Collaboration XXIV, 2016, *A&A*, 594, A24  
Rines K. J., Geller M. J., Diaferio A., Hwang H. S., Sohn J., 2018, *ApJ*, 862, 172  
Roza E., Rykoff E. S., 2014, *ApJ*, 783, 80  
Roza E. et al., 2010, *ApJ*, 708, 645  
Roza E., Rykoff E. S., Becker M., Reddick R. M., Wechsler R. H., 2015, *MNRAS*, 453, 38  
Rykoff E. S. et al., 2012, *ApJ*, 746, 178  
Rykoff E. S. et al., 2014, *ApJ*, 785, 104  
Rykoff E. S. et al., 2016, *ApJS*, 224, 1  
Salcedo A. N., Wibking B. D., Weinberg D. H., Wu H.-Y., Ferrer D., Eisenstein D., Pinto P., 2020, *MNRAS*, 491, 3061  
Saro A. et al., 2015, *MNRAS*, 454, 2305  
Sartoris B. et al., 2016, *MNRAS*, 459, 1764  
Simet M., McClintock T., Mandelbaum R., Roza E., Rykoff E., Sheldon E., Wechsler R. H., 2017, *MNRAS*, 466, 3103  
Sohn J., Geller M. J., Rines K. J., Hwang H. S., Utsumi Y., Diaferio A., 2018, *ApJ*, 856, 172  
Sunayama T. et al., 2020, *MNRAS*, 496, 4468 (S20)  
To C. H. et al., 2021a, *Phys. Rev. Lett.*, 126, 141301  
To C.-H. et al., 2021b, *MNRAS*, 502, 4093  
Umetsu K., 2020, *A&AR*, 28, 7  
Umetsu K. et al., 2014, *ApJ*, 795, 163  
Vikhlinin A. et al., 2009, *ApJ*, 692, 1060  
Wechsler R. H., DeRose J., Busha M. T., Becker M. R., Rykoff E., Evrard A., 2022, *ApJ*, 931, 145  
Weinberg D. H., Mortonson M. J., Eisenstein D. J., Hirata C., Riess A. G., Roza E., 2013, *Phys. Rep.*, 530, 87  
Wetzell V. et al., 2022, *MNRAS*, 514, 4696  
White M., Cohn J. D., Smit R., 2010, *MNRAS*, 408, 1818  
Wojtak R. et al., 2018, *MNRAS*, 481, 324  
Wu H.-Y., Roza E., Wechsler R. H., 2008, *ApJ*, 688, 729  
Wu H.-Y., Hahn O., Wechsler R. H., Mao Y.-Y., Behroozi P. S., 2013a, *ApJ*, 763, 70  
Wu H.-Y., Hahn O., Wechsler R. H., Behroozi P. S., Mao Y.-Y., 2013b, *ApJ*, 767, 23  
Wu H.-Y., Weinberg D. H., Salcedo A. N., Wibking B. D., 2021, *ApJ*, 910, 28  
Yoo J., Seljak U., 2012, *Phys. Rev. D*, 86, 083504  
Zhang Y. et al., 2019, *MNRAS*, 487, 2578  
Zhang Z. et al., 2022, preprint ([arXiv:2202.08211](https://arxiv.org/abs/2202.08211))

## APPENDIX A: COMPARING TWO VERSIONS OF THE BUZZARD SIMULATIONS

As described in Section 3, we use two versions of Buzzard: 12 realizations of the DES Y1 data (1120 deg<sup>2</sup>) based on Buzzard version 1.9.2+1 (presented in DeRose et al. 2019), and one realization of the DES Y3 data (4143 deg<sup>2</sup>) based on Buzzard version 1.9.8 (presented in DeRose et al. 2022). The main differences between the two versions are as follows.



