The Astrophysical Journal, 825:40 (25pp), 2016 July 1 doi:10.3847/0004-637X/825/1/40

CAN WE DETECT THE COLOR–DENSITY RELATION WITH PHOTOMETRIC REDSHIFTS?

Chuan-Chin Lai[1,2], Lihwai Lin[2], Hung-Yu Jian[2,3], Tzi-Hong Chiueh[1,3], Alex Merson[4], Carlton M. Baugh[5], Sebastien Foucaud[6], Chin-Wei Chen[2], and Wen-Ping Chen[7]

[1] Graduate Institute of Astrophysics, National Taiwan University, Taipei, 10617, Taiwan (R.O.C.); chuanchinlai@gmail.com
[2] Institute of Astronomy & Astrophysics, Academia Sinica, Taipei 10617, Taiwan (R.O.C.)
[3] Department of Physics, National Taiwan University, Taipei, 10617, Taiwan (R.O.C.)
[4] Department of Physics and Astronomy, University College London, Gower Street, London WC1E 6BT, UK
[5] Institute for Computational Cosmology, Department of Physics, Durham University, South Road, Durham DH1 3LE, UK
[6] Center for Astronomy and Astrophysics, Shanghai Jiao Tong University, Shanghai 200240, China
[7] Graduate Institute of Astronomy, National Central University, Chung-Li 32054, Taiwan (R.O.C.)

Received 2015 January 6; revised 2016 April 10; accepted 2016 May 1; published 2016 June 27

ABSTRACT

A variety of methods have been proposed to define and to quantify galaxy environments. While these techniques work well in general with spectroscopic redshift samples, their application to photometric redshift surveys remains uncertain. To investigate whether galaxy environments can be robustly measured with photo-z samples, we quantify how the density measured with the nearest-neighbor approach is affected by photo-z uncertainties by using the Durham mock galaxy catalogs in which the 3D real-space environments and the properties of galaxies are known exactly. Furthermore, we present an optimization scheme in the choice of parameters used in the 2D projected measurements that yield the tightest correlation with respect to the 3D real-space environments. By adopting the optimized parameters in the density measurements, we show that the correlation between the 2D projected optimized density and the real-space density can still be revealed, and the color–density relation is also visible out to $z \sim 0.8$ even for a photo-z uncertainty ($\sigma_{\Delta_z/(1+z)}$) up to 0.06. We find that at redshifts $0.3 < z < 0.5$ a deep ($i \sim 25$) photometric redshift survey with $\sigma_{\Delta_z/(1+z)} = 0.02$ yields a performance in small-scale density measurement that is comparable to a shallower $i \sim 22.5$ spectroscopic sample with ~10% sampling rate. Finally, we discuss the application of the local density measurements to the Pan-STARRS1 Medium Deep Survey (PS-MDS), one of the largest deep optical imaging surveys. Using data from ~5 square degrees of survey area, our results show that it is possible to measure local density and to probe the color–density relation with 3σ confidence level out to $z \sim 0.8$ in the PS-MDS. The color–density relation, however, quickly degrades for data covering smaller areas.

Key words: galaxies: evolution – galaxies: formation – galaxies: halos

1. INTRODUCTION

Recent observations have shown that various galaxy properties such as star formation rate, color, and morphology are strongly correlated with galaxy environment (Hogg et al. 2004; Blanton et al. 2006; Cooper et al. 2006, 2007, 2012; Haines et al. 2006; Capak et al. 2007; Elbaz et al. 2007; Gallazzi et al. 2009; Mostek et al. 2013; Darvish et al. 2014; Lin et al. 2014). These studies indicate that galaxies located in dense environments, such as galaxy groups and clusters, tend to be redder, elliptical, and with lower star formation rates. Several physical processes, including ram pressure stripping (Gunn & Gott 1972; Quilis et al. 2000), high-speed galaxy encounters (galaxy harassment; Moore et al. 1996), galaxy–galaxy mergers (Mihos & Hernquist 1994), and removal of warm and hot gas (strangulation; Larson et al. 1980; Balogh et al. 2000; McCarthy et al. 2008) have been proposed to explain the observed relation between environment and galaxy properties. Yet exactly how the environment affects the evolution of galaxies and how important it is in relation to the internal properties of galaxies (e.g., stellar mass) is still unclear. Part of the discrepancy between previous studies may come from the differences in sample selection as well as the definition of environment, which makes the comparisons non-trivial.

