# Galaxy correlation function and local density from photometric redshifts using the stochastic order redshift technique (SORT)

James Kakos ⬤,[1]★ Joel R. Primack,[1] Aldo Rodríguez-Puebla ⬤,[2] Nicolas Tejos ⬤,[3] L. Y. Aaron Yung ⬤[4] and Rachel S. Somerville[5]

[1]*Physics Department, University of California, Santa Cruz, CA 95060, USA*
[2]*Instituto de Astronomía, Universidad Nacional Autónoma de México, A.P. 70-264, 04510, CDMX, México*
[3]*Instituto de Física, Pontificia Universidad Católica de Valparaíso, Casilla 4059, Valparaíso, Chile*
[4]*Astrophysics Science Division, NASA Goddard Space Flight Center, 8800 Greenbelt Rd, Greenbelt, MD 20771, USA*
[5]*Center for Computational Astrophysics, Flatiron Institute, 162 5th Ave, New York, NY 10010, USA*

## ABSTRACT

The stochastic order redshift technique (SORT) is a simple, efficient, and robust method to improve cosmological redshift measurements. The method relies upon having a small (∼10 per cent) reference sample of high-quality redshifts. Within pencil-beam-like sub-volumes surrounding each galaxy, we use the precise $dN/dz$ distribution of the reference sample to recover new redshifts and assign them one-to-one to galaxies such that the original rank order of redshifts is preserved. Preserving the rank order is motivated by the fact that random variables drawn from Gaussian probability density functions with different means but equal standard deviations satisfy stochastic ordering. This process is repeated for sub-volumes surrounding each galaxy in the survey. This results in every galaxy being assigned multiple 'recovered' redshifts from which a new redshift estimate is determined. An earlier paper applied SORT to a mock Sloan Digital Sky Survey at $z \lesssim 0.2$ and accurately recovered the two-point correlation function (2PCF) on scales $\gtrsim 4~h^{-1}$Mpc. In this paper, we test the performance of SORT in surveys spanning the redshift range $0.75 < z < 2.25$. We used two mock surveys extracted from the Small MultiDark–Planck and Bolshoi–Planck $N$-body simulations with dark matter haloes that were populated by the Santa Cruz semi-analytic model. We find that SORT overall improves redshift estimates, accurately recovers the redshift-space 2PCF $\xi(s)$ on scales $\gtrsim 2.5~h^{-1}$Mpc, and provides improved local density estimates in regions of average or higher density, which may allow for improved understanding of how galaxy properties relate to their environments.

**Key words:** methods: data analysis – methods: statistical – techniques: photometric – techniques: spectroscopic – large-scale structure of Universe.

## 1 INTRODUCTION

In modern cosmology, the large-scale distribution of galaxies arises from the gravitational evolution and hierarchical clustering of primordial fluctuations. Large ΛCDM $N$-body simulations of cold dark matter and dark energy predict how these structures evolve. Such simulations show that many properties of dark matter haloes are correlated with the local density of the regions in which they form on scales of a few Mpc (e.g. Lee et al. 2017). But baryonic physics is complex, and we are still seeking to understand how galaxies form and evolve and how that is connected with the properties of their host dark matter haloes and the environments in which they reside (e.g. Somerville & Davé 2015; Wechsler & Tinker 2018, and references therein). This can perhaps be clarified by comparing how halo properties and galaxy properties, such as stellar radius, depend on local density and small-scale clustering (e.g. Behroozi, Hearin & Moster 2022).

We anticipate that this effort will be tremendously aided by the immense quantity of data that will flow from the giant LSST imaging survey at the Vera Rubin Observatory (Ivezić et al. 2019) and the higher resolution imaging surveys by the *Euclid Space Telescope*[1] and the *Roman Space Telescope* (Spergel et al. 2015). These surveys will provide multiwaveband photometric redshifts for billions of galaxies, of accuracy $\sigma_z/(1 + z) \approx 0.02$ or better. *Euclid* will also measure grism redshifts of accuracy $\sigma_z/(1 + z) \approx 10^{-3}$ for ∼30 million galaxies (Scaramella et al. 2021). Meanwhile, the Dark Energy Spectroscopic Instrument (DESI Collaboration et al. 2016) will measure redshifts of accuracy $\sigma_z/(1 + z) \approx 10^{-4}$ for ∼10 million QSOs and ∼20 million galaxies, including ∼17 million emission-line galaxies in the redshift interval $0.6 < z < 1.6$. In the same regions of the sky where these accurate spectroscopic redshifts are being measured, the imaging surveys will produce more than an order of magnitude more photometric redshifts. It is therefore very important to develop methods that can make efficient use of the combination of a small fraction of spectroscopic redshifts and a much larger fraction of photometric or grism redshifts in order to measure the local environments and correlations of distant galaxies. The present paper discusses one such method.

★ E-mail: jkakos@ucsc.edu

[1]https://sci.esa.int/web/euclid

The basic idea behind these methods is that galaxies cluster, especially on scales of a few Mpc. The idea of estimating redshifts using clustering was first developed by Seldner & Peebles (1979), Phillipps & Shanks (1987), and Landy, Szalay & Koo (1996). More recently, Ménard et al. (2013) proposed a method using a small set of reference galaxies with spectroscopic redshifts to estimate redshifts for a larger set of galaxies that are nearby on the sky to the reference galaxies. This was tested with simulations by Schmidt et al. (2013), compared with spectroscopic redshifts by Rahman et al. (2015), used to reconstruct redshift distributions from measurement of the angular clustering of galaxies using a subset of spectroscopic redshifts by Scottez et al. (2016), and tested with simulations by Scottez et al. (2018). A related method was proposed by Morrison et al. (2017). See also Gatti et al. (2022), Cawthon et al. (2020), and Hildebrandt et al. (2021) for more recent applications of clustering redshifts.

A method to estimate redshifts of galaxies with photometric redshifts using proximity to the cosmic web defined by a subset of galaxies with spectroscopic redshifts was proposed by Aragon-Calvo et al. (2015), who applied this PhotoWeb method to the SDSS out to redshift $z \approx 0.12$. Shuntov et al. (2020) applied this method to a larger sample of galaxies with spectroscopic redshifts from the SDSS and BOSS surveys out to redshift $z = 0.4$ to reconstruct the cosmic web using the DisPerSE algorithm (Sousbie 2011), and they used a convolutional neural network (CNN) trained with the SDSS and GAMA surveys to obtain photometric redshifts with mean absolute deviation $\sigma_{MAD} \approx 0.01$ out to redshift $z \approx 0.3$ for bright galaxies with $r < 17.8$. They claimed that their version of the PhotoWeb method improved the accuracy of the redshifts by about a factor of 2, to $\sigma/(1 + z) \approx 0.004$.

The stochastic order redshift technique (SORT; Tejos, Rodríguez-Puebla & Primack 2018) is complementary to these approaches. It considers a patch on the sky where initially two kinds of galaxy redshift measurements exist: less accurate (e.g. photometric) and precise (spectroscopic). The galaxies with precise redshifts are used as a 'reference sample,' and it is of course expected that these correspond to a small fraction of the total number of galaxies. New 'recovered' redshifts are sampled from the distribution of precise redshifts and matched one-to-one with the uncertain redshifts such that the rank order of the uncertain redshifts is preserved. This step is motivated by the fact that random variables drawn from Gaussian probability density functions (PDFs) with equal, arbitrarily-large standard deviations satisfy stochastic ordering. By construction, SORT is non-parametric as it does not need to assume any functional form for either the distribution of redshifts or the relationship between spectroscopic and photometric redshifts. Thus, the power of SORT relies on its simplicity and versatility.

In this paper, we test how well the SORT method can use photometric redshifts plus a smaller set of reference galaxies with spectroscopic redshifts to estimate the correlations of galaxies and the local densities of their environments out to high redshifts (here we focus on a redshift interval from 0.5 to 2.5). We test the SORT method using mock galaxy surveys from backward light cones extracted from the Small MultiDark–Planck and Bolshoi–Planck cosmological ΛCDM simulations (Klypin et al. 2016; Rodríguez-Puebla et al. 2016). The dark matter haloes were populated with central and satellite galaxies using a current version of the Santa Cruz semi-analytic model (SAM), which has been shown to reproduce well the properties of observed galaxies out to high redshifts (Somerville et al. 2021, and references therein). We show that SORT is indeed robust and that it can provide unbiased measurement of the redshift-space two-point correlation function (2PCF) on scales $\gtrsim 2.5\,h^{-1}$Mpc

while also recovering the local galaxy and mass density, especially in regions of higher than average density where most galaxies reside.

This paper is organized as follows. In Section 2, we briefly describe the method, while in Section 3 we describe the mock galaxy surveys used to study its performance. In Section 4, we present the results of applying SORT to a 2 deg$^2$, mock wide-field light cone including galaxy redshifts, 2PCFs, and inferred three-dimensional densities of galaxy neighbourhoods. In Section 5, we provide a discussion regarding preservation of the redshift rank order, the effects of the SORT parameters, limitations of the method, and potential future improvements. Section 6 presents a summary and main conclusions. Appendix A provides the results of applying SORT to a mock CANDELS light cone of area 0.2 deg$^2$. Appendix B describes our method for assigning 3D coordinates to satellite galaxies in the Santa Cruz SAM. Appendix C describes SORT performance with larger photometric redshift uncertainties. Appendix D describes the method we used to make vectorized figures with many points that are nevertheless of a small file size. Appendix E provides several supplementary figures. All reported distances hereafter are comoving unless specified otherwise.

## 2 THE SORT METHOD

Here we present a brief overview of the SORT method. For a more complete discussion with illustrative figures, we refer the reader to Tejos et al. (2018).

### 2.1 General idea

Consider a set of $N$ galaxies comprised of a mixture of low-quality (referred to as photometric) and high-quality (referred to as spectroscopic) redshifts within a volume $V$. Assume that there are $N_{ph}$ galaxies with photometric redshifts and $N_{sp}$ galaxies with spectroscopic redshifts. When observing galaxies along some pencil-beam-like sub-volume, each of these subsets of galaxies will have a redshift probability distribution, $P_{ph}$ or $P_{sp}$, dependent upon their respective redshift uncertainties. In principle, both $P_{ph}$ and $P_{sp}$ can be considered representations of the same underlying true probability distribution with different levels of noise. Due to the greater expense of obtaining spectroscopic redshifts than photometric redshifts, the statistics for $P_{sp}$ are comparatively limited. However, if $N_{sp}$ is large enough to be statistically relevant to the total set of $N$ galaxies – i.e. accurately traces the cosmic structure within the volume – the higher quality of the spectroscopic redshifts will provide us with a higher resolution look at the true galaxy distribution. In this way, $P_{ph}$ can be considered a noisier version of $P_{sp}$.

We can leverage the relationship between $P_{ph}$ and $P_{sp}$ to try to improve the estimates of the photometric redshifts. To do this, we rely on stochastic ordering, which is defined as follows. Given two PDFs $P_A(x)$ and $P_B(x)$, the variable $X_A$ is stochastically less than $X_B$ if

$$P_A(X_A > x) \leq P_B(X_B > x) \quad \forall x. \qquad (1)$$

To relate this to redshift estimates, consider two observed photometric redshifts $z_i^{ph}$ and $z_j^{ph}$ where $z_i^{ph} < z_j^{ph}$. We can think of each of these as being random variables sampled from Gaussian[2] PDFs centred on $z_i^{true}$ and $z_j^{true}$, respectively, with equal standard

---

[2]The redshift PDFs are not required to be Gaussian, but this is used for simplicity.

deviations determined by the measurement uncertainties.[3] Even with potentially overlapping PDFs, $z_i^{\mathrm{ph}}$ and $z_j^{\mathrm{ph}}$ will satisfy stochastic ordering. Therefore, we can say the most likely scenario is that the underlying true redshifts satisfy $z_i^{\mathrm{true}} \leq z_j^{\mathrm{true}}$. By extension, if we have $N_{\mathrm{ph}}$ redshift estimates ordered such that $z_1^{\mathrm{ph}} \leq z_2^{\mathrm{ph}} \leq \ldots \leq z_{N_{\mathrm{ph}}}^{\mathrm{ph}}$, we would also expect the true redshifts to most likely have the same rank ordering such that $z_1^{\mathrm{true}} \leq z_2^{\mathrm{true}} \leq \ldots \leq z_{N_{\mathrm{ph}}}^{\mathrm{true}}$.

Of course, we do not have the 'true' redshifts for galaxies, so we rely on high-quality spectroscopic redshifts instead. To apply the idea, we search in pencil-beam-like sub-volumes to determine $P_{\mathrm{sp}}$ in that sub-volume. We then randomly sample $N_{\mathrm{ph}}$ 'recovered' redshifts, $z_i^{\mathrm{rec}}$, from $P_{\mathrm{sp}}$. Both the photometric redshifts and the recovered redshifts are rank ordered and matched one-to-one such that $z_i^{\mathrm{rec}} \leftrightarrow z_i^{\mathrm{ph}}$ for all $N_{\mathrm{ph}}$ redshifts. In doing this, we take advantage of the higher resolution provided by $P_{\mathrm{sp}}$ and simultaneously preserve the rank ordering.

We note that there may be cases where equation (1) does not hold true, e.g. in situations where there is high variance among the individual photometric PDF widths and/or shapes. However, we can expect that for state-of-the-art photometric redshift uncertainties, the PDFs will be well behaved and obey stochastic ordering for the majority of cases. We also note that SORT is a statistical model that should only be applied to sets of galaxies rather than individual measurements. Overall, SORT can improve redshift estimates of a set, but it can also make individual measurements worse than the original photometric estimates. Indeed, in some cases, SORT may return individual measurements that are inconsistent with a galaxy's original PDF (i.e. redshifts with errors larger than three times the photometric uncertainty).

### 2.2 The SORT algorithm

For each galaxy $i$ in the sample with photometric redshifts, the following steps are taken (see also fig. 1 from Tejos et al. 2018):

(i) A circle with radius $R$ is drawn on the sky around the $i$th galaxy.

(ii) Galaxies that fall within a cylinder defined by the radius $R$ and a redshift range $z_i \pm \Delta_z$ are selected and used for the remaining steps.

(iii) From the selected galaxies, a check is made to ensure there are at least $N_{\mathrm{ref}}^{\mathrm{min}}$ galaxies with spectroscopic redshifts. If there are not at least $N_{\mathrm{ref}}^{\mathrm{min}}$ spectroscopic redshifts, the circle radius is incremented by $\delta R$ until the criterion is met or $R$ exceeds some $R_{\mathrm{max}}$. If $R$ exceeds $R_{\mathrm{max}}$, SORT is considered to have failed and does not return any redshifts. The algorithm then moves to the next galaxy.

(iv) A redshift histogram of the spectroscopic galaxies is made using a binning of $\mathrm{d}z/3$. The histogram is then convolved with a Gaussian with $\sigma = \mathrm{d}z$.[4]

(v) For each of the photometric galaxies within the cylinder, a new recovered redshift is sampled from the histogram of spectroscopic redshifts.

(vi) The selected galaxies' redshifts and the recovered redshifts are each rank ordered and matched one-to-one so each photometric galaxy is assigned a recovered redshift. This is the key 'sorting' step of SORT.

As this procedure is carried out for the remaining galaxies, every time a given galaxy is within the cylinder of one of its neighbours, it will gain another recovered redshift based on that selection. In the data presented, there was a median of 43 recovered redshifts for any given galaxy. After the algorithm completes, each galaxy is assigned the median of all its recovered redshifts as its sorted redshift, $z_{\mathrm{sort}}$. The values used for the algorithm parameters are discussed in Section 4.

## 3 MOCK GALAXY SURVEYS

### 3.1 Simulations and backward light cones

We use mock galaxy surveys constructed by extracting dark matter haloes along backwards light cones from the Small MultiDark–Planck and Bolshoi–Planck $N$-body simulations (Klypin et al. 2016; Rodríguez-Puebla et al. 2016). A brief summary of the light cones and their respective simulations is shown in Table 1. The dark matter haloes in the simulations were identified using ROCKSTAR (Behroozi, Wechsler & Wu 2013). The backward light cones were constructed using the LIGHTCONE package[5] released by Behroozi et al. (2019), and further details are described in Somerville et al. (2021) and Yang et al. (2021).