**Figure A1.** Comparison between two versions of Buzzard and different realizations. The colour curves correspond to the 12 DES Y1 realizations, and the black curve corresponds to the single DES Y3 realization. The statistical fluctuation is larger than the difference between the two versions. Therefore, in the main text, we show the area-weighted means and standard deviations of all 13 realizations.

(i) For the subhalo abundance matching, the former uses the DES Y1 luminosity functions, while the latter uses the DES Y3 luminosity functions (both are modified from the SDSS luminosity functions).

(ii) For the environmental proxy for the SED assignment, the former uses a galaxy’s projected distance to the fifth nearest neighbour, while the latter uses a galaxy’s 3D distance to the nearest halo above a given mass.

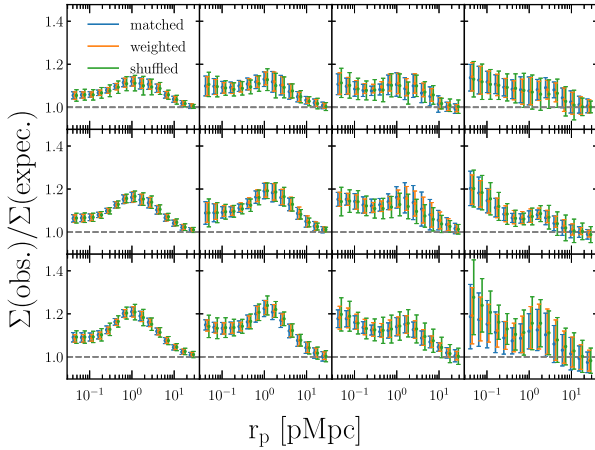
(iii) The former has a narrower red sequence compared with DES data, while the latter explicitly matches the mean and scatter of the red sequence observed in DES Y3.

Fig. A1 compares the lensing bias derived from the two versions and their individual realizations: the colour curves show the 12 DES Y1 realizations, and the black curve shows the one DES Y3 realization. Overall, the statistical fluctuations associated with different realizations are larger than the difference between the two versions. While To et al. (2021b) found that the galaxy clustering of realization ‘3b’ is problematic and removed it from their analysis, we do not find such a discrepancy in cluster lensing. Therefore, unless otherwise noted, throughout this paper we combine all 13 realizations and calculate area-weighted means and standard deviations.

## APPENDIX B: COMPARING DIAGNOSIS METHODS FOR LENSING SELECTION BIAS

Here we compare three diagnosis methods for cluster lensing selection bias, which is quantified by the ratio between (1) the lensing signal from a sample selected in a richness and redshift bin, similar to the DES cluster analyses (we call this the ‘richness-selected’ sample), and (2) the lensing signal we expect from the underlying halo mass PDF of the richness-selected sample. The former is straightforward to compute. Below we present three methods to compute the latter. For all cases, we split each redshift bin  $\Delta z = 0.15$  into three narrower bins  $\Delta z = 0.05$  to account for the redshift dependence.

(i) *Shuffling richness at a given halo mass.* We start with the full halo catalogue and put haloes in narrow mass bins. For haloes in a given mass bin, we shuffle their richness values. This procedure washes out any correlated residuals between lensing and richness at a given mass. Each halo is assigned a new richness  $\lambda^{\text{shuff}}$ . We calculate



**Figure B1.** Comparison between the three diagnosis methods for lensing selection bias detailed in Appendix B. The three methods agree well with each other.

the ratio between the stacked lensing signal of clusters selected with  $\lambda$  and that of clusters selected with  $\lambda^{\text{shuff}}$ . This ratio corresponds to the biased lensing signal due to the correlated residuals between richness and lensing.

(ii) *Matching the underlying mass PDF.* From the full halo catalogue, we select *random* haloes to match exactly the same mass PDF as the richness-selected sample. We construct this random sample to be five times the number of the richness-selected sample, although for massive haloes we need to draw with replacement. We call this the ‘mass-matched’ sample and calculate its mean weak lensing signal.