One of the common approaches to characterizing galaxy environments is to use the local overdensity of matter. For the rest of this paper, we use the observed overdensity of galaxies as a proxy for the galaxy environment. A variety of methods have been used to define the density field of a galaxy, for example: (1) the Fixed Aperture method, which counts the number of neighboring galaxies in a fixed volume around each galaxy (e.g., Gallazzi et al. 2009; Grützbauch et al. 2011); (2) the Annulus method, which counts the number of neighboring galaxies within a circular ring around each galaxy (e.g., Wilman et al. 2010); and (3) the Nth nearest neighbor, which defines the local density by finding the distance from the individual reference galaxies to the Nth nearest galaxy (e.g., Casertano & Hut 1985; Gómez et al. 2003; Baldry et al. 2006; Haas et al. 2012). A fundamental and crucial quantity for these methods to work is the redshift, which provides the information about the line-of-sight separation (in the absence of peculiar velocities) of two given galaxies. Observationally, there are two types of redshift that are used widely: spectroscopic redshifts and photometric redshifts (hereafter spectral-z and photo-z respectively). While the spectral-z samples have greater precision in the redshift measurement, they suffer from incompleteness and are observationally expensive for high-redshift galaxies.

To date, studies of environments using large spectroscopic surveys such as SDSS (York et al. 2000), DEEP2 (Davis et al. 2003; Newman et al. 2013), and zCOSMOS (Lilly et al. 2007) have been limited to redshifts lower than $z \sim 1.5$ (see Tanaka et al. 2004; Cooper et al. 2006, 2007; Haines et al. 2006; Mostek et al. 2013). In contrast, photo-z surveys provide