The merger histories of the dark matter haloes were constructed using an algorithm based on the extended Press–Schechter formalism (Somerville & Kolatt 1999; Somerville et al. 2008). The formation and evolution of galaxies within these haloes was then modeled using the Santa Cruz SAM (Somerville & Primack 1999; Somerville et al. 2008; Somerville, Popping & Trager 2015). Somerville et al. (2021) presented a suite of light cones that was designed to represent the geometry and approximate areas of the five fields from the Cosmic Assembly Near-infrared Deep Extragalactic Legacy Survey (CANDELS).[6] They compared the mock survey predictions with the CANDELS observed counts, stellar mass functions, rest-frame luminosity functions from $0.1 \lesssim z \lesssim 2$, and found generally good agreement. Yung et al. (in preparation) present a suite of 2 deg$^2$ mock light cones that have been populated with galaxies using the same approach (Yung et al. 2019a,b). In this work, we make use of one of the 2 deg$^2$ mock light cones and one of the mock CANDELS catalogues with field geometry similar to the COSMOS field, covering an area on the sky of $17 \times 41$ arcmin$^2$ in Right Ascension and Declination.

The Santa Cruz SAM does not make use of the $N$-body positions and velocities for dark matter haloes once they become 'sub-haloes' (or satellites) within a larger halo. Instead, it estimates the galactocentric radius of each satellite from the centre of the halo and its decay due to dynamical friction using an analytic model (see Somerville et al. 2008, 2021). As a result, in order to compute separate redshifts for the satellites, the 3D positions and velocities must be assigned in post-processing. For details on our method for assigning these properties to the satellite galaxies, see Appendix B.

The full catalogues span the range $0 < z < 10$, though this work uses only galaxies in the range $0.75 < z < 2.25$ based on their mock observed redshifts. The lower redshift limit was imposed to ensure the light cones had large enough cross-sectional areas to measure the 2PCF on scales of $\gtrsim 3~h^{-1}$Mpc. Mock observed redshifts were calculated using

$$z_{\mathrm{obs}} = z_{\mathrm{los}} + \delta_z(1 + z_{\mathrm{los}}), \tag{2}$$

---

[3] Alternatively, one can think of the Gaussian PDFs as centred on $z^{\mathrm{ph}}$ where the PDF corresponds to the probability of finding $z^{\mathrm{true}}$ at a given $z$.

[4] The motivation for this step is to have a smooth version of the discrete $\mathrm{d}N/\mathrm{d}z$ distribution associated with the reference sample.

[5] https://bitbucket.org/pbehroozi/universemachine/src/master/.

[6] http://arcoiris.ucolick.org/candels/.

**Table 1.** Comparison of the two mock light cones used.

| Light cone | Size (deg$^2$) | Galaxies | Completeness | Simulation | $\Omega_{\Lambda, 0}$ | $\Omega_{M, 0}$ | $\Omega_{B, 0}$ | $h$ | $n_s$ | $\sigma_8$ |
|---|---|---|---|---|---|---|---|---|---|---|
| Wide field | 2 | 1058 366 | $H < 27$ | Small MultiDark–Planck | 0.693 | 0.307 | 0.048 | 0.678 | 0.96 | 0.829 |
| CANDELS | 0.2 | 47 404 | $H < 25.5$ | Bolshoi–Planck | 0.693 | 0.307 | 0.048 | 0.678 | 0.96 | 0.823 |

*Notes.* Each light cone was restricted to the redshift range $0.75 < z < 2.25$. Galaxies were selected from three complete redshift bins (as shown in Fig. 1). The light cones were extracted from different simulations, though the cosmological parameters are the same for both with the exception of $\sigma_8$.
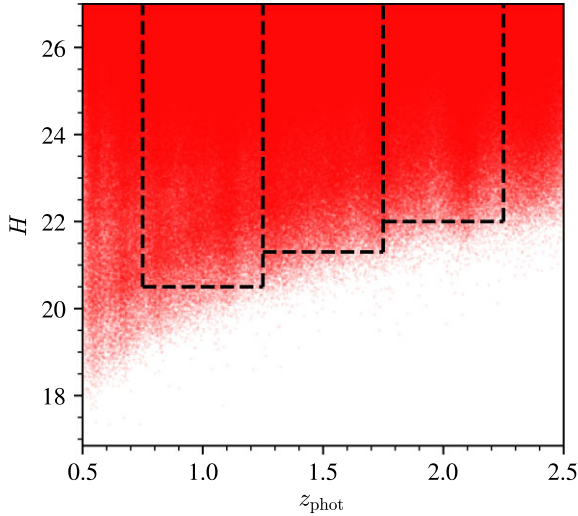


**Figure 1.** Galaxies were selected from the three volume-complete regions defined by the dashed lines. Spectroscopic redshifts were assigned within each region randomly to 10 per cent of the galaxies in that region. See Appendix D for a note on how this figure (and others in this paper) were made with a combination of vectorized and rasterized elements that maintain clarity while reducing file sizes.

where $z_{\mathrm{los}}$ is the redshift that includes distortions from peculiar velocities along the line of sight and $\delta_z$ is a random sample from a Gaussian centred at zero with standard deviation $\sigma_z$ (either photometric or spectroscopic). Note that we do not model catastrophic failures in the photometric sample as we do not expect them to have a significant effect on the net result of the SORT method.

Apparent magnitudes are provided in the mock light cones. For this work, we use $H$-band magnitudes given by the 'wfc3f160w_dust' output of the Santa Cruz SAM. The observed-frame IR luminosities are calculated based on the star formation histories predicted by the Santa Cruz SAM and stellar population synthesis models of Bruzual & Charlot (2003). Dust attenuation is modeled using a standard 'slab' model as described in Somerville et al. (2012). For more details, we refer the reader to Somerville et al. (2021).

We adopt a completeness of $H < 25.5$ for the mock CANDELS light cone, which is a rough limit to which we expect CANDELS photometric redshifts to be accurate. We expect future surveys to improve this and thus adopt $H < 27$ for the wide-field light cone. Galaxies were selected from three volume-complete regions within the light cone, as shown in Fig. 1. In each of the three regions, 10 per cent of galaxies were randomly chosen to have mock spectroscopic redshifts while the remaining 90 per cent were given mock photometric redshifts. After preparing the mock catalogues, the 2 deg$^2$ wide-field light cone had 1058 366 galaxies ($\sim$147 galaxies per arcmin$^2$) and the mock CANDELS light cone had 47 404 galaxies ($\sim$68 galaxies per arcmin$^2$). All results in the main text of this paper are drawn from the wide-field light cone as this provides better overall statistics. Parallel results for the mock CANDELS light cone

are shown in Appendix A and are more representative of present galaxy surveys.

### 3.2 Redshift types

Here we define several different redshift types that will be discussed in our results:

(i) $z_{\mathrm{cos}}$ : These are redshifts that are purely cosmological and include neither redshift-space distortions from line-of-sight peculiar velocities nor measurement uncertainty.

(ii) $z_{\mathrm{spec}}$ : These are simulated spectroscopic redshifts that include a small measurement uncertainty according to equation (2).

(iii) $z_{\mathrm{ref}}$ : These are the reference sample redshifts. They comprise a relatively small fraction of the total number of redshifts and have spectroscopic quality.

(iv) $z_{\mathrm{phot}}$ : These are simulated photometric redshifts. They are generated the same way as $z_{\mathrm{spec}}$ but with larger uncertainties.

(v) $z_{\mathrm{sort}}$ : These are the results of running the SORT method.

(vi) $z_{\mathrm{ctrl}}$ : These are the results of the controlled SORT algorithm that excludes rank ordering (see Section 5.1 for details).

### 4 RESULTS

Most of the results in this paper were obtained assuming a spectroscopic redshift fraction of 10 per cent, although we also explored larger and smaller spectroscopic fractions (see Fig. E2). The spectroscopic and photometric uncertainties used were $\sigma_z^{\mathrm{sp}}/(1 + z) = 0.0001$ and $\sigma_z^{\mathrm{ph}}/(1 + z) = 0.01$, respectively, but we also provide results for $\sigma_z^{\mathrm{ph}}/(1 + z) = 0.02$ in Fig. 7 and Appendix C. The minimum required number of reference galaxies for each sub-volume was set to $N_{\mathrm{ref}}^{\mathrm{min}} = 4$. We found SORT to be effective with this value as low as $N_{\mathrm{ref}}^{\mathrm{min}} = 2$, but increasing to 4 provided a better overall estimate of the 2PCF while other results remained similar. The initial search radius was set to $R = 0°.01$ and the redshift bin width was set to $\mathrm{d}z = 0.0003$. These correspond to length-scales of around 0.3–0.7 $h^{-1}$Mpc for $R$ and 0.3–0.6 $h^{-1}$Mpc for $\mathrm{d}z$ in the range $0.75 < z < 2.25$. These values were chosen to be able to capture relevant scales of the cosmic web. The search radius increment was set to $\delta R = 0.1R$ with a maximum possible radius of $R_{\mathrm{max}} = 0.1°$. If the $N_{\mathrm{ref}}^{\mathrm{min}}$ criterion was not met within $R \leq R_{\mathrm{max}}$ for a given galaxy, that galaxy was removed from the results.[7] The search depth was limited to $z_i \pm \Delta_z$ with $\Delta_z = 2.5\sigma_z^{\mathrm{ph}}$. This depth was chosen to be large enough to capture nearly all photometric redshifts and their true environments within the same sub-volume. See Section 5.2 for details on these parameters.

Our primary comparison for the results of SORT is to $z_{\mathrm{spec}}$, as spectroscopic redshifts represent our best estimates of galaxy redshifts and SORT uses these to trace the cosmic web. However, in some cases, we also show results of $z_{\mathrm{cos}}$ despite these redshifts not being directly observable due to redshift-space distortions. These
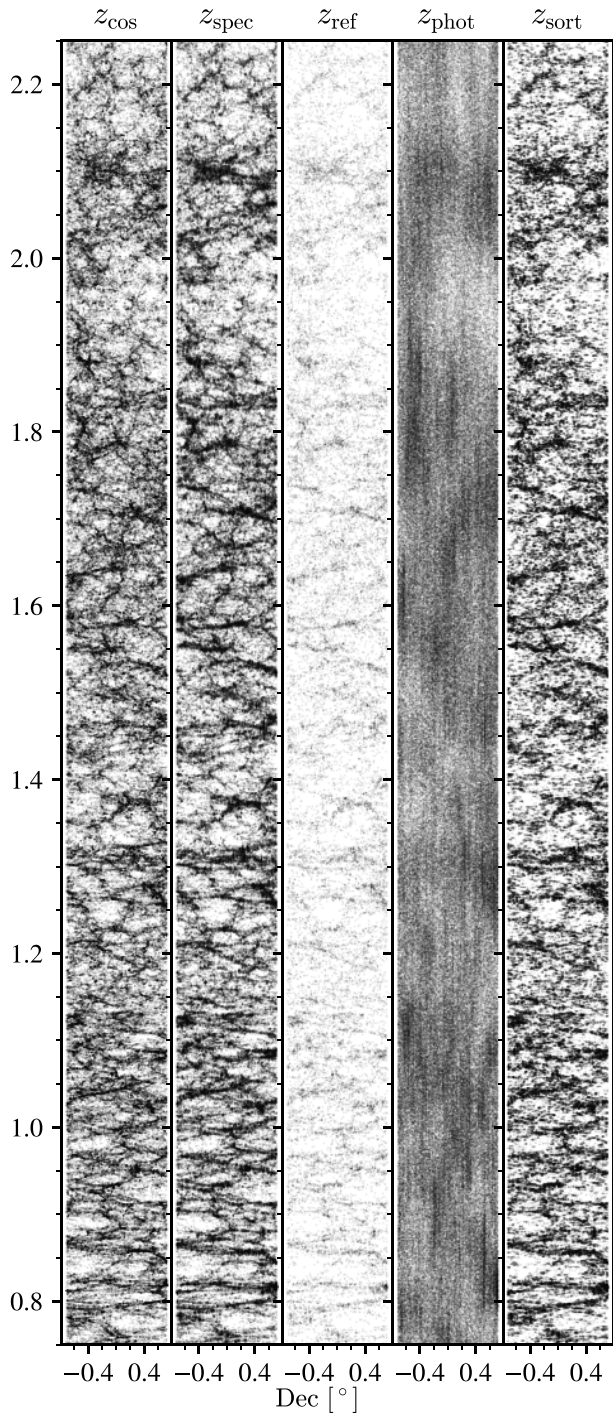
---

[7]In the data presented, no such galaxies were removed.

**Figure 2.** Scatter plot of the projected two-dimensional distribution of galaxies for $z_{cos}$, $z_{spec}$, $z_{ref}$, $z_{phot}$, and $z_{sort}$. Each panel shows a 0°5 slice in Right Ascension and the full Declination of the light cone. The middle panel, $z_{ref}$, corresponds to 10 per cent of the total galaxies, and the remaining panels show the 90 per cent non-reference galaxies. The large-scale features of the cosmic web are much more identifiable with $z_{sort}$ than $z_{phot}$. However, SORT's tendency to group galaxies closely together means that it struggles to recover low-density regions. Note that the horizontal cuts slightly visible in the $z_{phot}$ panel are a result of the completeness condition shown in Fig. 1.

results are shown for comparison as they represent the true underlying distribution of galaxies. Figs 3 and E4 show $z_{cos}$ to illustrate the effects of redshift-space distortions and the alignment of reference galaxies with the true cosmic web. Density estimates of $z_{sort}$ are

compared to those of $z_{cos}$ in Fig. 8 as Lee et al. (2017) showed that many halo properties correlate with local densities using the true $N$-body positions of haloes (which are replicated by using $z_{cos}$, not $z_{spec}$).

### 4.1 Improving redshift estimates

A general look at how well SORT is able to improve redshift estimates can be seen in Fig. 2. Each panel shows a different redshift type plotted against the full Declination of the light cone. The middle panel shows the reference sample, $z_{ref}$, which is comprised of 10 per cent of the spectroscopic sample and is assumed to be known when SORT is applied. This is the structural outline that SORT uses to reassign redshifts. In the $z_{phot}$ panel, the cosmic structure is almost entirely smoothed out. Even with an optimistic photometric uncertainty of $\sigma_z^{ph}/(1 + z) = 0.01$, one can only get a very rough sense of high- or low-density regions. The $z_{sort}$ panel shows a significant improvement on $z_{phot}$. We see more accurate clustering of galaxies, as well as signs of filamentary structure and voids. We note that SORT's reconstruction of low-density regions is not particularly good. This is primarily due to SORT's tendency to place galaxies near other galaxies. Lower density regions will be populated with few galaxies, and only a small fraction of those will be reference galaxies.

A more zoomed-in view of the different redshift types can be seen in Fig. 3. Each panel shows a square region of space, roughly $75 \times 75\,h^{-1}$Mpc. The red and blue colouring represents the direction of the peculiar velocities along the line of sight; red points have positive velocities and blue points have negative velocities. When comparing $z_{cos}$ to $z_{spec}$, we see that galaxies in denser regions become spread out vertically. The severity of these distortions will directly impact SORT's ability to reconstruct the cosmic web. Redshift-space distortions in the reference sample will inherently affect how SORT assigns redshifts. For example, there is a dense cluster of galaxies in Fig. 3 in front of a void. The redshift-space distortions cause a number of galaxies, including some reference galaxies, to be shifted into the void. This results in SORT placing galaxies in the void where they otherwise should not be placed.

We notice also how SORT clusters galaxies tightly to the reference sample. In the lowest density environments, there are cases where galaxies build up around one or two reference galaxies – e.g. around (0.5, 1.46) in the $z_{sort}$ panel. Galaxies are pulled along the line of sight to a nearby reference galaxy, leading to horizontal structures in a plane perpendicular to the line of sight. This is most prominent in low-density environments because SORT has to increase its search radius to find reference galaxies. This allows galaxies at a wider range of angular separations to be placed at roughly the same redshift. Overall, though, we see that SORT does a fairly good job at reconstructing the main features in this region of space, especially compared to the photometric redshifts. In the $z_{phot}$ panel on the lower left-hand panel, any sign of the main features in this region is almost completely lost.

This is further shown when looking at the recovery of the spectroscopic $dN/dz$ distribution. The one-dimensional redshift distributions are shown in Fig. 4 for $z_{spec}$ (grey region), $z_{sort}$ (red), and $z_{phot}$ (blue). The peaks and valleys are smoothed out in the photometric distribution while $z_{sort}$ shows significant improvement in ability to outline large-scale structure along the line of sight. By construction, $z_{sort}$ is meant to follow the $dN/dz$ distribution of the spectroscopic reference sample, and that is what we observe here. Some of the discrepancy is a result of SORT pulling galaxies from low-density regions, where reference galaxies are scarce, to high-density regions.