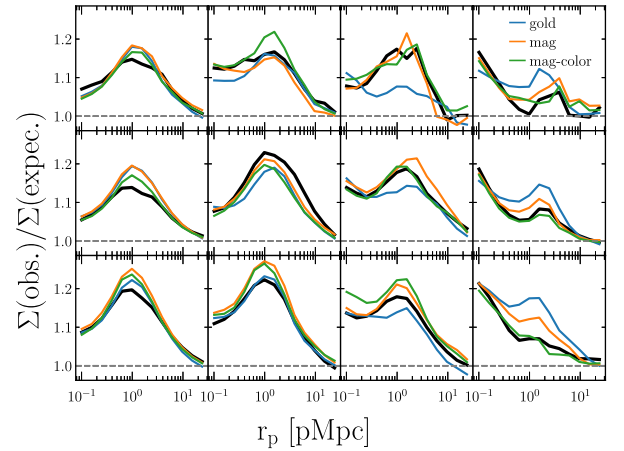
(iii) *Weighting by the underlying mass PDF.* This method is analogous to the matching method, but instead of constructing a random halo sample to match the mass PDF of the richness-selected sample, we use all haloes in the catalogue weighted by this PDF. We use the full halo catalogue and put haloes in narrow mass bins and calculate the mean lensing signal from this bin. Using the PDF of the richness-selected sample, we can calculate the weight associated with each narrow mass bin. We then perform a weighted average of the lensing signal from all mass bins.

Fig. B1 shows that all three methods give consistent results. The shuffling method is the easiest to understand but is also the noisiest because it uses the smallest number of haloes. The weighting method is the least noisy because it averages over the lensing of all haloes in the catalogue. In the main text, we present the results calculated from the weighting method. As an additional sanity check, we have constructed catalogues with no selection bias by assigning to each halo a random richness with a lognormal scatter, and we have recovered unbiased results.

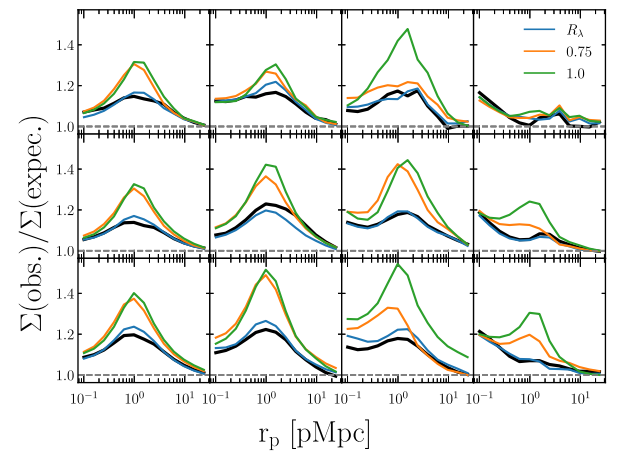
### APPENDIX C: IMPACT OF MAGNITUDE, COLOUR, AND APERTURE ON SELECTION BIAS

In the main text, we present the impact of projection depth on lensing selection bias. In this appendix, we compare different magnitude cuts, colour cuts, and apertures for selecting member galaxies.

Fig. C1 shows the selection bias using galaxies selected with different magnitude and colour criteria. For each cluster, we use a cylinder of radius  $R_\lambda$  and depth  $\pm 30 h^{-1}$  Mpc to define its richness. We compare the redMaPPer results with (1) galaxies in the GOLD



**Figure C1.** Impact of member galaxy magnitude and colour cuts on lensing selection bias. We select members using a cylinder of radius  $R_\lambda$  and length  $\pm 30 h^{-1}$  Mpc and compare the richness calculated with (1) using all galaxies in the GOLD catalogue, (2) setting a magnitude cut at  $0.2L_*$ , and (3) setting an additional colour cut at  $\chi^2_{\text{colour}} < 9$ . While the magnitude and colour cuts significantly reduce the selection bias for high-richness clusters, they have weak impact on low-richness clusters.