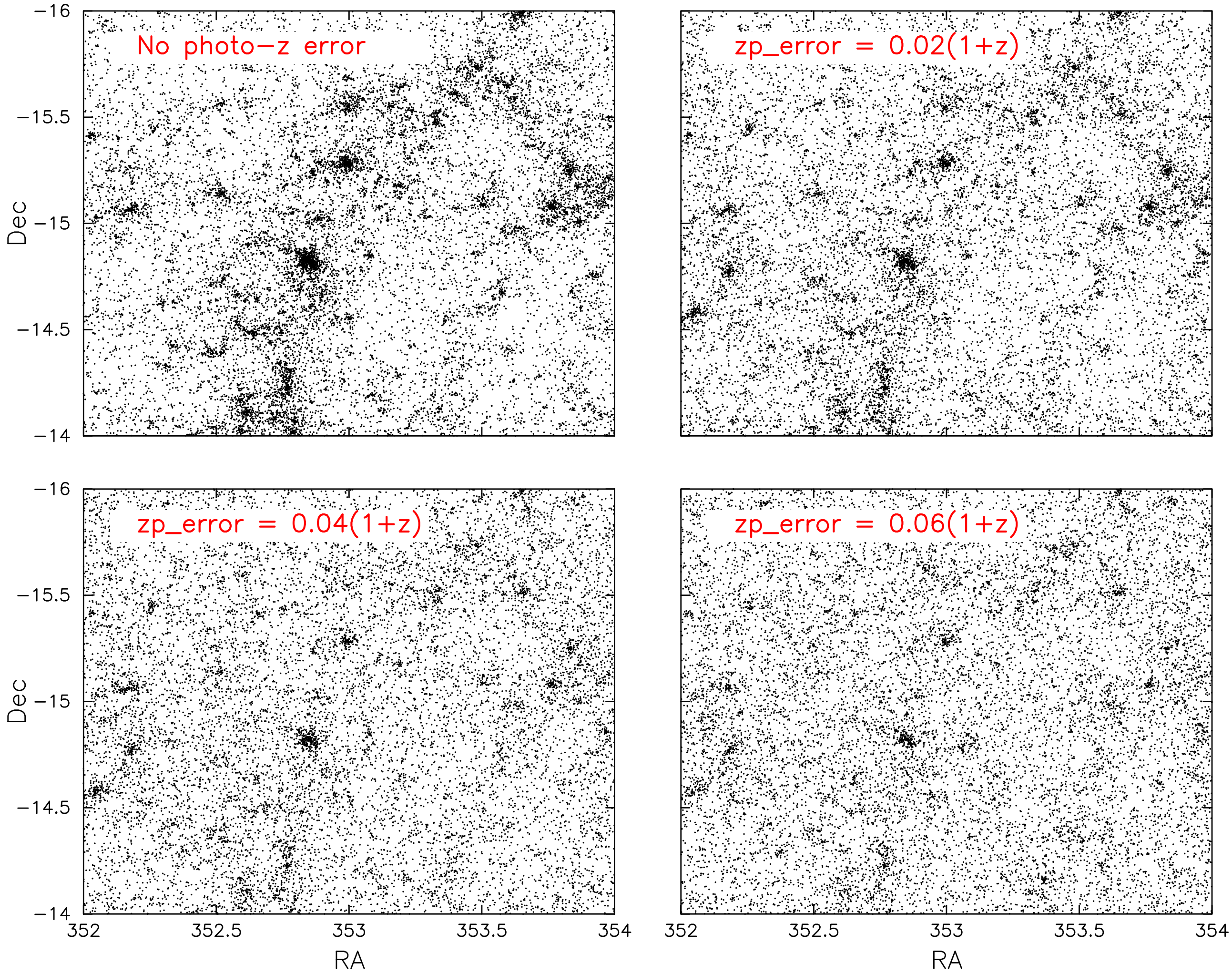

Figure 1. Spatial distribution of mock galaxies projected onto the plane of the sky with redshifts perturbed corresponding to different photo-z errors: 0.00, $0.02(1 + z)$, $0.04(1 + z)$, and $0.06(1 + z)$ at $0.3 < z_{\rm photo} < 0.35$.

larger sample sizes and reach higher redshifts, but they suffer from poorer redshift resolution. Figures 1–3 show how the photo-z uncertainties distort the real galaxy environment from different viewpoints. The large-scale structures are clearly revealed in the case without photo-z error but become less prominent as the photo-z error increases. Despite this problem, there have been some attempts to measure the galaxy environment for various studies using photo-z samples (Capak et al. 2007; Quadri et al. 2012; Scoville et al. 2013; Chiang et al. 2014; Lin et al. 2016). Several works have provided viable methods that can be used to recover the density fields of galaxies from photo-z samples (Etherington & Thomas 2015; Lin et al. 2016; Malavasi et al. 2016). Moreover, Arnalte-Mur et al. (2009) and Schlagenhaufer et al. (2012) both demonstrated that the two-point correlation function of galaxies can also be successfully recovered from photometric samples, and they also discussed the influence of photo-z errors on their measurements.