Fig. 5 shows the error $\Delta z/(1 + z)$ with respect to $z_{spec}$. The grey-shaded region shows the error distribution for $z_{phot}$. Since the
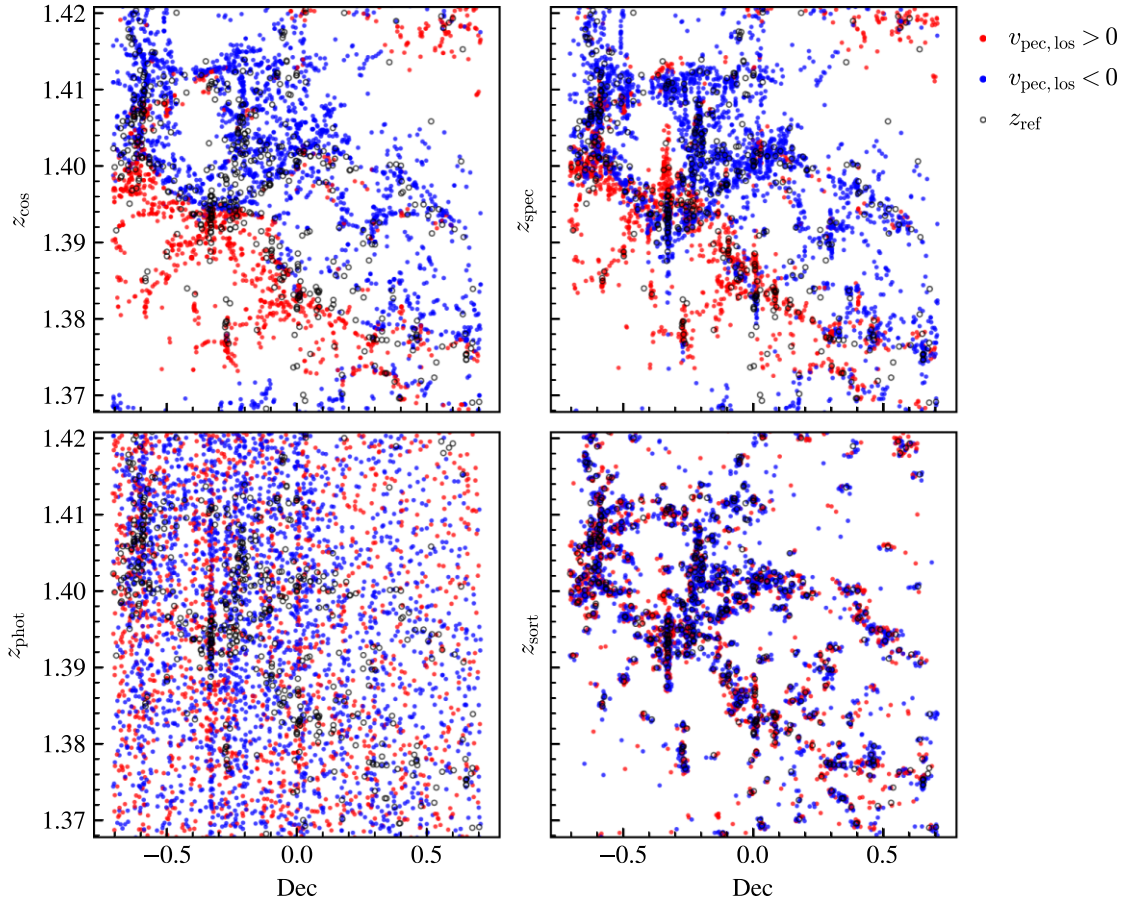
**Figure 3.** Right Ascension slices (thickness $0°1$) of galaxy distributions using different redshifts in a roughly $75 \times 75\ h^{-1}$Mpc region of space. The red and blue colouring denotes the direction of the peculiar velocity along the line of sight (red is positive and blue is negative). The black rings with empty centres are reference galaxies. This region is dominated by a high-density ring of galaxies that surrounds a void in the upper left-hand quadrant. We can see that the accurate tracing of this ring by the reference sample allows SORT to recreate it while also preserving the void in the centre. We expect that such voids surrounded by a sufficiently high density of galaxies should largely be preserved in $z_{\mathrm{sort}}$. Reference galaxies are rarely found in voids, but may be shifted into them by redshift-space distortions in cases where dense clusters are positioned along the line of sight to the void. As such, SORT will primarily place galaxies around the voids where the reference galaxies reside. On the other hand, with photometric redshifts, the large uncertainties in the measurements of galaxies surrounding the void smooth out the region, obscuring the underlying structure as shown in the bottom left-hand panel.
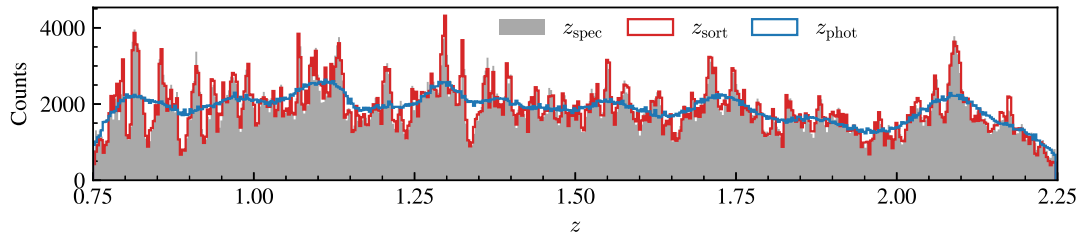


**Figure 4.** Redshift distributions for $z_{\mathrm{spec}}$ (grey), $z_{\mathrm{sort}}$ (red), and $z_{\mathrm{phot}}$ (blue) with arbitrary binning of 0.003. The large uncertainty of the photometric redshifts blurs out the structure of the distribution, which becomes more or less flat over the entire range. The distribution produced by SORT much more closely follows the distribution obtained with spectroscopic measurements. This is by design, as SORT samples new redshifts based on the distribution of the spectroscopic-quality reference sample within each sub-volume.

photometric redshifts were generated using a Gaussian distribution, the recovered distribution is Gaussian with a standard deviation of $\sim 0.01(1 + z)$. In red, the results of SORT show a significant fraction of redshifts that have been improved. Overall, $\Delta z_{\mathrm{sort}}$ and $\Delta z_{\mathrm{phot}}$ share a similar standard deviation; however, the large peak shows that $z_{\mathrm{sort}}$ provides much more information than $z_{\mathrm{phot}}$. This is shown clearly in both Figs 2 and 4 as $z_{\mathrm{sort}}$ is able to more accurately outline large-scale structure that is washed out by $z_{\mathrm{phot}}$.

A direct comparison of redshifts can be seen in Fig. 6. The left- and right-hand panels show the two-dimensional histograms of $z_{\mathrm{phot}}$ and $z_{\mathrm{sort}}$ compared to $z_{\mathrm{spec}}$. We continue to see improvement in redshift estimates after applying SORT. The large peak shown in Fig. 5 is now represented by a bright, narrow band of redshifts along the line of equality where errors are small. This improvement is seen in all redshift bins, which are shown in Fig. E5.
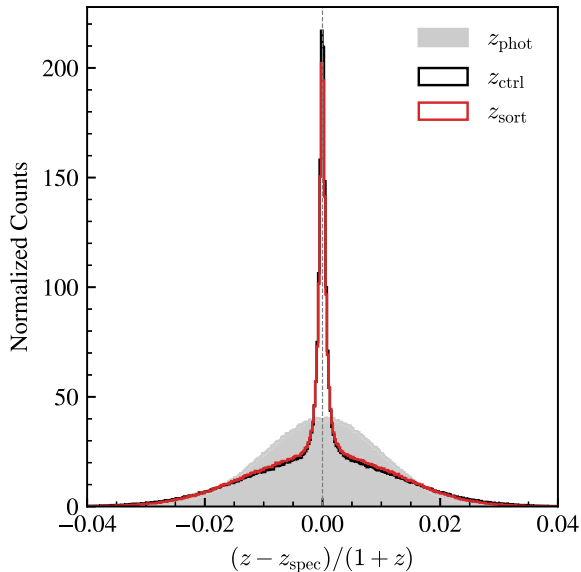
**Figure 5.** Normalized distribution of $\Delta z$ (excluding the spectroscopic sample) for $z_{\rm sort}$, $z_{\rm phot}$, and $z_{\rm ctrl}$ (see Section 5.1 for details on $z_{\rm ctrl}$.). The photometric distribution essentially recovers the Gaussian used to create the photometric sample. SORT is able to produce a tall peak surrounding $\Delta z = 0$ where a significant fraction of redshifts have been improved. The overall standard deviation of $\Delta z_{\rm sort}$ is comparable to $\Delta z_{\rm phot}$ as shown by the broader base of the distribution.

### 4.2 Recovering the 2PCF

The 2PCF is a relatively simple metric that provides information about the three-dimensional spatial clustering of galaxies. The large uncertainties associated with photometric redshifts lead to smoothing of spatial clustering and a highly biased estimate of the 3D 2PCF on relevant scales. As a result, using only photometric redshifts, one typically calculates the 2D angular 2PCF. Here we test SORT's ability to recover the 3D 2PCF. We note, however, that this test is somewhat conservative because redshift distortions and SORT only affect positioning along the line of sight. Angular correlations do not deviate from their true values.

Estimates of the 2PCF were calculated using various redshift types as a function of redshift-space distance $s$. Both $\xi_{\rm cos}(s)$ and $\xi_{\rm spec}(s)$ assume 100 per cent of the galaxies have a known cosmological or spectroscopic redshift.[8] $\xi_{\rm ref}(s)$ uses only the reference sample – i.e. only 10 per cent of galaxies with spectroscopic redshifts. For $\xi_{\rm phot}(s)$, $\xi_{\rm sort}(s)$, and $\xi_{\rm ctrl}(s)$ (see Section 5.1 for details on $\xi_{\rm ctrl}(s)$), 2PCFs were calculated using their respective 90 per cent non-reference sample redshifts plus the 10 per cent spectroscopic-quality reference sample. The 2PCFs were calculated using CORRFUNC (Sinha & Garrison 2020) from scales of $\sim$1 to $\sim$18–30 $h^{-1}$Mpc (larger scales are calculated in higher redshift bins). Fig. 7 shows 2PCF results for $0.75 < z < 1.25$ (see Fig. E3 for 2PCFs in all redshift bins). We note that because SORT is dependent upon $z_{\rm ref}$, which comprises a small fraction of the total redshifts, SORT's ability to reconstruct the three-dimensional distribution of galaxies is susceptible to sample variance in $z_{\rm ref}$. As such, SORT was run on the same light cone with 10 different random seeds (which determine the selection of $z_{\rm ref}$) to find an average result for the 2PCF. These averages are shown in Fig. 7 along with error bars calculated as the standard deviation

across the 10 random seeds. We also note that averaging over the 10 random seeds was done for all redshift types, though results will not vary much when using $z_{\rm cos}$, $z_{\rm spec}$, and $z_{\rm phot}$.

We see that using only photometric redshifts leads to a very poor estimate of the 2PCF. This is not surprising due to the large uncertainty associated with $z_{\rm phot}$. On the other hand, we observe that SORT accurately recovers the spectroscopic 2PCF on scales of $s \gtrsim 2.5$ $h^{-1}$Mpc. At the smallest scales, however, SORT overestimates the 2PCF. This result stems from the design of the SORT algorithm – namely, galaxies will be placed near other galaxies, resulting in high clustering on smaller scales.[9] The lower limit to which SORT is accurate will depend on the choice of binning the method uses. As our chosen bin width corresponds roughly to 1 $h^{-1}$Mpc, we can only expect to be reasonably accurate at scales larger than this.

### 4.3 Estimating local densities

Densities were calculated by searching for neighbours within cylindrical sub-volumes surrounding each galaxy. The total length of each cylinder was set to 4 $h^{-1}$Mpc. Photometric redshift uncertainties correspond to scales of $\sigma_z^{\rm ph} = 0.01(1 + z) \approx 30$ $h^{-1}$Mpc for $0.75 < z < 2.25$. This is much larger than the scale of the cylinder and thus photometric density estimates will be particularly poor. Nevertheless, we use this cylinder length to test the limits of SORT. We also provide a sample of density estimates using a much longer cylinder defined by $l = 2(\Delta v/c)(1 + z)$ with $\Delta v = 1000$ km s$^{-1}$ in Fig. E8.

The radius of the cylinder is initially set to $r = 0°\!.02$. If there are not at least $n$ galaxies within the cylinder, the radius is incremented by $\delta r = 0°\!.001$ until that condition is met or $r$ reaches $r_{\rm max} = 0°\!.04$. This method was chosen to make the calculation adaptive. The range of densities across the entire light cone is large, and having an adaptive aperture allows for probing different scales. The radius can start small to inspect high-density regions and expand in low-density regions to estimate an average density where there may otherwise be only one or two galaxies in the cylinder. The values for $r$, $\delta r$, and $r_{\rm max}$ were arbitrarily chosen to be similar to the parameters used by SORT. Likewise, the minimum threshold of neighbours was arbitrarily chosen to be $n = 5$. In principle, these parameters are all adjustable depending on how much ones wishes to constrain the densities. The results of SORT's estimations of local densities compared to photometric estimates are generally not sensitive to the choice in these parameters, though.

Fig. 8 shows three different density estimates in one redshift bin with $\rho_{\rm phot}$ in the top panels and $\rho_{\rm sort}$ in the bottom panels. The left-hand panels show number densities, the middle panels show stellar mass densities, and the right-hand panels show halo mass densities only considering central galaxies. The colour and contours are proportional to the maximum bin value within each of the individual subplots. The dashed contour (in red) is set to a limit equal to the minimum contour level in the corresponding $\rho_{\rm phot}$ subplot.

As expected, the photometric densities are underestimated in high-density regions. The high uncertainty of the photometric redshifts has the effect of smoothing out high- and low-density regions causing them to take on a more average density. After applying SORT, low bias in the high-density regime is greatly improved and the distributions become more aligned with the line of equality. As shown by the dashed contour, SORT struggles in the low-density regime, and the

---

[8]It is not expected that SORT should recover $\xi_{\rm cos}(s)$ since $z_{\rm ref}$ traces $z_{\rm spec}$, not $z_{\rm cos}$. These results are shown for the sake of comparison.

[9]In the previous SORT paper (Tejos et al. 2018), the 2PCF was *underestimated* on small scales. This difference stems from the fact that the previous paper did not include satellite galaxies while this one does.
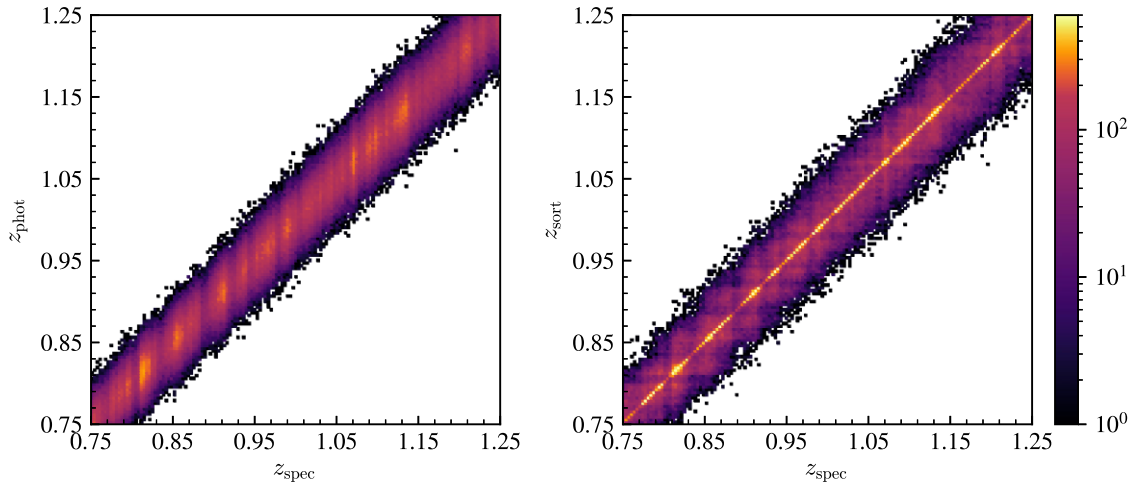
**Figure 6.** Two-dimensional redshift histograms for $z_{phot}$ and $z_{sort}$ relative to $z_{spec}$ with binning of 0.003. The colour bar represents the total number of counts in each bin. We observe significant improvement in redshift estimates by $z_{sort}$ compared to $z_{phot}$. There are much higher counts along the line of equality for $z_{sort}$, and this effect is consistent across the entire redshift range of the light cone. All redshift bins can be seen in Fig. E5.

scatter is comparable to the photometric estimates. SORT tends to overestimate its lowest local densities, which is a side effect of the clustering nature of the method. This is likely not something that can easily be remedied due to the simplicity of the SORT method. By design, SORT places galaxies near where it finds spectroscopic redshifts. Most spectroscopic redshifts will be in areas of higher density because this is where most galaxies are located. This tends to develop a cosmic structure that is highly clustered. SORT's ability to reconstruct low-density regions is dependent upon the quantity of high-quality redshifts found there, which will tend to be fairly limited.