**Figure C2.** Impact of member galaxy selection aperture on lensing selection bias. We compare  $R_\lambda$  with fixed, richness-independent apertures: 0.75 and 1 physical  $h^{-1}$  Mpc. A fixed aperture increases the selection bias.

catalogue, which correspond to an *i*-band magnitude limit of  $\approx 26$  and no colour cut, (2) a magnitude cut at  $0.2L_*$  and no colour cut, and (3) our fiducial magnitude and colour cut ( $0.2L_*$  and  $\chi^2_{\text{colour}} < 9$ , see equation 9). Releasing the magnitude and colour cut increases the selection bias for high-richness clusters but has small effects on low-richness clusters.

Fig. C2 shows the selection bias associated with richness defined by different apertures, using our fiducial magnitude–colour selection and cylinder depth  $\pm 30 h^{-1}$  Mpc (comoving). We compare  $R_\lambda = 1(\lambda/100)^{0.2} h^{-1}$  Mpc (physical) with a fixed, richness-independent aperture, 0.75 and 1 physical  $h^{-1}$  Mpc. We find that using a fixed aperture leads to a higher selection bias compared with  $R_\lambda$ . This increased selection bias is related to the increased scatter of mass at a given richness when using a fixed aperture. Although  $R_\lambda$  and fixed apertures give very similar scatter in  $\lambda$  at a given mass, the latter gives a shallower slope of  $\lambda$ –mass relation, which leads to a larger scatter in mass at a given  $\lambda$ . Rykoff et al. (2012) chose the relation

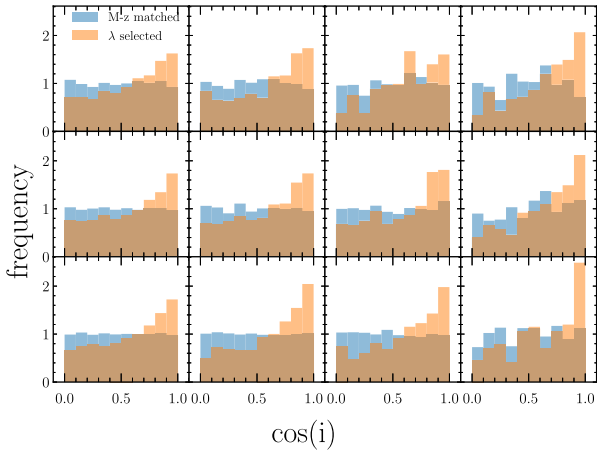
between  $\lambda$  and  $R_\lambda$  to minimize the scatter of  $L_X$  at a fixed  $\lambda$ , and it is encouraging that this radius also gives a weaker selection bias.

## APPENDIX D: SELECTION BIAS ASSOCIATED WITH HALO ORIENTATION

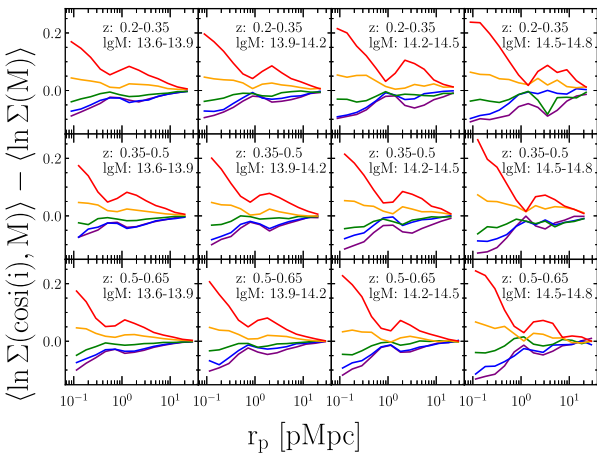
### D1 Measuring the triaxial shape of haloes in simulations

We use a triaxial ellipsoid to describe the 3D shape of a halo, and its orientation is described by the angle between its major axis and the line of sight,  $i$ . For a halo sample with random orientations,  $\cos(i)$  follows a uniform distribution, because the surface area element on a sphere is given by  $d(\cos(i))d\phi$ , where  $\phi$  is the azimuthal angle and runs from 0 to  $2\pi$ .