Several ongoing large sky surveys such as the Panoramic Survey Telescope and Rapid Response System (Pan-STARRS: Onaka et al. 2008; Kaiser et al. 2010), Dark Energy Survey (The Dark Energy Survey Collaboration 2005; Albrecht et al. 2006), Hyper Suprime-Cam Survey (Miyazaki et al. 2012), and the upcoming Large Synoptic Survey Telescope (LSST Science Collaboration et al. 2009) will yield large galaxy samples with photometric redshift measurements, so it is important to understand the potential and limitations of the photo-z method in the studies of galaxy evolution, especially the environmental effects. Can we use photo-z samples to measure environment reliably? What are the systematics in the environment measurement between spectral-z samples and photo-z samples? What is the optimal choice for density measurement that can reliably recover the underlying environments? These are the questions that we aim to answer. In particular, we focus on the measurement of the density field of a galaxy. We first study the difference between 3D real-space density and 2D projected density measurements by using mock galaxy catalogs. We adopt Spearman's rank correlation coefficient (Spearman 1904), r_s, as a measure of the correlation between the 2D and the 3D real-space density. The optimized parameters for the density measurement are obtained by maximizing r_s. We then use the results of optimized density measurements to show the dependence of galaxy properties on environments from the mock catalog. Finally, we apply our optimized scheme to the Pan-STARRS1 data and compare the results with the measurements by Cooper et al. (2006), who use the DEEP2 spectroscopic sample in the same field.

This paper is structured as follows. In Sections 2.1 and 2.2 we describe the simulation and observational data used in our