SORT's estimates for central halo mass densities are not quite as good as its number and stellar mass densities, particularly for the mock CANDELS light cone where statistics are more limited (see Fig. A4). The likely cause of this is the removal of satellite galaxies. Halo masses for satellites are not tracked once they become sub-haloes and therefore were not considered in these calculations. The problem with this is that SORT is effective on average for the *full ensemble* of galaxies and does not discriminate different demographics (e.g. centrals versus satellites). We would not expect results to be as effective for a given subset of data since there is no mechanism within the method to treat different subsets differently. By removing satellites, we are decreasing the reliability of SORT, particularly in high-density regions where most satellites reside. However, we note that despite this, SORT still shows improvement over photometric density estimates.

## 5 DISCUSSION

### 5.1 The effects of preserving the rank order

One of the key aspects of SORT is the sorting itself. While it is clear that the reference sample provides a significant amount of information to SORT, one might wonder what the effect of sorting is (in step (vi) of Section 2.2). To test this, we ran a control algorithm that excluded step (vi) where the rank ordering is done. The control results were computed simultaneously with the standard SORT results and are identical in every way with the exclusion of the sorting; hence, the only difference in these two sets of results lies solely in the rank ordering. We call the results of this control algorithm $z_{ctrl}$.

The first result to consider is the $\Delta z$ histogram shown in Fig. 5. Performing a two-sided Kolmogorov–Smirnov test on $\Delta z_{sort}$ and $\Delta z_{ctrl}$ yields a *p*-value of $p < 0.001$. This indicates with a high level of certainty that sorting indeed changes the distribution of $\Delta z$. To understand the differences, we consider the two features of $\Delta z$: the narrow, central peak and the broader base.

When looking at the peak around $\Delta z = \pm 0.001$, we see the control sample outperforming SORT. To interpret this result, we plotted the same diagram with the data broken into a set of central galaxies and a set of satellite galaxies, shown in Fig. 9. The left-hand panel shows only satellites and the right-hand panel shows only centrals. Looking at the peaks shows that the difference between $\Delta z_{sort}$ and $\Delta z_{ctrl}$ arises in the satellites. Satellite galaxies will tend to be situated more closely to their neighbours than a central galaxy. As such, the peak of $\Delta z^{sat}$ will tend to favour environments that are more densely packed. This is precisely what the control sample provides.

To illustrate this, consider some region of space containing a dense cluster of galaxies. If we assume there are *N* galaxies along a pencil-beam-like sub-volume encompassing this dense cluster, we would expect each of those galaxies to have $\sim N$ recovered redshifts after SORT is complete. In other words, since the cluster is dense, we expect most of the galaxies to fall within the sub-volumes of their neighbours. The sorting aspect of SORT will always assign the lowest redshift galaxies in this region the lowest recovered redshifts, and likewise assign the highest redshift galaxies the highest recovered redshifts. This is simply following the condition laid out by stochastic ordering. Recalling that $z_{sort}$ is taken to be the median of a galaxy's assigned recovered redshifts, galaxies on the lower redshift end of the cluster are biased to have a lower $z_{sort}$ and vice-versa at the higher redshift end. In contrast to this, the control sample has no such bias. Each of the galaxies in the cluster will receive a random recovered redshift. After the algorithm completes, each galaxy will have a mixture of $\sim N$ high and low recovered redshifts that will tend to have a median towards the centre of the cluster. This centralization makes the cluster more dense than SORT would make it, thus favoring $\Delta z^{sat}$.

Let us now consider the right-hand panel of Fig. 9, $\Delta z^{cent}$. In this case, we have observed that there is no appreciable difference between $\Delta z_{sort}$ and $\Delta z_{ctrl}$ when it comes to the peak. However, if we look beyond the peak, we can see that SORT performs better than the
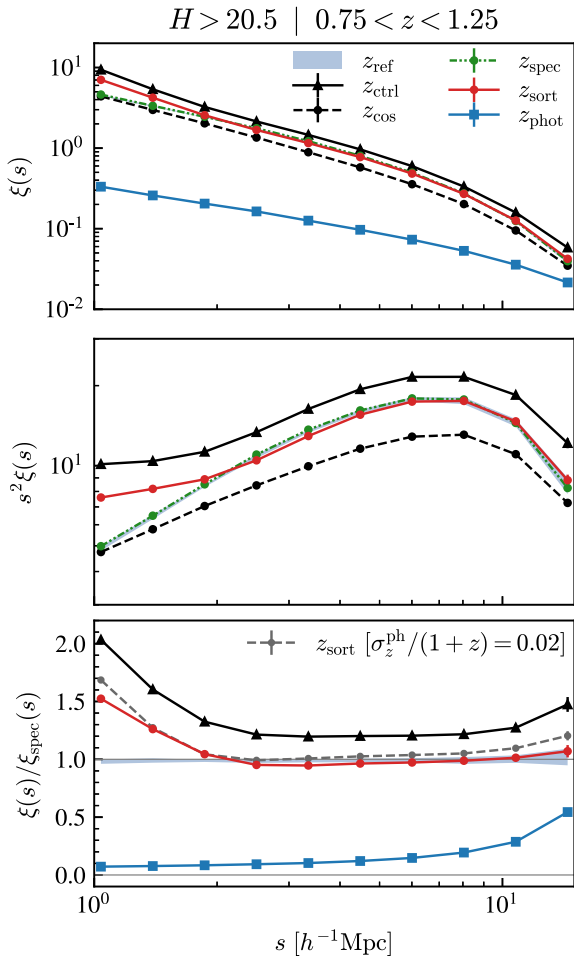
**Figure 7.** 2PCFs of various redshift types as a function of redshift-space distance $s$ shown in three different ways. The results show the mean value of the 2PCFs along with $1\sigma$ error bars after running SORT with 10 different random seeds to determine the reference sample selection. Note that the error bars are too small to be seen. We observe that $\xi_{phot}(s)$ is a poor estimate of the 2PCF at all shown length-scales, and $z_{ctrl}$ consistently overestimates the 2PCF while $\xi_{sort}(s)$ is accurate (relative to $\xi_{spec}(s)$) for $s \gtrsim 2.5\ h^{-1}$Mpc. See Section 5.1 for details on $z_{ctrl}$ (shown as black triangles) and Appendix C for details on $\xi_{sort}(s)$ using $\sigma_z^{ph}/(1+z) = 0.02$ (shown as the grey dashed line in the bottom panel).

control sample. $\Delta z_{sort}$ tends to have higher counts than $\Delta z_{ctrl}$ up until the point where the tails of their distributions become broader than that of $\Delta z_{phot}$, around $\Delta z^{cent} = \pm 0.015$. Beyond this point, $\Delta z_{sort}$ has a steeper distribution, signifying its better overall recovery of redshift estimates. This relative shape is also present for the $\Delta z$ histogram of the entire set of galaxies, though difficult to see in Fig. 5.

To further investigate the effects of sorting, we can consider the 2PCF. This metric provides a better characterization of the full three-dimensional distribution of galaxies than $\Delta z$. Fig. 7 shows a clear distinction between $z_{sort}$ and $z_{ctrl}$. SORT is able to accurately reproduce $\xi_{spec}(s)$ on scales of $s \gtrsim 2.5\ h^{-1}$Mpc. Due to the centralization and higher density produced by the control sample, $\xi_{ctrl}(s)$ ends up on average around 25 per cent higher than $\xi_{sort}(s)$. In other words, $z_{ctrl}$ is overestimating the clustering while $z_{sort}$ is not.

We conclude that while much of the information is provided by the reference sample, the sorting aspect of SORT certainly provides useful information as well. This information is most evident when considering the 2PCF where the lack of sorting leads to overclus-

tering by around 25 per cent. The only drawback to sorting comes with the $\Delta z$ histogram of satellite galaxies. This is a difficult issue to resolve because SORT uses one prescription to treat two distinct demographics, and information about which galaxies are centrals or satellites is not readily available for real observations.

### 5.2 Sub-volume parameters

The SORT parameters $N_{ref}^{min}$, $R$, and $\Delta_z$ determine the sizes of the cylindrical sub-volumes that surround each galaxy during the SORT procedure. A balance must be struck for these parameters in order for SORT to produce reasonable results.

Having a larger cylinder radius allows for more of the environment to be taken into consideration when looking for reference galaxies. This can be useful in cases where galaxies are near the outer edge of a large cluster of galaxies. If the radius is too small, the inner region of the cluster may not be detected by the pencil-beam-like sub-volume. This leaves the galaxy more susceptible to being pulled toward denser regions that may be close on the sky but not in redshift. On the other hand, making the radius too large can also be problematic. With a large radius, galaxies that are not particularly nearby on the sky, but still within the sub-volume, will be pulled toward the redshifts of denser regions. Because SORT only moves galaxies along the line of sight, these galaxies will be placed around the same redshift as another group of galaxies but with a seemingly 'incorrect' position on the sky. The result is a distribution of galaxies that becomes elongated in a plane perpendicular to the line of sight.

This effect can be seen in Figs 3 and E4 and is a signature of the SORT method. In low-density regions (e.g. the upper right corner of the $z_{sort}$ panel), we see horizontal formations of galaxies. The magnitude of this effect can be limited by adjusting $R_{max}$ or $N_{ref}^{min}$. As $N_{ref}^{min}$ becomes smaller, the radius of the average sub-volume will also be smaller, leading to narrower horizontal formations. This may come at the expense of SORT's overall performance, though. We found that increasing $N_{ref}^{min}$ from two to four provided a better estimate of the 2PCF, for example. If, instead, $R_{max}$ is adjusted, one must take care to not make it too small relative to $N_{ref}^{min}$. If $R_{max}$ is too small, the fail rate of the SORT algorithm will increase as the maximum sub-volume size is too constrained to find enough reference galaxies. Likewise, the fail rate will also increase if $N_{ref}^{min}$ is too large for a given sub-volume size.

The redshift cut imposed by $\Delta_z$ is a new addition to SORT. In the original SORT paper (Tejos et al. 2018), which looked at nearby galaxies in a wider, shallower field, an apparent magnitude cutoff was imposed such that only galaxies within $\pm \delta m$ of the $i$th galaxy were considered in the sub-volume. When SORT is applied to a deeper field, the magnitude cut is not sufficient to allow SORT to perform well. The same range of magnitudes can be found at opposite ends of the light cone, which leaves too large of a range of possible recovered redshifts for a given galaxy.

To limit the range of redshifts that are considered neighbours of the $i$th galaxy, a redshift cut based on a galaxy's photometric redshift was implemented to replace the magnitude cut. The length chosen for the cylinder should be based on the photometric uncertainty. In this case, we have assumed the photometric uncertainties are Gaussian. As such, we have chosen $\Delta_z = 2.5\sigma_z^{ph}$ to allow the majority of photometric galaxies the potential of recovering their true redshift. The value of this parameter was not thoroughly tested, however, and may not be optimal. We emphasize that this parameter, as well as other SORT parameters, should be tested to find optimal values when applied to different surveys. The values used in this paper correspond to sensible length-scales, but optimal values will likely vary depending on the metric one wishes to optimize.
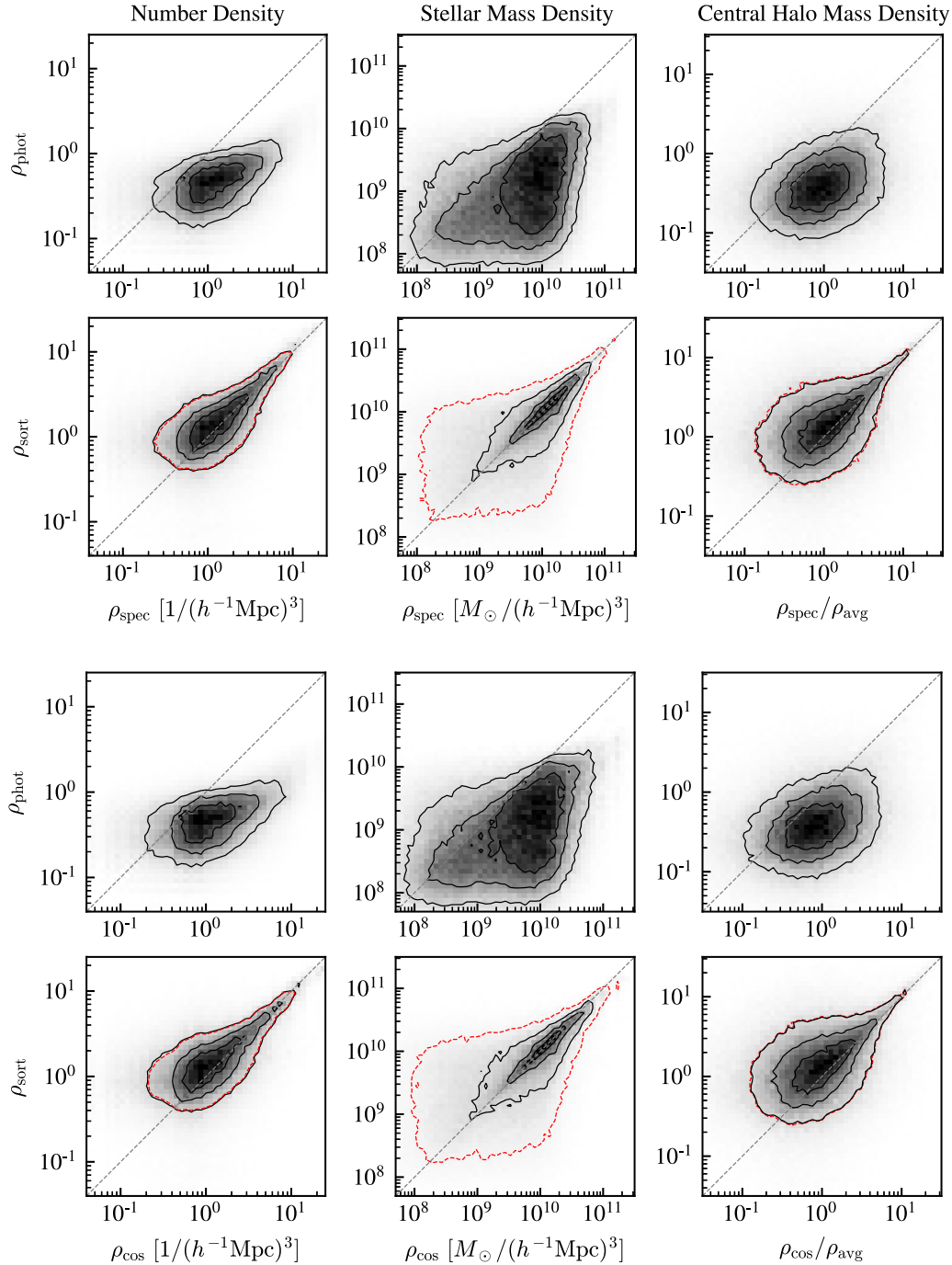
**Figure 8.** Two-dimensional density histograms for $\rho_{\mathrm{phot}}$ and $\rho_{\mathrm{sort}}$ in the range $0.75 < z < 1.25$. The left-hand panels show number densities, the middle panels show stellar mass densities, and the right-hand panels show halo mass densities using only central galaxies. The top six panels compare densities to estimates using $z_{\mathrm{spec}}$, and the bottom six panels compare densities to estimates using $z_{\mathrm{cos}}$. The solid contours represent limits of 25, 50, and 75 per cent of the maximum bin value in each subplot. The dashed contour (red) is set at a limit equal to the minimum contour level in the corresponding $\rho_{\mathrm{phot}}$ subplot. Densities were calculated with a fixed cylinder length of $4\,h^{-1}\mathrm{Mpc}$ and a radius starting at $\sim 1\,h^{-1}\mathrm{Mpc}$ and expanding up to $\sim 2\,h^{-1}\mathrm{Mpc}$ as needed to encompass at least five galaxies. As expected, the photometric densities estimates are all poor as the cylinder length-scales are much smaller than the typical photometric redshift error. SORT densities show significant improvement in regions with average or higher density. Overall scatter is similar when comparing $\rho_{\mathrm{sort}}$ to $\rho_{\mathrm{phot}}$, but $\rho_{\mathrm{sort}}$ displays much better alignment with the line of equality and a more peaked distribution surrounding it. See Section 4.3 for details on the central halo mass densities.