For haloes in Buzzard, we measure the triaxial shapes and axis orientations following the method described in Osato et al. (2018) with slight modifications. We use all dark matter particles inside  $R_{\text{vir}}$  to iteratively measure the reduced inertia tensor and its eigensystems. We adopt the convention  $a \leq b \leq c$  for an ellipsoid, following Osato



**Figure D1.** Probability distribution of halo orientation  $\cos(i)$  for the richness-selected sample (orange) and the sample constructed to match the underlying mass and redshift distribution (blue). The panels match the redshift and richness bins in Fig. 2. The richness selection preferentially selects haloes with high  $\cos(i)$ .



**Figure D2.** Dependence of  $\Sigma$  on halo orientation. Each panel represents a mass and redshift bin, and the  $\Sigma$  profiles are split into five  $\cos(i)$  bins between 0 and 1.

et al. (2018) and Jing & Suto (2002). Unlike Bett (2012), we do not trim particles in each iteration. We have tested that using slightly different radii, trimming particles in each iteration, or using a non-reduced tensor changes the  $\cos(i)$  by less than 0.1.

The iterative calculation of the halo shape starts with

$$\begin{aligned} q &= 1, \quad s = 1, \\ R_{p,1} &= x, \quad R_{p,2} = y, \quad R_{p,3} = z, \\ R_p^2 &= \left(\frac{R_{p,1}}{q}\right)^2 + \left(\frac{R_{p,2}}{s}\right)^2 + R_{p,3}^2, \\ \mathcal{M}_{ij} &= \frac{1}{N_p} \sum_{p=1}^{N_p} \frac{R_{p,i} R_{p,j}}{R_p^2} \quad \text{for } i, j \in (1, 2, 3), \end{aligned} \quad (\text{D1})$$

where  $(x, y, z)$  are the positions of individual particles,  $N_p$  is the number of particles, and the subscript  $p$  runs through all particles. We calculate the eigenvalues  $(\lambda_1, \lambda_2, \lambda_3)$ , sorted from small to large, and the corresponding eigenvectors  $(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$ ,

$$\begin{aligned} a &= \sqrt{\lambda_1}, \quad \mathbf{v}_{\text{minor}} = \mathbf{v}_1, \\ b &= \sqrt{\lambda_2}, \quad \mathbf{v}_{\text{int}} = \mathbf{v}_2, \\ c &= \sqrt{\lambda_3}, \quad \mathbf{v}_{\text{major}} = \mathbf{v}_3. \end{aligned} \quad (\text{D2})$$

We then update the values:

$$\begin{aligned} q &= a/c, \\ s &= b/c, \\ R_{p,1} &= \mathbf{R}_p \cdot \mathbf{v}_1, \\ R_{p,2} &= \mathbf{R}_p \cdot \mathbf{v}_2, \\ R_{p,3} &= \mathbf{R}_p \cdot \mathbf{v}_3, \\ R_p^2 &= \left(\frac{R_{p,1}}{q}\right)^2 + \left(\frac{R_{p,2}}{s}\right)^2 + R_{p,3}^2. \end{aligned} \quad (\text{D3})$$

These numbers are used for the next iteration.