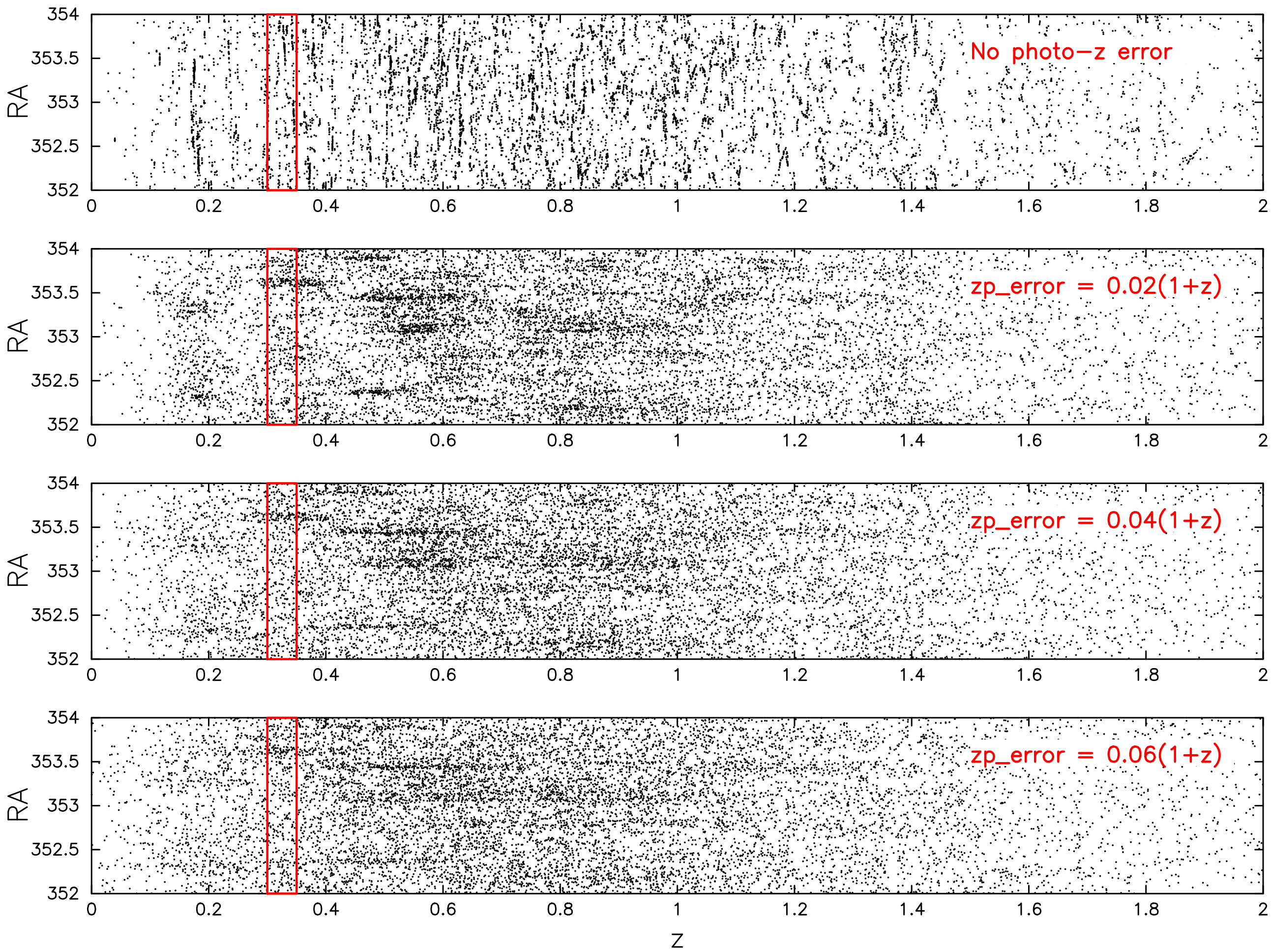

Figure 2. Spatial distribution of mock galaxies seen in one line-of-sight projection and redshift. A 0°.05 interval in decl. is used when projecting galaxies onto the plane. The redshifts of galaxies are perturbed according to different photo-z errors: 0.00, $0.02(1+z)$, $0.04(1+z)$, and $0.06(1+z)$. The red rectangle indicates the redshift range $0.3 < z_{\rm photo} < 0.35$ used in Figure 1.

study. The environment measurements used in this study are introduced in Section 3. In Section 4 we compare the 3D real-space density with the 2D projected environment, and demonstrate how to optimize the choice of Nth nearest neighbor to improve the 2D projected density measurement. In Section 5 we show the relation between galaxy environment and galaxy properties in the mock galaxy catalog to verify whether or not our optimized scheme is applicable. We discuss several possible factors that might limit our optimized scheme and apply it to observations in Section 6. Finally we summarize our results in Section 7. In this paper, we adopt the following cosmological parameters: $H_0 = 100h$ km s^{-1} Mpc^{-1}, $h = 0.73$, $\Omega_0 = 0.25$, and $\Omega_\Lambda = 0.75$.

2. DATA

2.1. Simulation Data

In this work, we use a theoretical mock galaxy catalog to understand the systematics in the local density estimates. The advantage of using a mock galaxy catalog compared to real spectroscopic survey data is that the real-space density can be directly measured and compared with the projected density. Moreover, the mock sample does not suffer from the incompleteness that often affects real observations. On the other hand, one needs to be cautious when interpreting the results since the properties of galaxies in the simulation may not be a perfect representation of the real universe.

The mock galaxy catalog used in this work is built based on the Millennium simulation with $N = 2160^3$ in a box with volume $= 500^3\ h^{-3}$ Mpc^3 from redshift $z = 127$ to the present day at $z = 0$ by adopting the following cosmological parameters: a baryon matter density $\Omega_b = 0.045$, a total matter density $\Omega_0 = 0.25$, a dark energy density $\Omega_\Lambda = 0.75$, and a Hubble constant $H_0 = 100h$ km s^{-1} Mpc^{-1} where $h = 0.73$. These cosmological parameters match the first-year results of the *Wilkinson Microwave Anisotropy Probe* (Spergel et al. 2003). Galaxies are put into halos using the GALFORM semi-analytical model (Cole et al. 2000), which takes into account various galaxy formation processes including gas accretion and cooling, star formation in galactic disks, and galaxy mergers. The mock catalog adopts the model of Lagos et al. (2012), which takes advantage of the extension to the treatment of star formation introduced into GALFORM in Lagos et al. (2011) to populate galaxies, and is then assembled into a lightcone (Merson et al. 2013). Further detailed information is given in Cole et al. (2000), Springel et al. (2005), Bower et al. (2006), Lagos et al. (2011, 2012), and Merson et al. (2013).