## 5.3 Limitations

One main limitation of the SORT method is its dependence on a reference sample. The limitations of this dependence are twofold. First, there is a limitation to the length-scale that SORT will be able to properly recover. Dispersion velocities of $v \approx 200\,\mathrm{km\,s^{-1}}$ correspond

to lengths of $\sim 4$–$6\,h^{-1}\mathrm{Mpc}$ for $z = 1$–$2$. This is a rough limit of SORT's accuracy relative to the *true* distribution of galaxies (i.e. not the spectroscopic distribution to which results were compared in this paper) due to redshift-space distortions. Second, SORT requires a structural outline by the reference sample to recover an accurate
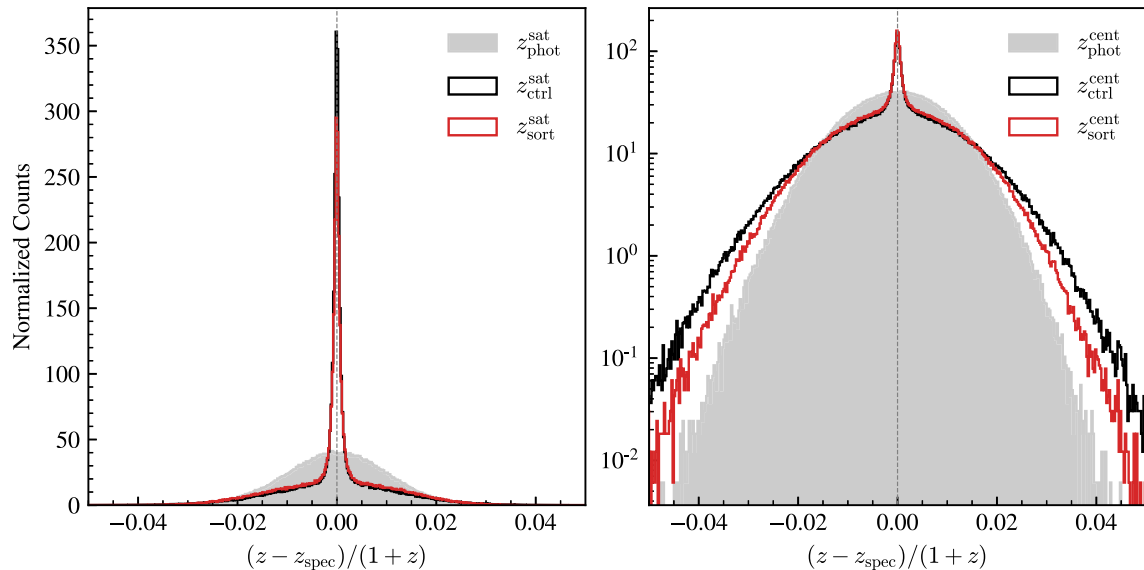
**Figure 9.** Normalized distribution of $\Delta z$ for $z_{sort}$, $z_{ctrl}$, and $z_{phot}$. The left-hand panel shows the results using only satellite galaxies and the right-hand panel shows the results using only central galaxies. See Section 5.1 for details. The left-hand panel shows that the more highly-clustered results of $z_{ctrl}$ favour satellite galaxies. The right-hand panel shows similar results between $z_{sort}$ and $z_{ctrl}$ at smaller errors; however, SORT displays better treatment at larger errors ($\Delta z/(1 + z) \gtrsim 0.02$) with more rapidly declining tails.

distribution of galaxies. If the reference sample does not outline a feature of the cosmic web, then SORT will not be able to reconstruct it. This effect is most significant in low-density regions. The fraction of galaxies found in low-density regions will naturally be low. An even lower fraction of those galaxies will be reference galaxies. Without reference galaxies, SORT will not be able to reproduce an accurate distribution of galaxies in these regions.

A second limitation of the SORT method is the fact that it can only improve redshift estimates collectively for ensembles of galaxies. Fig. 5 shows a tall peak surrounding $\Delta z = 0$ where a significant fraction of redshifts have been improved, but there is no way to tell which galaxies are in this peak or which galaxies are in the tails. Despite SORT doing a fairly good job of recovering the large-scale structure of galaxies, there will undoubtedly be some galaxies placed in the wrong environments. Higher accuracy redshift estimates are still required to properly place galaxies on an individual basis. However, as shown in the previous section, SORT can still be used to infer local densities (particularly, average or higher densities) reasonably well.

### 5.4 Future considerations

One possible next step for SORT is updating the assignment of spectroscopic redshifts to create the reference sample. In this paper, spectroscopic redshifts were assigned randomly to 10 per cent of galaxies within three complete redshift bins. To make tests of the SORT method more realistic, one could model the reference galaxy selection using the methods chosen by large imaging surveys (e.g. selecting a mixture of brighter galaxies and galaxies with high star formations rates that produce strong emission lines).

Another possible step is to improve redshift assignments within sub-volumes as the SORT method is carried out. There are currently no considerations given to the angular correlations within each sub-volume. As shown in the left-hand panel of Fig. 9, the current treatment of satellite galaxies by SORT is not optimal.[10] This could

potentially be improved by assigning similar recovered redshifts to galaxies that appear highly clustered on the sky. This could also be implemented when determining the final $z_{sort}$ redshift of a galaxy. The final selection from a galaxy's pool of recovered redshifts at the end of the method could be biased to redshifts where the galaxy appears more clustered on the sky, as opposed to taking a simple median. Such considerations could also prove useful in reducing the horizontal structures produced by SORT discussed in Section 5.2.

Yet another improvement would be to treat satellite galaxies more realistically than we have done, as described in Appendix B. For example, an improved semi-analytic treatment of satellite galaxies could be based on the recent SatGen papers (Green, van den Bosch & Jiang 2021a,b; Jiang et al. 2021).

Recently, it has been shown that correlations between galaxy and halo properties create observable signatures in local environments (Behroozi et al. 2022). In particular, halo spin, concentration, growth rate, and interaction history have all been shown to leave scale-dependent signatures in both 2PCFs and the distributions of distances to galaxies' $k$th nearest neighbours out to $z \sim 2.5$. These determinations were based on projected two-dimensional environments so as to make them observationally accessible with low-resolution spectroscopy ($\sigma_z/(1 + z) \lesssim 0.005$). We have shown that SORT is able to recover the full three-dimensional 2PCF as estimated with high-resolution spectroscopy. We also provide in Appendix E results using SORT to estimate three-dimensional distances to $k$th nearest neighbours (see Fig. E9). We expect SORT's ability to reasonably-well recover local environments may allow for further observationally-accessible detections of environmental signatures that result from galaxy–halo property correlations.

In this paper, we have not taken advantage of the fact that galaxy properties could depend on environment and/or location within the cosmic web. Indeed, it is well known that, for example, star-forming/blue galaxies are less clustered than more quiescent / red galaxies (see e.g. Li et al. 2006; Zehavi et al. 2011; Coil et al. 2017; Berti et al. 2019, 2021) and that more spheroid-like morphologies are more frequently in denser environments (e.g. Dressler 1980; Pearson et al. 2021). Thus, a natural next step within the framework of SORT

---

[10]That is not to say that $z_{ctrl}$ *is* optimal, but it is enough to demonstrate that SORT is not.

would be to divide the reference sample by galaxy properties. By doing so, it is expected that SORT would be able to determine even more accurate redshifts than when not considering galaxy properties.

Finally, we expect to test the performance of SORT using real data sets from highly complete spectroscopic galaxy surveys, e.g. GAMA (Baldry et al. 2018) and DESI (DESI Collaboration et al. 2016), in order to account for systematics that are present in real surveys but not properly modelled by our mock experiment. For instance, we expect that the so-called 'catastrophic redshift' failures in photometric redshift methods will have a minor effect in the performance of SORT as these are typically only a small fraction of the total sample. Other systematic differences include having non-Gaussian PDFs for the photometric redshifts and having a set of photometric galaxies with variable $\sigma_z^{\rm ph}$ in the sample. For example, COSMOS2020 (Weaver et al. 2021) obtained photometric redshift precision of ~4 per cent for the faintest galaxies and better than 1 per cent for the brightest galaxies.

## 6 SUMMARY AND CONCLUSIONS

In this paper, we have tested the performance of the SORT method (Tejos et al. 2018) in mock high-redshift surveys. SORT is a simple, efficient, and robust method that can be used to improve redshift estimates. It relies upon a reference sample of high-quality spectroscopic redshifts for which a precise distribution $dN/dz$ is known within pencil-beam-like sub-volumes of the survey. Within each sub-volume, we:

(i) sample new 'recovered' redshifts from the $dN/dz$ distribution of high-quality redshifts; and

(ii) match the recovered redshifts one-to-one with the low-quality (photometric) redshifts such that the rank order is preserved.

The second step is motivated by the fact that random variables drawn from Gaussian PDFs with equal, arbitrarily-large standard deviations satisfy stochastic ordering. In other words, if two redshift estimates $z_i$ and $z_j$ satisfy $z_i < z_j$, then their *true* redshift values most likely satisfy $z_i^{\rm true} \leq z_j^{\rm true}$. Thus, preserving the rank order makes the assigned recovered redshifts more likely to be close to their underlying true value. This process is repeated for sub-volumes surrounding each galaxy in the survey. The result is every galaxy with a low-quality redshift is assigned multiple recovered redshifts from which a new redshift estimate can be determined.

We ran the SORT method on a wide-field $2 \deg^2$ mock light cone and a mock CANDELS light cone extracted from the Small MultiDark–Planck and Bolshoi–Planck $N$-body simulations, respectively, to test its performance in a pencil-beam-like survey spanning the redshift range $0.75 < z < 2.25$. After applying SORT, we observe similar improvement from both mock catalogues and make the following determinations:

(i) We observed overall improvement in redshift estimates, allowing for better reconstruction of the three-dimensional distribution of galaxies than photometric redshifts alone provide. This can be seen broadly in Fig. 2 or more close up in Fig. 3.

(ii) We also observed that SORT produces much better agreement in the one-dimensional $dN/dz$ distribution (by design), allowing it to better identify the large-scale structure along the line of sight as shown in Fig. 4.

(iii) Distributions of redshift errors with respect to spectroscopic estimates were overall improved throughout the light cone, while the standard deviations remained about the same. One- and two-

dimensional histograms of this are shown in Figs 5 and 6, respectively.

(iv) SORT accurately recovers the spectroscopic redshift-space 2PCF down to scales of $\gtrsim 2.5\ h^{-1}$Mpc while photometric redshifts (with errors of $\sigma_z^{\rm ph}/(1+z) = 0.01$ corresponding to scales of ~20–30 $h^{-1}$Mpc) drastically underestimate galaxy clustering. This is shown clearly in the top and bottom panels of Fig. 7.

(v) SORT is able to recover three-dimensional local densities in regions of average or higher density at scales of $\gtrsim 4\ h^{-1}$Mpc. Three different density histograms are shown in Fig. 8, and additional histograms are shown in Fig. E8 at a larger length-scale of $l = 2(1000\ {\rm km\ s^{-1}}/c)(1+z)$.

We expect that such improved determinations of local galaxy environments will help to distinguish the effects of environmental properties (e.g. local density) on galaxy evolution from other effects, such as galaxy stellar or halo mass (e.g. Peng et al. 2010; Woo et al. 2013; Chartab et al. 2020; Behroozi et al. 2022).

## DATA AVAILABILITY

The mock CANDELS backward light cone used in this paper, as well as additional light cones for other CANDELS fields, are available at https://www.simonsfoundation.org/candels-survey.

## REFERENCES

Aragon-Calvo M. A., van de Weygaert R., Jones B. J. T., Mobasher B., 2015, MNRAS, 454, 463
Astropy Collaboration et al., 2013, A&A, 558, A33
Astropy Collaboration et al., 2018, AJ, 156, 123
Baldry I. K. et al., 2018, MNRAS, 474, 3875
Behroozi P. S., Wechsler R. H., Wu H.-Y., 2013, ApJ, 762, 109
Behroozi P., Wechsler R. H., Hearin A. P., Conroy C., 2019, MNRAS, 488, 3143
Behroozi P., Hearin A., Moster B. P., 2022, MNRAS, 509, 2800
Berlind A. A., Weinberg D. H., 2002, ApJ, 575, 587
Berti A. M., Coil A. L., Hearin A. P., Moustakas J., 2019, ApJ, 884, 76
Berti A. M., Coil A. L., Hearin A. P., Behroozi P. S., 2021, AJ, 161, 49

[11]https://www.python.org/.
[12]https://www.scipy.org/ (Virtanen et al. 2020).
[13]https://numpy.org/ (Harris et al. 2020).
[14]https://matplotlib.org/ (Hunter 2007).
[15]https://www.astropy.org/ (Astropy Collaboration et al. 2013, 2018).

Bruzual G., Charlot S., 2003, MNRAS, 344, 1000
Cawthon R. et al., 2020, preprint (arXiv:2012.12826)
Chartab N. et al., 2020, ApJ, 890, 7
Coil A. L., Mendez A. J., Eisenstein D. J., Moustakas J., 2017, ApJ, 838, 87
Cooray A., Sheth R., 2002, Phys. Rep., 372, 1
DESI Collaboration et al., 2016, preprint (arXiv:1611.00036)
Dressler A., 1980, ApJ, 236, 351
Gatti M. et al., 2022, MNRAS, 510, 1223
Green S. B., van den Bosch F. C., Jiang F., 2021a, MNRAS, 508, 2944
Green S. B., van den Bosch F. C., Jiang F., 2021b, MNRAS, 509, 2624
Harris C. R. et al., 2020, Nature, 585, 357
Hildebrandt H. et al., 2021, A&A, 647, A124
Hunter J. D., 2007, Comput. Sci. Eng., 9, 90
Ivezić Ž. et al., 2019, ApJ, 873, 111
Jiang F., Dekel A., Freundlich J., van den Bosch F. C., Green S. B., Hopkins
    P. F., Benson A., Du X., 2021, MNRAS, 502, 621
Klypin A. A., Trujillo-Gomez S., Primack J., 2011, ApJ, 740, 102
Klypin A., Yepes G., Gottlöber S., Prada F., Heß S., 2016, MNRAS, 457,
    4340
Landy S. D., Szalay A. S., Koo D. C., 1996, ApJ, 460, 94
Lange J. U., van den Bosch F. C., Zentner A. R., Wang K., Villarreal A. S.,
    2019, MNRAS, 487, 3112
Lee C. T., Primack J. R., Behroozi P., Rodríguez-Puebla A., Hellinger D.,
    Dekel A., 2017, MNRAS, 466, 3834
Li C., Kauffmann G., Jing Y. P., White S. D. M., Börner G., Cheng F. Z.,
    2006, MNRAS, 368, 21
Ménard B., Scranton R., Schmidt S., Morrison C., Jeong D., Budavari T.,
    Rahman M., 2013, preprint (arXiv:1303.4722)
Morrison C. B., Hildebrandt H., Schmidt S. J., Baldry I. K., Bilicki M., Choi
    A., Erben T., Schneider P., 2017, MNRAS, 467, 3576
Navarro J. F., Frenk C. S., White S. D. M., 1996, ApJ, 462, 563
Navarro J. F., Frenk C. S., White S. D. M., 1997, ApJ, 490, 493
Pearson W. J., Wang L., Brough S., Holwerda B. W., Hopkins A. M., Loveday
    J., 2021, A&A, 646, A151
Peng Y.-j. et al., 2010, ApJ, 721, 193
Phillipps S., Shanks T., 1987, MNRAS, 229, 621
Rahman M., Ménard B., Scranton R., Schmidt S. J., Morrison C. B., 2015,
    MNRAS, 447, 3500
Rodríguez-Puebla A., Behroozi P., Primack J., Klypin A., Lee C., Hellinger
    D., 2016, MNRAS, 462, 893
Scaramella R. et al., 2021, preprint (arXiv:2108.01201)
Schmidt S. J., Ménard B., Scranton R., Morrison C., McBride C. K., 2013,
    MNRAS, 431, 3307
Scottez V. et al., 2016, MNRAS, 462, 1683
Scottez V., Benoit-Lévy A., Coupon J., Ilbert O., Mellier Y., 2018, MNRAS,
    474, 3921
Seldner M., Peebles P. J. E., 1979, ApJ, 227, 30
Shuntov M. et al., 2020, A&A, 636, A90
Sinha M., Garrison L. H., 2020, MNRAS, 491, 3022
Somerville R. S., Davé R., 2015, ARA&A, 53, 51
Somerville R. S., Kolatt T. S., 1999, MNRAS, 305, 1
Somerville R. S., Primack J. R., 1999, MNRAS, 310, 1087
Somerville R. S., Hopkins P. F., Cox T. J., Robertson B. E., Hernquist L.,
    2008, MNRAS, 391, 481
Somerville R. S., Gilmore R. C., Primack J. R., Domínguez A., 2012,
    MNRAS, 423, 1992
Somerville R. S., Popping G., Trager S. C., 2015, MNRAS, 453, 4337
Somerville R. S. et al., 2021, MNRAS, 502, 4858
Sousbie T., 2011, MNRAS, 414, 350
Spergel D. et al., 2015, preprint (arXiv:1503.03757)
Tejos N., Rodríguez-Puebla A., Primack J. R., 2018, MNRAS, 473, 366
van den Bosch F. C., Lange J. U., Zentner A. R., 2019, MNRAS, 488, 4984
Virtanen P. et al., 2020, Nat. Methods, 17, 261
Watson D. F., Berlind A. A., McBride C. K., Hogg D. W., Jiang T., 2012,
    ApJ, 749, 83
Weaver J. R. et al., 2021, ApJS, 258, 11
Wechsler R. H., Tinker J. L., 2018, ARA&A, 56, 435
Woo J. et al., 2013, MNRAS, 428, 3306