In each iteration, we transform the coordinate system using the matrix

$$\mathbf{R}_i = \begin{bmatrix} \mathbf{v}_1^T \\ \mathbf{v}_2^T \\ \mathbf{v}_3^T \end{bmatrix}.$$

We multiply the rotation matrix in each step:

$$\mathbf{R}_{\text{final}} = \mathbf{R}_n \cdots \mathbf{R}_1 \mathbf{R}_0, \quad (\text{D4})$$

and the third row of  $\mathbf{R}_{\text{final}}$  is the major axis, denoted as  $\mathbf{v}_3^{\text{final}}$ . The orientation with respect to the line of sight is given by

$$\cos(i) = |\mathbf{v}_3^{\text{final}} \cdot \mathbf{v}^{\text{LOS}}|. \quad (\text{D5})$$

This  $\cos(i)$  is very similar to the result using  $\mathbf{v}_3$  in the initial step. The iteration ends when the fractional changes in both  $q$  and  $s$  are less than  $10^{-7}$ . In Buzzard, the observer is placed at the origin of the  $z = 0$   $N$ -body simulation box, and thus the line-of-sight direction is the same as the position vector.

### D2 Orientation PDF for the Buzzard redMaPPer clusters

One of the possible sources of the selection bias is associated with preferentially selecting haloes with major axes parallel to the line of sight. At a given halo mass, haloes with high  $\cos(i)$  tend to have higher lensing signal than haloes of the same mass. Osato et al. (2018) show that the halo orientation can be associated with the enhanced lensing signal at both small and large scales (100 comoving  $h^{-1}$  Mpc along the line of sight). Using Buzzard simulations, Zhang et al.

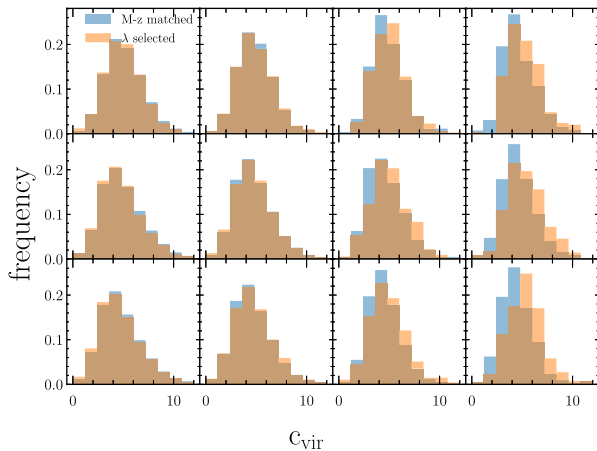
(2022) show that the redMaPPer clusters indeed preferentially select haloes with major axes aligned along the line of sight.

Fig. D1 presents the  $\cos(i)$  distribution of our cluster sample, combining both Y1 and Y3 Buzzard realizations. The orange histograms show the probability distribution of  $\cos(i)$  for our richness-selected sample, and the blue histograms show that for the sample with the same mass and redshift distribution (based on the matching method in Appendix B). The richness-selected sample includes more high- $\cos(i)$  haloes.