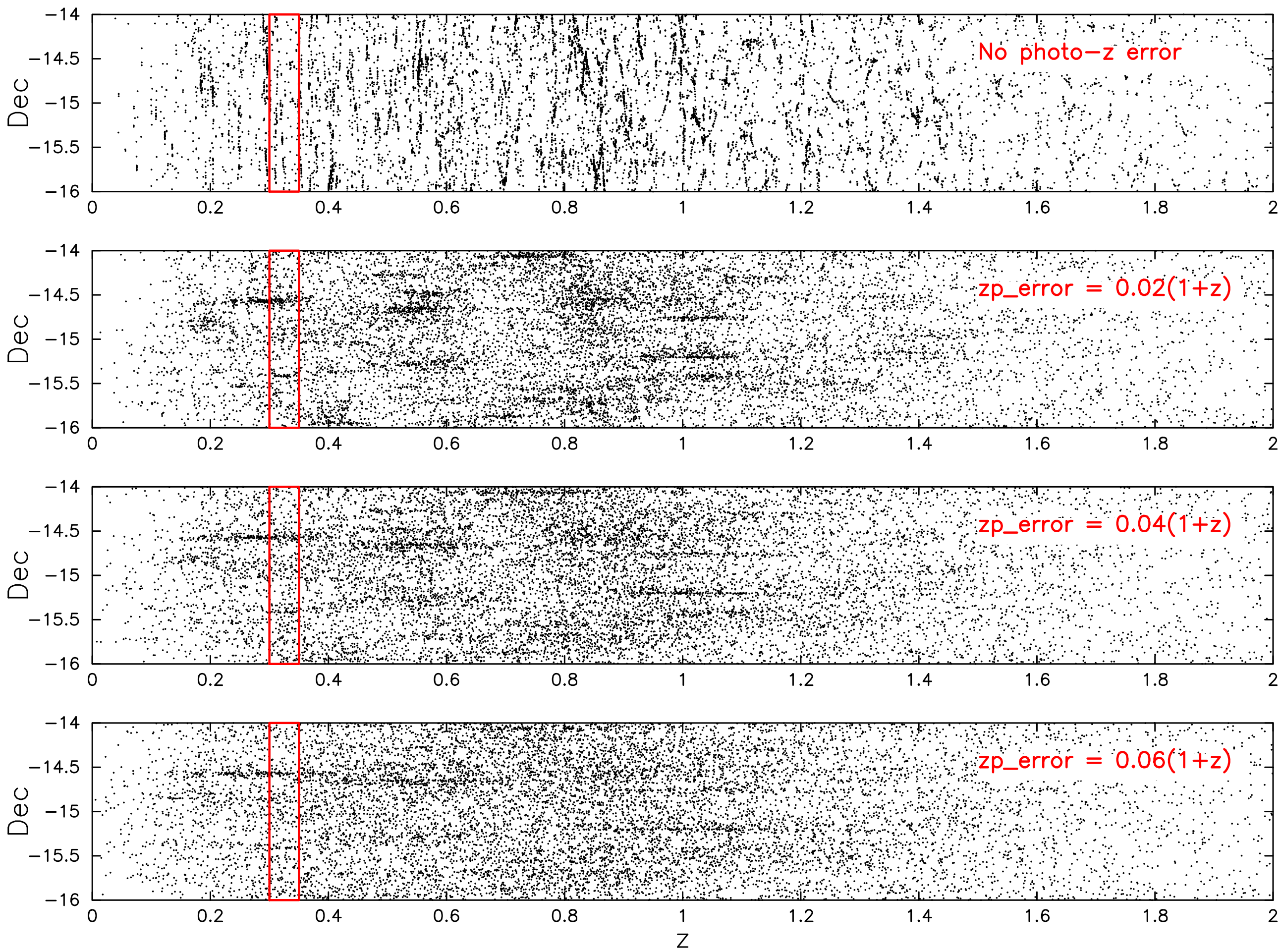

Figure 3. Spatial distribution of mock galaxies seen in one line-of-sight projection and redshift. A 0°.05 interval in R.A. is used when projecting galaxies onto the plane. The redshifts of galaxies are perturbed according to different photo-z errors: 0.00, $0.02(1 + z)$, $0.04(1 + z)$, and $0.06(1 + z)$. The red rectangle indicates the redshift range $0.3 < z_{\rm photo} < 0.35$ used in Figure 1.

We constructed two types of mock catalogs that mimic the observed spectral-z and photo-z catalogs. The mock spectral-z catalog can be obtained from primitive simulation data, which store the intrinsic line-of-sight positions of galaxies. To generate mock photo-z catalogs, we perturb the position of galaxies along the line-of-sight direction by making a random shift that follows a Gaussian distribution with a standard deviation that matches the photo-z error in each case, in order to simulate cases with observed redshift uncertainties. The new redshift obtained can be viewed as the "observed redshift," and used to compute the local density. Although the photometric redshift model adopted here is oversimplified because it does not take into account the effect of catastrophic redshift failures, this simplistic model allows us to understand the effect of redshift dispersion. Later, in Section 5.4, we consider more realistic situations in which the outlier effect is included. In this study we consider several photo-z cases with uncertainties of $\sigma_{\Delta_z/(1+z)} = 0.00$, 0.02, 0.04, and 0.06. We restrict our environment study to the redshift range $0.3 < z < 0.5$ for most of our analysis. A central