Yang S., Somerville R. S., Pullen A. R., Popping G., Breysse P. C., Maniyar
    A. S., 2021, ApJ, 911, 132
Yung L. Y. A., Somerville R. S., Finkelstein S. L., Popping G., Davé R.,
    2019a, MNRAS, 483, 2983
Yung L. Y. A., Somerville R. S., Popping G., Finkelstein S. L., Ferguson H.
    C., Davé R., 2019b, MNRAS, 490, 2855
Zehavi I. et al., 2011, ApJ, 736, 59

## APPENDIX A: SORT PERFORMANCE IN A MOCK CANDELS LIGHT CONE

In addition to the 2 deg$^2$ wide-field light cone, SORT was tested on a narrower $17 \times 41$ arcmin$^2$ light cone with 47 404 galaxies ($\sim$68 galaxies per arcmin$^2$). All model parameters were kept the same, and we observed that SORT continues to perform well at improving redshift estimates and determining local galaxy environments. Here we provide parallel results of the main text for SORT applied to this light cone with minor adjustments detailed hereafter.

Figs A1 and A2 show the one-and two-dimensional redshift histograms comparing $z_{\text{sort}}$ and $z_{\text{phot}}$ to $z_{\text{spec}}$. Fig. A1 is a direct parallel to Fig. 5 and shows similar overall improvement in redshift errors. For Fig. A2, however, because statistics are more limited in the mock CANDELS light cone, all three redshift ranges of width $\Delta z = 0.5$ were stacked on top of each other to create composite histograms. This allows recovery of similar distributions to those shown in Fig. 6. In particular, the photometric redshifts maintain their Gaussian error distributions while SORT redshifts build up along the line of equality.

As before, SORT was run on the same light cone with 10 different random seeds (which determine the selection of $z_{\text{ref}}$) to find an average result for the 2PCF. The average 2PCF estimates are shown in Fig. A3 for each of the redshift types in the range $0.75 < z < 1.25$. The error bars represent the standard deviations of $\xi(s)$ for each of the 10 seeds within each bin. Due to the narrower geometry of this light cone compared to the wide-field light cone, the 2PCFs were only calculated out to 8 $h^{-1}$Mpc. Similar to Fig. 7, we see that SORT provides a much better estimate of the 2PCF than using photometric redshifts.

## APPENDIX B: ASSIGNING THREE-DIMENSIONAL COORDINATES TO SATELLITE GALAXIES

The mock galaxy surveys used for this paper do not provide full three-dimensional coordinates for satellite galaxies. Instead, all satellite galaxies are assigned the same redshift as the dark matter halo they occupy. We used the following procedure to assign new three-dimensional coordinates and velocities to each of the satellites.

### B1 Calculating the position

New positions were calculated for each satellite assuming that the satellites have the same radial distribution as the dark matter (see e.g. Berlind & Weinberg 2002; Cooray & Sheth 2002).[16] The radial density profile was approximated using the NFW formula (Navarro, Frenk & White 1996, 1997):

$$\rho_{\text{NFW}}(r) = \frac{4\rho_{\text{s}}}{(r/R_{\text{s}})(1 + r/R_{\text{s}})^2}. \tag{B1}$$

[16]Improved treatments based on observations are discussed in Watson et al. (2012), Wechsler & Tinker (2018), Lange et al. (2019), and van den Bosch, Lange & Zentner (2019).
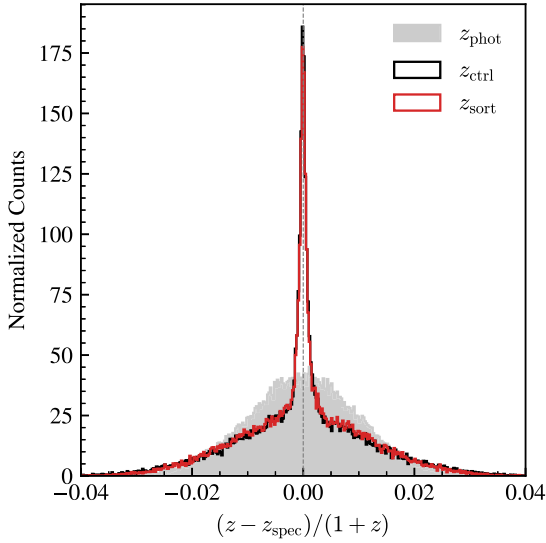
**Figure A1.** Normalized distribution of $\Delta z$ (excluding the spectroscopic sample) for $z_{\rm sort}$, $z_{\rm phot}$, and $z_{\rm ctrl}$ using the mock CANDELS light cone (see Section 5.1 for details on $z_{\rm ctrl}$). We recover a distribution of redshift errors similar to Fig. 5 using the wide-field light cone. In particular, the $\Delta z_{\rm sort}$ distribution is dominated by a tall, central peak of improved redshifts.

It is determined by two parameters, in this case $\rho_{\rm s}$ and $R_{\rm s}$. Alternatively, it can be determined by the halo mass, $M_{\rm vir}$, and the halo concentration, $c_{\rm vir}$, which is defined as

$$c_{\rm vir} = \frac{R_{\rm vir}}{R_{\rm s}}. \tag{B2}$$

The scaleradius, $R_{\rm s}$, is the radius at which the log-space derivative of $\rho_{\rm NFW}(r)$ is -2. This could be found by fitting the NFW profile to each halo in the simulation. However, a more robust method is to find the Klypin scaleradius using the $M_{\rm vir}$–$V_{\rm max}$ relation under the assumption of an NFW profile (Klypin, Trujillo-Gomez & Primack 2011). The parameter $V_{\rm max}$ is the maximum circular velocity of the halo (i.e. the maximum of $\sqrt{GM(r)/r}$, where $M(r)$ is the mass enclosed within a radial distance $r$). For an NFW profile, the maximum circular velocity occurs at $R_{\rm max} = 2.1626 R_{\rm s}$ (Klypin et al. 2011; Behroozi et al. 2013). With this, we calculated the Klypin concentration, $c_{\rm vir, K}$, by numerically solving

$$\frac{c_{\rm vir,K}}{f(c_{\rm vir,K})} = V_{\rm max}^2 \frac{R_{\rm vir}}{GM_{\rm vir}} \frac{2.1626}{f(2.1626)}, \tag{B3}$$

where

$$f(x) = \ln(1 + x) - x/(1 + x). \tag{B4}$$

The radial distribution for an NFW profile can also be written in terms of the halo's mass as

$$M_{\rm h}(r) = M_{\rm vir} \times u_{\rm vir}(r), \tag{B5}$$

where $u_{\rm vir}(r)$ is

$$u_{\rm vir}(r) = \frac{\ln(1 + c_{\rm vir,K}x) - c_{\rm vir,K}x/(1 + c_{\rm vir,K}x)}{\ln(1 + c_{\rm vir,K}) - c_{\rm vir,K}/(1 + c_{\rm vir,K})}, \tag{B6}$$

with $x = r/R_{\rm vir}$. We can use this to sample new radial positions for the satellites within a halo. For each satellite in a given halo, the following procedure was followed:

(i) Generate three random numbers $U_r$, $U_\theta$, and $U_\phi$, each uniformly distributed between 0 and 1.

(ii) Sample a radius from the distribution $u_{\rm vir}(r)$. This can be done by finding the value $r$ such that $U_r - u_{\rm vir}(r) = 0$.

(iii) Assign the new spherical coordinates $(r, \theta, \phi)$ to the satellite relative to the halo's centre where $\theta = \pi U_\theta$ and $\phi = 2\pi U_\phi$.

(iv) Assign new Cartesian coordinates $\boldsymbol{r} = (x, y, z)$ relative to the halo using

$$x = r \sin\theta \cos\phi, \tag{B7}$$

$$y = r \sin\theta \sin\phi, \tag{B8}$$

$$z = r \cos\theta. \tag{B9}$$

(v) Get the position of the satellite relative to the box of the simulation using $\boldsymbol{R} = \boldsymbol{R}_{\rm h} + \boldsymbol{r}$, where $\boldsymbol{R}_{\rm h}$ is the position of the halo relative to the simulation box. A cosmological redshift, $z_{\rm cos}$, can be inferred from the new satellite position.

## B2 Calculating the velocity

To find the line-of-sight redshift, $z_{\rm los}$, of each satellite, we must account for the effects of the peculiar velocity along the line of sight. The peculiar velocity of each satellite will depend on its position within a halo. Using the new radial distribution of satellites, as well as properties of the haloes within which the satellites reside, we can estimate new satellite velocities:

(i) By assuming that the satellite velocities trace the dark matter particle velocities within an NFW halo, we can calculate the velocity dispersion of the satellites at a distance $r$ from the halo's centre using

$$\sigma^2(r) = \frac{c_{\rm vir,K} V_{\rm vir}^2}{\mu(c_{\rm vir,K})} \frac{r}{R_{\rm s}} \left(1 + \frac{r}{R_{\rm s}}\right)^2 \int_{r/R_{\rm s}}^{\infty} \frac{\mu(x){\rm d}x}{x^3(1 + x)^2}, \tag{B10}$$

where $\mu(x)$ is defined as

$$\mu(x) = \ln(1 + c_{\rm vir,K}x) - c_{\rm vir,K}x/(1 + c_{\rm vir,K}x). \tag{B11}$$

(ii) Sample a velocity $v$ from the Gaussian distribution

$$P(v) = \frac{1}{\sqrt{2\pi\sigma^2(r)}} \exp\left(-\frac{v^2}{2\sigma^2(r)}\right). \tag{B12}$$

(iii) Generate two random numbers $U_\theta$ and $U_\phi$, each uniformly distributed between 0 and 1.

(iv) Using $\theta = \pi U_\theta$ and $\phi = 2\pi U_\phi$, the components of the satellite's velocity vector $\boldsymbol{v} = (v_x, v_y, v_z)$ relative to the halo's centre are

$$v_x = v \sin\theta \cos\phi, \tag{B13}$$

$$v_y = v \sin\theta \sin\phi, \tag{B14}$$

$$v_z = v \cos\theta. \tag{B15}$$

(v) With respect to the box of the simulation, the satellite's velocity is $\boldsymbol{V} = \boldsymbol{V}_{\rm h} + \boldsymbol{v}$, where $\boldsymbol{V}_{\rm h}$ is the halo's velocity with respect to the simulation box.

The component of the velocity along the line of sight can be found by the new position and velocity vectors:

$$v_{\rm los} = \boldsymbol{V}\cdot\hat{\boldsymbol{R}}, \tag{B16}$$

where $\hat{\boldsymbol{R}}$ is the unit vector pointing to the satellite's position. The final redshift can be calculated using

$$z_{\rm los} = z_{\rm cos} + \frac{v_{\rm los}}{c}(1 + z_{\rm cos}), \tag{B17}$$
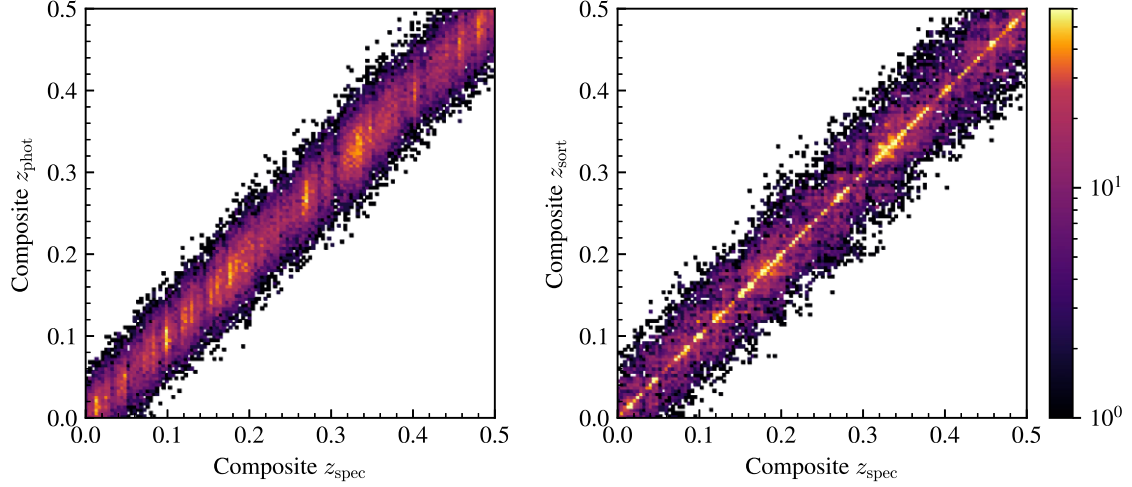
where $c$ is the speed of light.