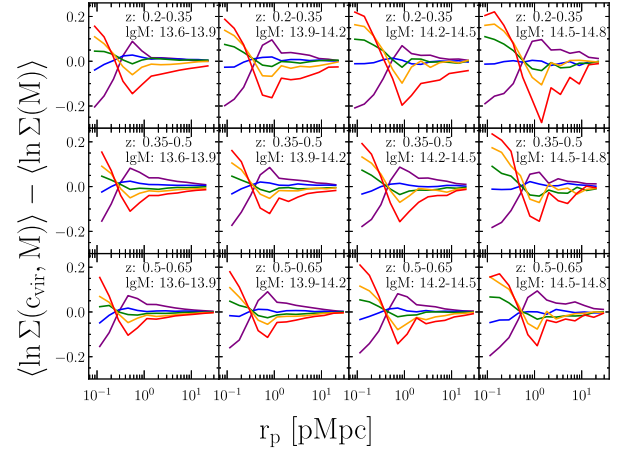
Fig. D2 shows how the  $\Sigma$  profile depends on halo orientation. Each panel corresponds to a redshift and mass bin. For a given mass and redshift, we divide  $\cos(i)$  into five bins of  $\Delta(\cos(i)) = 0.2$  (colour curves). Let us focus on  $0.8 < \cos(i) < 1$  (red), i.e. haloes with their major axes almost perfectly aligned with the line of sight. Their average  $\Sigma$  profile is significantly boosted at small scales, has a dip at  $\approx 1$  pMpc, and has another peak at approximately 2–3 pMpc. The selection bias we have seen in Fig. 2, however, has a different scale dependence. For high-richness clusters, the scale dependence of selection bias is similar to what we see here. However, for low-richness clusters, the selection bias is usually small at small radii and peaks at approximately 1 pMpc. In addition, the 2–3 pMpc peak in Fig. D2 is at 5–10 per cent level, which is lower than the amplitude in Fig. 2. Therefore, while the orientation bias can account for part of the selection bias for the high-richness clusters, it cannot account for the selection bias for low-richness clusters because of the disagreement in the scale dependence and the amplitude. It is possible that other orientation bias proxies (e.g. Dietrich et al. 2014; Herbonnet et al. 2022) could capture the selection bias more fully. For example, we have only calculated halo triaxiality at  $R_{\text{vir}}$ ; it is possible that a halo triaxiality proxy calculated at larger radii or a proxy of large-scale filaments connected to the halo could lead to a better model for selection bias.

## APPENDIX E: LENSING SELECTION BIAS ASSOCIATED WITH HALO CONCENTRATION

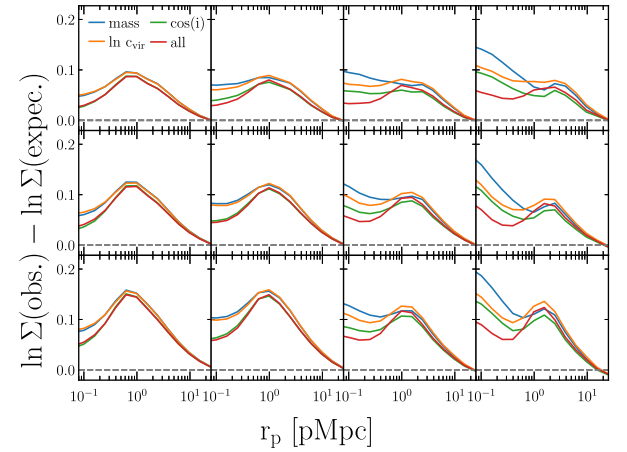
In the main text, we have shown that high-richness clusters exhibit a selection bias in the 3D density profiles (Fig. 3). In this appendix, we investigate the role of halo concentration. Fig. E1 shows the PDF



**Figure E1.** Distribution for halo concentration for the richness-selected sample (orange) and a sample selected to match its mass and redshift distribution (blue). High-richness clusters show a stronger selection bias towards haloes with high concentration, while low-richness clusters do not show such a bias.



**Figure E2.** Dependence of  $\Sigma$  on halo concentration. For a given mass and redshift, we split the haloes into five concentration quintiles. Haloes in the highest concentration quintile have high  $\Sigma$  for at  $r_p \lesssim 0.5$  pMpc, which is compensated by a drop at  $r_p \approx 1$  pMpc.



**Figure E3.** Reduced lensing selection bias by accounting for the mass,  $c_{\text{vir}}$ , and  $\cos(i)$  distribution. The four curves correspond to taking into account mass (blue), mass and concentration (orange), mass and orientation (green), and all three properties (red). The lensing selection bias is not eliminated.

of halo concentration,  $c_{\text{vir}} = R_{\text{vir}}/r_s$ . Here  $r_s$  comes from the fitting of NFW profiles provided by the ROCKSTARhalo finder (Behroozi, Wechsler & Wu 2013). As in Fig. D1, the orange histograms correspond to the richness-selected sample, and the blue histograms correspond to the mass-matched sample. For high-richness clusters (the two right-hand columns), we see a clear preference for high-concentration clusters. Such preference does not exist for low-richness clusters (the two left-hand columns).