area of the catalogs of ~16 square degrees is selected for our studies, containing ~1,900,000 galaxies with $i < 25.8$. It is worth noting that the redshift that we use for computing the 3D overdensity with the Millennium mock spectroscopic catalogs refers to the intrinsic redshift of galaxies, which does not include the effect from peculiar velocity. Therefore, the results shown for the spectroscopic redshift sample may be too optimistic. However, since our main focus is to understand the performance of density recovery in the case of photometric errors, this mock spectroscopic catalog does provide a "real" answer for the 3D density.

2.2. Observation Data

2.2.1. Spectroscopic Observation

The DEEP2 Galaxy Survey (Davis et al. 2003; Newman et al. 2013) was designed to study the galaxy population and large-scale structure at $z \sim 1$. It uses the Keck II telescopes with the DEIMOS spectrograph (Faber et al. 2003), covers ~3.5 square degrees of the sky with measured spectra, and has targeted ~60,000 galaxies down to a limiting magnitude of $R_{\rm AB} < 24.1$. About ~60% of the galaxies are sampled over the redshift interval $0.2 < z < 1.4$. The overall redshift success rate is about ~70%. DEEP2 comprises four widely separated fields. One of the DEEP2 fields, the Extended Groth Strip (EGS), is enclosed by the Pan-STARRS1 Medium Deep Survey Field (MD07). In this study we match the DEEP2 spectroscopic redshift catalog to the Pan-STARRS1 MD07 catalog. For galaxies that are common to the two catalogs, we compare the local density measurements computed using the DEEP2 spectral-z and Pan-STARRS1 photo-z respectively (see Section 6).