**Figure A2.** Two-dimensional redshift histograms for $z_{phot}$ and $z_{sort}$ relative to $z_{spec}$ with binning of 0.004 using the mock CANDELS light cone. The colour bar shows the total number of counts in each bin. The data represent the full catalogue of redshifts broken into the three complete redshift bins of size $\Delta z = 0.5$ that have been stacked on top of each other. In doing so, we are able to observe similar improvement in redshift estimates to the wide-field light cone after applying SORT.
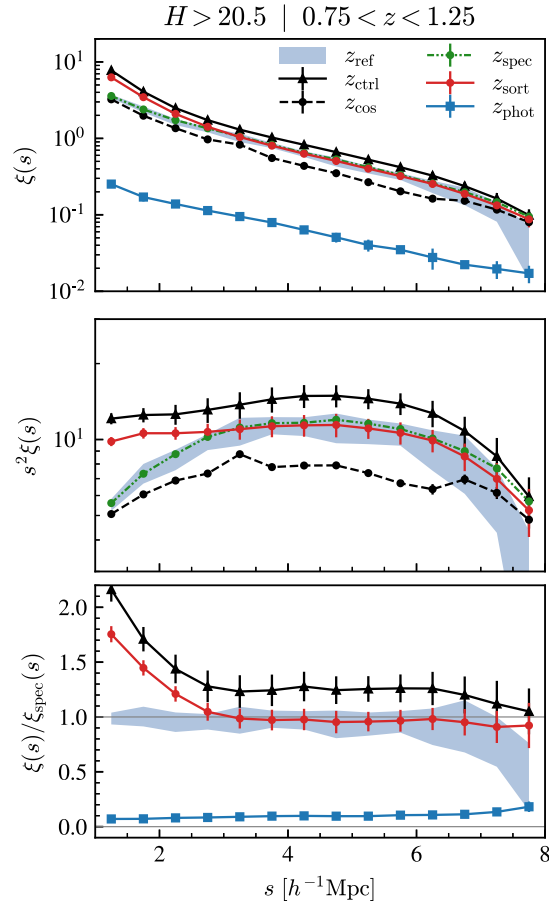


**Figure A3.** 2PCFs of various redshift types as a function of redshift-space distance $s$ shown in three different ways using the mock CANDELS light cone. In each panel, the values plotted represent the mean result of running SORT with 10 different random seeds to average out sample variance when selecting the reference galaxies. We continue to see that $\xi_{phot}(s)$ is a poor estimate of the 2PCF and $z_{ctrl}$ overestimates the 2PCF while $\xi_{sort}(s)$ is accurate (relative to $\xi_{spec}(s)$) for $s \gtrsim 2.5\ h^{-1}$Mpc. See Section 5.1 for details on $z_{ctrl}$ (shown as black triangles).
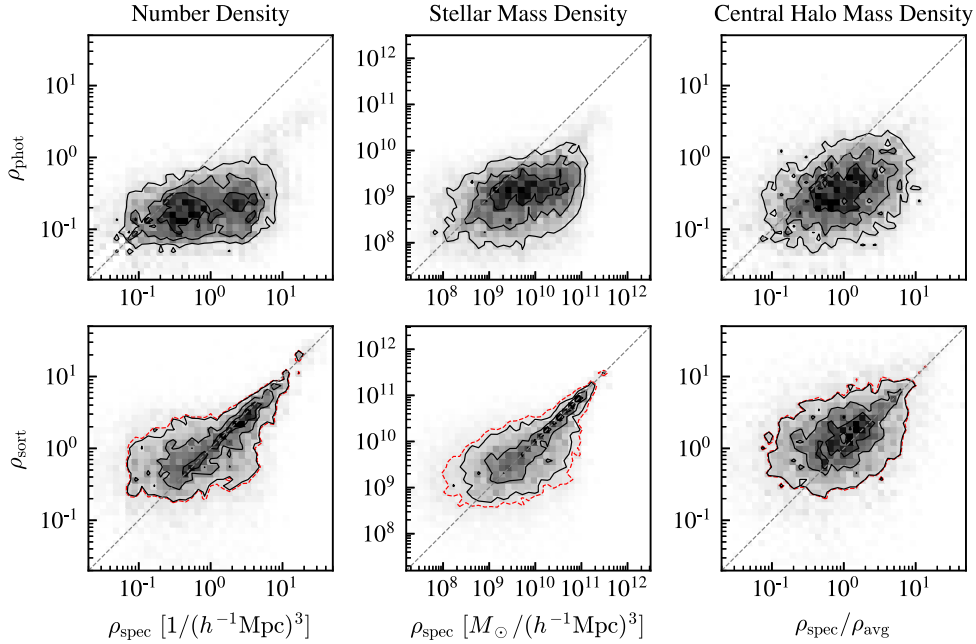
**Figure A4.** Two-dimensional density histograms for $\rho_{\rm phot}$ and $\rho_{\rm sort}$ in the range $0.75 < z < 1.25$ using the mock CANDELS light cone. The left-hand panels show number densities, the middle panels show stellar mass densities, and the right-hand panels show halo mass densities using only central galaxies. The solid contours represent limits of 25, 50, and 75 per cent of the maximum bin value in each subplot. The dashed contour (red) is set at a limit equal to the minimum contour level in the corresponding $\rho_{\rm phot}$ subplot. Densities were calculated with a fixed cylinder length of $4\,h^{-1}{\rm Mpc}$ and a radius check this starting at $\sim 0.5\,h^{-1}{\rm Mpc}$ and expanding up to $\sim 2\,h^{-1}{\rm Mpc}$ as needed to encompass at least five galaxies. As with the wide-field light cone, we continue to see improvement in density estimates by SORT compared to photometric estimates. See Section 4.3 for details on the central halo mass densities.

## APPENDIX C: SORT PERFORMANCE WITH LARGER PHOTOMETRIC UNCERTAINTIES

Our fiducial photometric uncertainty is somewhat optimistic at $\sigma_z^{\rm ph}/(1 + z) = 0.01$, though not entirely unrealistic as future redshift estimates are expected to have photometric uncertainties of $\sigma_z^{\rm ph}/(1 + z) \approx 0.02$ or better. Nevertheless, here we present brief results of SORT for larger photometric uncertainties.

We reiterate that stochastic ordering holds true for Gaussian PDFs with arbitrarily-large standard deviations. We should therefore expect to see similar redshift improvement when increasing the photometric uncertainty. The results for $\Delta z$ are shown in Fig. C1 with the fiducial results in black and the results with higher photometric uncertainties in red and blue. The histograms have been normalized by the photometric uncertainty to show the *relative* performance of SORT as $\sigma_z^{\rm ph}$ is increased. To deal with biases from sample variance in $z_{\rm ref}$, the histograms show the collection of SORT results using 10 different random seeds. We observe that SORT's improvement of redshifts with respect to a given photometric uncertainty remains largely unchanged as $\sigma_z^{\rm ph}$ increases. In all three cases we see the same general features: (i) a similar overall standard deviation in $\Delta z_{\rm sort}$ and $\Delta z_{\rm phot}$, (ii) a modest increase in scatter at the tail ends of $\Delta z_{\rm sort}$ compared to $\Delta z_{\rm phot}$, and (iii) a tall central peak of improved redshifts.

The two most notable differences are an increase in asymmetry in the tails of the histograms and a decrease in peak width as $\sigma_z^{\rm ph}$ increases. The asymmetry of the tails will be mostly irrelevant to the net result of SORT as the counts are around two orders of magnitude lower than the peak that dominates the distribution. Though we do not test this here, the width of the peak is likely more relevant to the final results of SORT. However, even with a photometric uncertainty of $\sigma_z^{\rm ph}/(1 + z) = 0.02$, SORT is still able to fairly well recover the
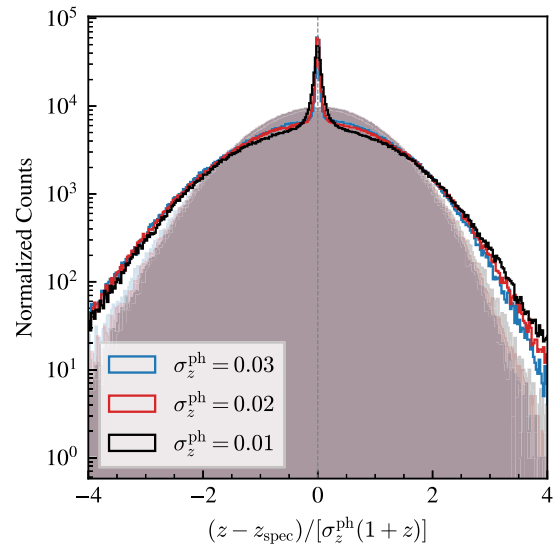


**Figure C1.** Normalized distributions of $\Delta z$ (excluding the spectroscopic sample) for $z_{\rm sort}$ and $z_{\rm phot}$ using three different photometric uncertainties. The distributions are normalized by their respective photometric uncertainties. We observe that the *relative* improvement of redshifts by SORT with respect to a given $\sigma_z^{\rm ph}$ is generally independent of $\sigma_z^{\rm ph}$. There are some dissimilarities, however. In particular, asymmetry in the tails of the distributions grows with $\sigma_z^{\rm ph}$, and the width of the peak decreases with increasing $\sigma_z^{\rm ph}$.

2PCF at similar scales of $s \gtrsim 2.5\,h^{-1}{\rm Mpc}$ (shown in the bottom panel of Fig. 7).

```
import matplotlib.pyplot as plt
import numpy as np

x = np.linspace(0, 10, 50000)
y = x + np.random.normal(0, 1, len(x))

fig, ax = plt.subplots(dpi=100)
ax.scatter(x, y, s=1, rasterized=True)
ax.plot([0, 10], [0, 10], color='k')
fig.savefig('figure.pdf')
```

**Figure D1.** Sample code for producing figures with individually rasterized elements.

## APPENDIX D: NOTE ON MAKING FIGURES WITH MANY POINTS FOR PUBLICATION

This paper features several figures that contain of order thousands of points (e.g. Fig 2). One challenge when dealing with figures such as these is keeping them clear and vectorized while simultaneously managing the size of the file. One option is to rasterize the entire figure by saving it as a jpg or png. However, this sacrifices the vectorization of the axes and labels, which are not the cause of file size issues. The result is an overall blurry figure when inspected closely. We provide in Fig. D1 some example code (using MATLAB in PYTHON) that demonstrates a better solution the reader may find useful.

Each plotted element on the figure can individually be rasterized while still maintaining vectorized axes and labels (or other plotted elements, e.g. lines, histograms, etc). The dpi for the rasterized elements can be controlled when the figure is created. In this particular example, rasterizing the scattered points lowers the file size by over an order of magnitude.

## APPENDIX E: ADDITIONAL FIGURES

Here we provide additional supplementary figures that support the main text. Fig. E1 shows two-dimensional histograms of the redshift errors as a function of the chosen redshift – either $z_{phot}$ or $z_{sort}$. Fig. E2 shows redshift errors for SORT with varying spectroscopic fractions. Fig. E3 shows results of the 2PCF in each of the three complete redshift bins. Fig. E4 shows another square region of space (similar to Fig. 3) using different redshift types. Fig. E5 shows two-dimensional redshift histograms for each of the three complete redshift bins. Fig. E6 shows two-dimensional histograms that correspond to the errors of densities shown in Fig. 8. Fig. E7 shows stellar mass densities in each of the three complete redshift bins using cylinders of length $4\ h^{-1}$Mpc. Fig. E8 shows densities similar to Fig. 8 but calculated at a larger length-scale of $l = 2(1000\ \mathrm{km\ s^{-1}}/c)(1 + z)$. Fig. E9 show two-dimensional histograms of 3D distances to $k$th nearest neighbours for $k = 3$, 5, and 7.
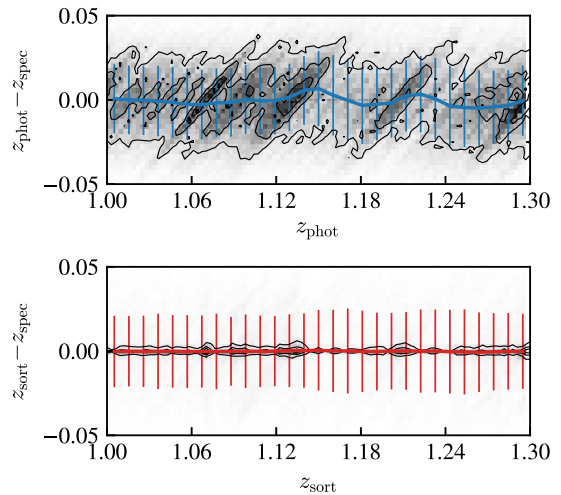


**Figure E1.** Normalized two-dimensional histograms for errors in $z_{phot}$ and $z_{sort}$ relative to $z_{spec}$. The contours show the limits where counts are at least 25, 50, or 75 per cent of the maximum value in each of the two subplots. While only the range $z = 1$–1.3 is shown, the results are representative of the entire light cone. There is a clear bias in the error of $z_{phot}$ in regions of higher density. This bias is shown as a blue line that designates the median value of all redshifts within a series of bins along with $1\sigma$ error bars. As with Fig. 5, both $\Delta z_{phot}$ and $\Delta z_{sort}$ have similar standard deviations. After applying SORT, though, the error bias is almost completely removed for the entire redshift range of the light cone.
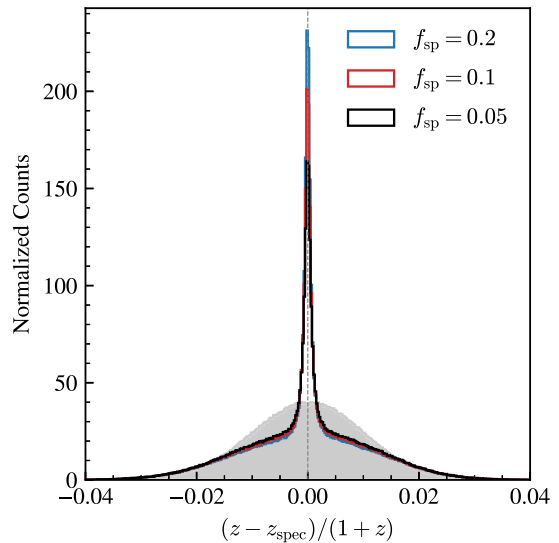