Fig. E2 shows the dependence of  $\Sigma$  on halo concentration. For a given mass and redshift bin, we split the haloes into five concentration quintiles. For high concentration haloes, the small-scale behaviour is similar to high- $\cos(i)$  haloes. However, we can see that the scale dependence for high-concentration haloes is very different from that of the selection bias. Therefore, while the concentration could account for some of the high-richness selection bias, it plays a negligible role in the selection bias of low-richness clusters.

Fig. E3 shows the reduction of selection bias when we match not only mass but also  $\cos(i)$  and  $\ln c_{\text{vir}}$  PDF when calculating the ‘expected’ signal. For this calculation, we generalize equation (8)

to include  $\ln c_{\text{vir}}$  and  $\cos(i)$  in the linear regression. The blue curves are our fiducial case that takes into account the mass, while the orange and green curves additionally take into account  $\ln c_{\text{vir}}$  and  $\cos(i)$ , respectively. The red curves take into account all three properties. The selection bias is slightly reduced but is not eliminated. This agrees with our reasoning that the lensing selection bias cannot be fully quantified by the biased selection of  $\cos(i)$  and  $c_{\text{vir}}$ .

<sup>1</sup>*Department of Physics, Boise State University, Boise, ID 83725, USA*

<sup>2</sup>*Astronomy Unit, Department of Physics, University of Trieste, Via Tiepolo 11, I-34131 Trieste, Italy*

<sup>3</sup>*INAF – Osservatorio Astronomico di Trieste, Via G. B. Tiepolo 11, I-34143 Trieste, Italy*

<sup>4</sup>*Institute for Fundamental Physics of the Universe, Via Beirut 2, I-34014 Trieste, Italy*

<sup>5</sup>*Center for Cosmology and Astro-Particle Physics, The Ohio State University, Columbus, OH 43210, USA*

<sup>6</sup>*Department of Physics, The Ohio State University, Columbus, OH 43210, USA*

<sup>7</sup>*Department of Astronomy, The Ohio State University, Columbus, OH 43210, USA*

<sup>8</sup>*Department of Physics, University of Arizona, Tucson, AZ 85721, USA*

<sup>9</sup>*Fermi National Accelerator Laboratory, PO Box 500, Batavia, IL 60510, USA*

<sup>10</sup>*University Observatory, Faculty of Physics, Ludwig-Maximilians-Universität, Scheinerstraße 1, D-81679 Munich, Germany*

<sup>11</sup>*Hamburger Sternwarte, Universität Hamburg, Gojenbergsweg 112, D-21029 Hamburg, Germany*

<sup>12</sup>*Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, CA 94720, USA*

<sup>13</sup>*Department of Physics, University of Michigan, Ann Arbor, MI 48109, USA*

<sup>14</sup>*Departments of Statistics and Data Science, University of Texas at Austin, Austin, TX 78757, USA*

<sup>15</sup>*Kavli Institute for Particle Astrophysics and Cosmology, Stanford University, PO Box 2450, Stanford, CA 94305, USA*

<sup>16</sup>*SLAC National Accelerator Laboratory, Menlo Park, CA 94025, USA*

<sup>17</sup>*Max Planck Institute for Extraterrestrial Physics, Giessenbachstrasse, D-85748 Garching, Germany*

<sup>18</sup>*Universitäts-Sternwarte, Fakultät für Physik, Ludwig-Maximilians Universität München, Scheinerstraße 1, D-81679 München, Germany*

<sup>19</sup>*Department of Physics, Stanford University, 382 Via Pueblo Mall, Stanford, CA 94305, USA*

<sup>20</sup>*Kavli Institute for Cosmological Physics, University of Chicago, Chicago, IL 60637, USA*

This paper has been typeset from a  $\text{\TeX}/\text{\LaTeX}$  file prepared by the author.