**Figure E2.** Normalized distribution of $\Delta z$ (excluding the spectroscopic sample) for $z_{phot}$ and $z_{sort}$ using three different spectroscopic fractions. As the spectroscopic fraction increases, SORT produces a taller peak surrounding $\Delta z = 0$. Even with a spectroscopic fraction as low as 5 per cent, SORT still improves redshift estimates for a significant fraction of galaxies. The efficiency of SORT is rooted in the fact that most galaxies will tend to occupy a relatively small volume. Therefore, it only takes a relatively small fraction of galaxies to reasonably trace the underlying distribution.
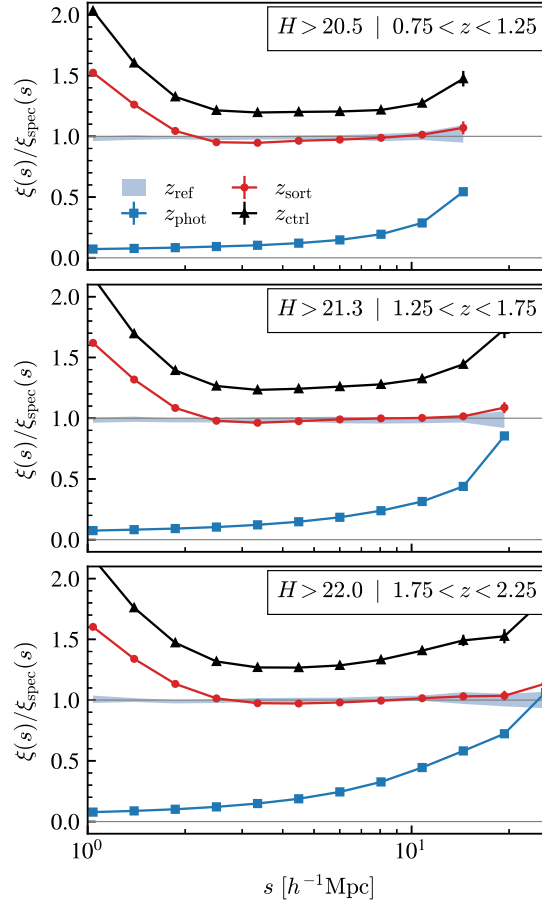
**Figure E3.** 2PCF ratios using $z_{ref}$, $z_{phot}$, $z_{sort}$, and $z_{ctrl}$ with respect to $z_{spec}$ as a function of redshift-space distance $s$ in three complete redshift bins. The 2PCFs were calculated out to distances of $\sim$18–30 $h^{-1}$Mpc, limited by the sizes of each redshift bin. The results show the mean value of the 2PCFs along with $1\sigma$ error bars after running SORT with 10 different random seeds to determine the reference sample selection. Note that the error bars are too small to be seen. The 2PCF estimates provided by $\xi_{sort}(s)$ show significant improvement over $\xi_{phot}(s)$ and accurately recover $\xi_{spec}(s)$ at scales of $s \gtrsim 2.5$ $h^{-1}$Mpc. We also observe the continued trend of $\xi_{ctrl}(s)$ overestimating the 2PCF at all scales relative to $\xi_{sort}(s)$ (see Section 5.1 for details).
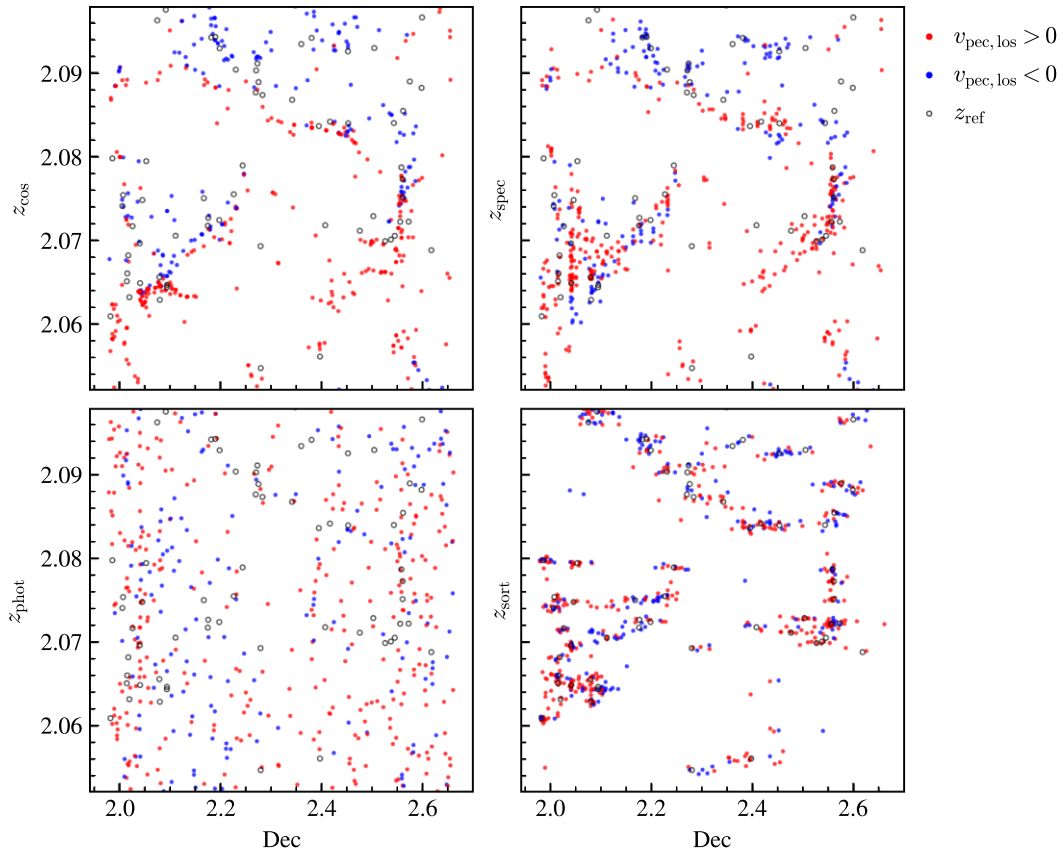
**Figure E4.** Right Ascension slices (thickness 0°.1) of galaxy distributions using different redshifts in a roughly $44 \times 44 \, h^{-1}$Mpc region of space. The red and blue colouring denotes the direction of the peculiar velocity along the line of sight (red is positive and blue is negative). The black rings with empty centres are reference galaxies. Using the outline of the reference galaxies, SORT is able to recover the distinctive features in this region – in particular, the large filamentary structure across the top and right side of the panels, as well as the more dense group of galaxies in the lower left-hand panel. We also note the presence of a characteristic feature found in SORT galaxy distributions – namely, horizontal rows of galaxies where there are few reference galaxies. In these areas, the radii of the sub-volumes within which SORT searches must expand to find reference galaxies. Galaxies are then pulled along the line of sight to an incorrect redshift, creating elongated features in a plane perpendicular to the line of sight. See Section 5.2 for details.
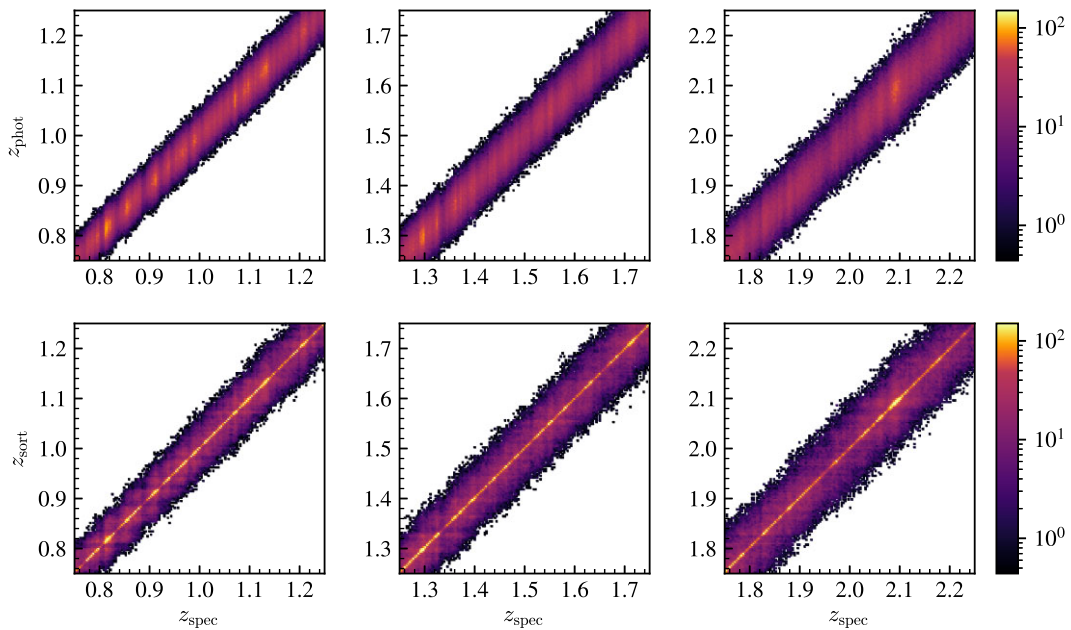


**Figure E5.** Normalized two-dimensional redshift histograms for $z_{phot}$ and $z_{sort}$ compared to $z_{spec}$ in all redshift bins. The $z_{sort}$ distributions show significant improvement as counts build up along the line of equality while the overall scatter for larger redshift errors is only modestly increased. This effect is consistent across all redshift ranges.
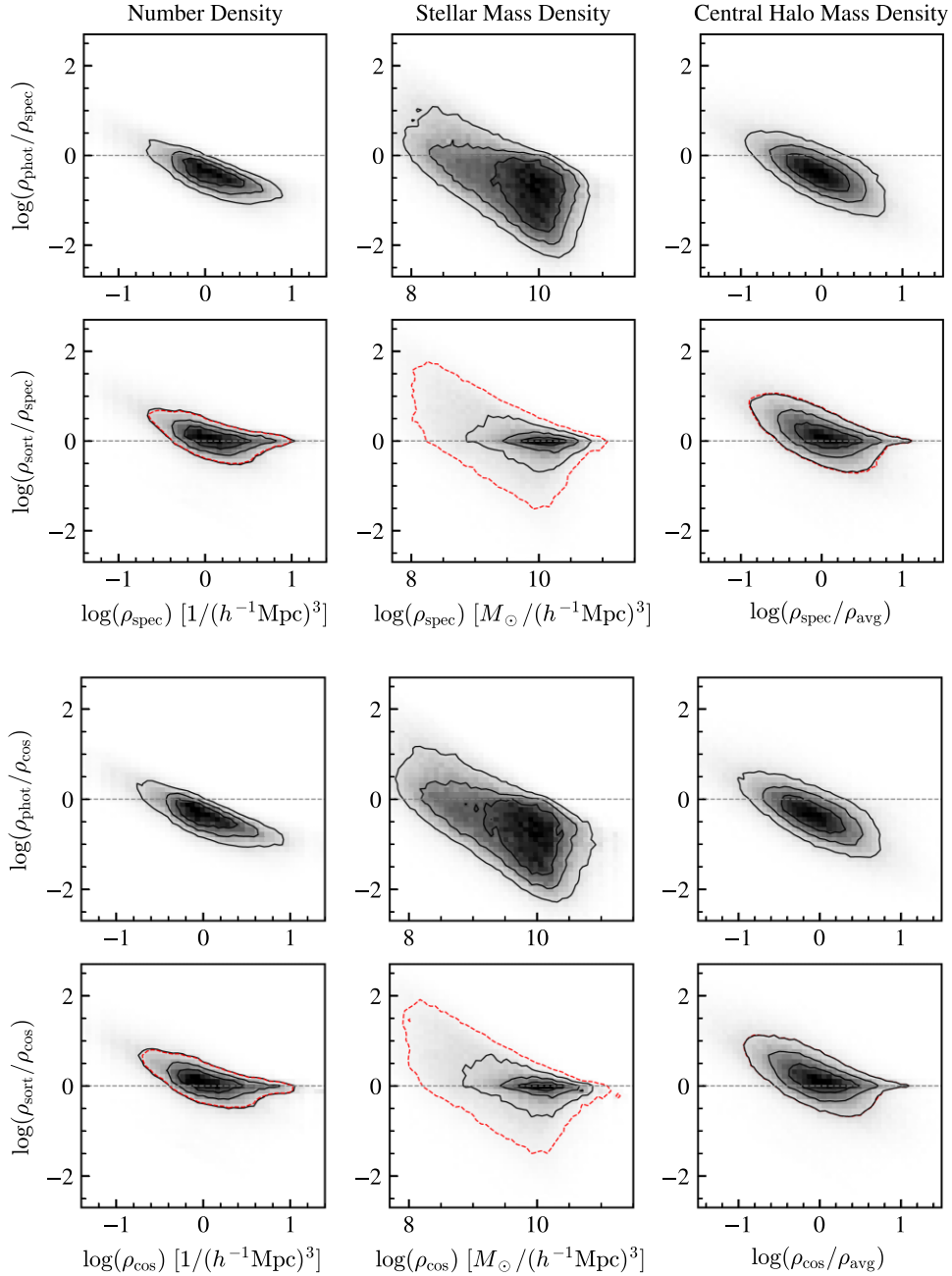
**Figure E6.** Two-dimensional density error histograms for $\rho_{\text{phot}}$ and $\rho_{\text{sort}}$. Densities were calculated in cylinders of length 4 $h^{-1}$Mpc. The left-hand panels show number density, the middle panels show stellar mass densities, and the right-hand panels show halo mass densities using only central galaxies. The solid contours represent limits of 25, 50, and 75 per cent of the maximum bin value in each subplot. The dashed contour (red) is set at a limit equal to the minimum contour level in the corresponding $\rho_{\text{phot}}$ subplot. The horizontal dashed line represents zero error. While SORT struggles with lower densities, we observe much improvement from the highest densities down to average densities. SORT distributions show better alignment with the zero error line while photometric densities all tend to be underestimated except in the lowest density environments.
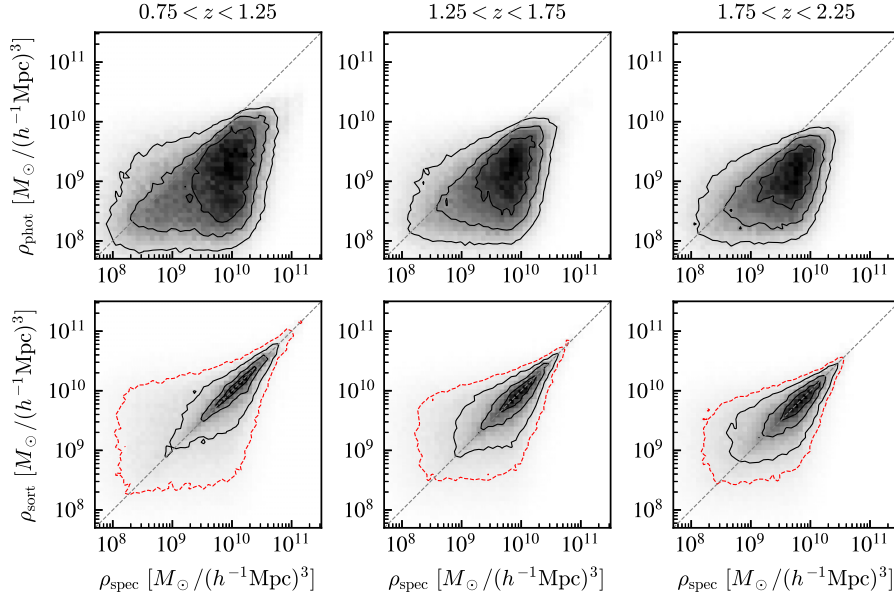
**Figure E7.** Two-dimensional stellar mass density histograms for $\rho_{\rm phot}$ (top panels) and $\rho_{\rm sort}$ (bottom panels) for all redshift ranges of the light cone using a cylinder length of 4 $h^{-1}$Mpc. The solid contours represent limits of 25, 50, and 75 per cent of the maximum bin value in each subplot. The dashed contour (red) is set at a limit equal to the minimum contour level in the corresponding $\rho_{\rm phot}$ subplot. We observe consistent improvement in density estimates with SORT at all redshifts. Biases in regions of average or higher density are greatly reduced. SORT distributions are more symmetric across the line of equality and the overall scatter is lower.
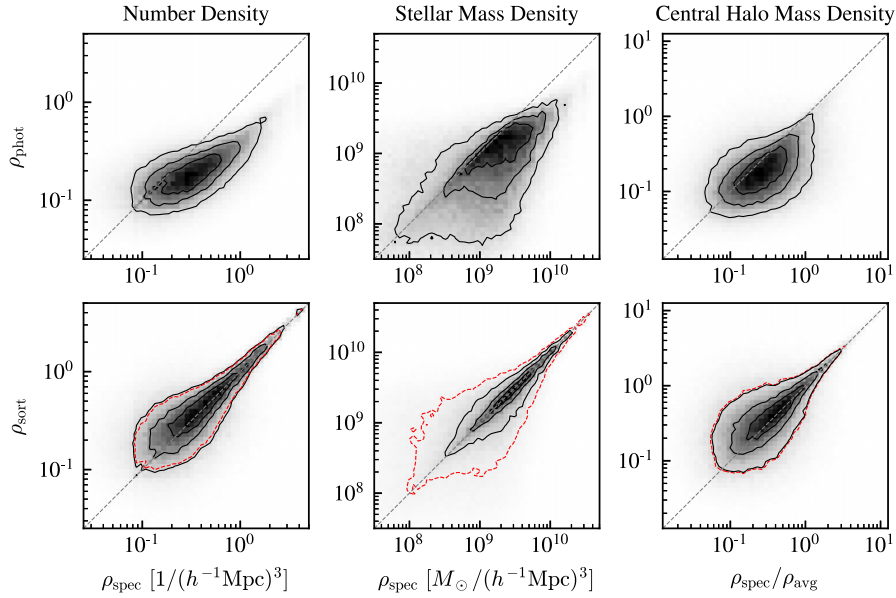


**Figure E8.** Two-dimensional density histograms for $\rho_{\rm phot}$ (top panels) and $\rho_{\rm sort}$ (bottom panels) in the range $0.75 < z < 1.25$. Densities were calculated within cylinders of length $l = 2(1000 \ {\rm km \ s}^{-1}/c)(1 + z)$. The left-hand panels show number density, the middle panels show stellar mass densities, and the right-hand panels show halo mass densities using only central galaxies. The solid contours represent limits of 25, 50, and 75 per cent of the maximum bin value in each subplot. The dashed contour (red) is set at a limit equal to the minimum contour level in the corresponding $\rho_{\rm phot}$ subplot. The longer length of the cylinder significantly improves SORT density estimates, most notably for stellar mass densities.

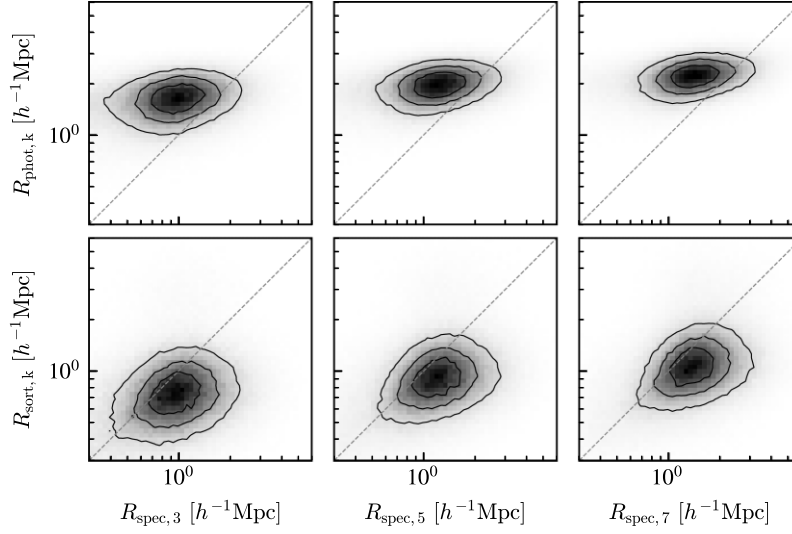**Figure E9.** Two-dimensional histograms of 3D distances to $k$th nearest neighbours using $z_{phot}$ (top panels) and $z_{sort}$ (bottom panels) compared to $z_{spec}$ for $k = 3$, 5, and 7. The overall scatter is slightly increased using $z_{sort}$, but alignment with the line of equality is improved, particularly at smaller scales. At larger scales (corresponding to lower densities), SORT underestimates $R_k$ as it packs galaxies too closely together in low-density environments. See Section 5.4 for details.

This paper has been typeset from a TEX/LATEX file prepared by the author.