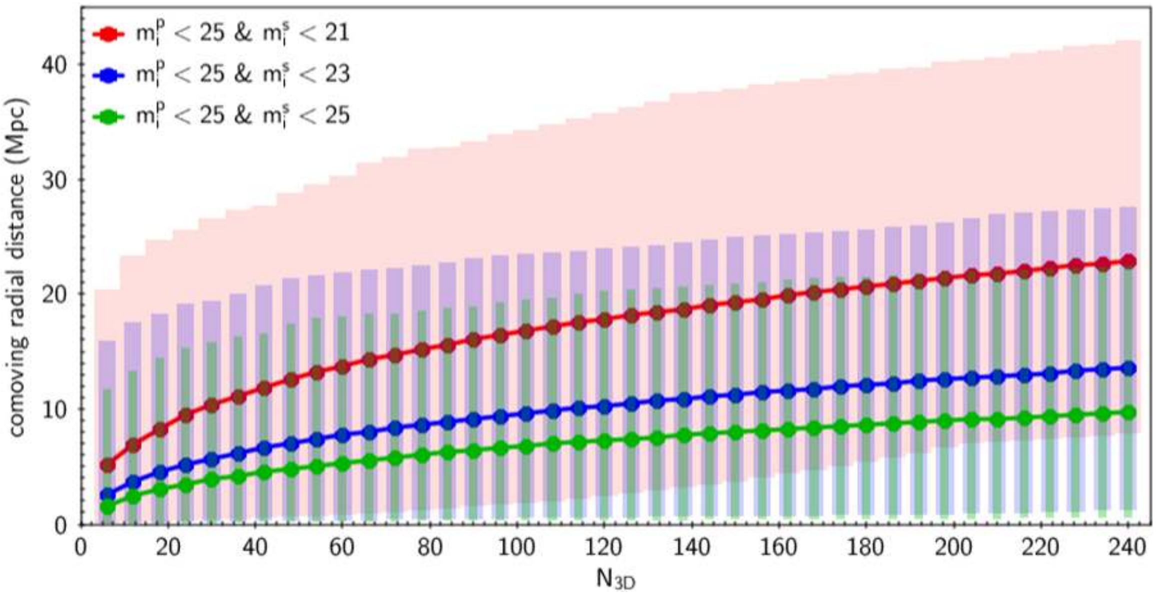

Figure 4. The median *N*th nearest-neighbor distance as a function of N_{3D} for various choices of the magnitude limit of the secondary sample (m_i^s). The areas shaded in different colors show the range between the minimum and maximum *N*th nearest-neighbor distances.

2.2.2. Photometric Observation

Pan-STARRS1 (hereafter PS1) is a 1.8 m telescope equipped with a CCD digital camera with 1.4 billion pixels and 3° field of view, located on the summit of Haleakala on Maui, Hawaii (Onaka et al. 2008; Kaiser et al. 2010). The PS1 observations are obtained in a set of five broadband filters, which we have designated as g_{p1}, r_{p1}, i_{p1}, z_{p1}, and y_{p1}. There are two major components of the PS1 survey, which started observations in 2010: the 3π survey and the Medium Deep Survey (MDS), which comprises ten fields spread across the sky. One of the MDS fields, namely MD07, is chosen for this study because it overlaps with the EGS field, which has the spectroscopic data from the DEEP2 survey, and enables a direct comparison with the environment measurements using the DEEP2 spectral-*z* sample (Cooper et al. 2006). Photo-*z* redshifts in MD07 are computed by running the EAZY code (Brammer et al. 2008) on PS1 five-band photometry plus the u^*-band data taken by Eugene Magnier et al. with CFHT MEGACAM as part of the PS1 efforts. Comparisons against the DEEP2 spectroscopic redshifts (Newman et al. 2013) show that the PS1 photo-*z* reaches an uncertainty of 0.05 with an outlier rate of 7% down to $r_{p1} < 24.1$. More details on the data processing and photo-*z* characteristics in the PS1 MD07 sample are given in Lin et al. (2014) and S. Foucaud et al. (2016, in preparation).

3. MEASUREMENTS OF GALAXY ENVIRONMENT

In this study, we adopt the *N*th nearest-neighbor method to quantify galaxy environment. This method defines the local density of each galaxy using the distance to the *N*th nearest-neighbor galaxy. In other words, whether the galaxy is located in an overdense or underdense environment depends on how far it is from its *N*th nearest galaxy. In simulations, the cosmological redshift reflects the "real" distance along the line of sight, enabling the environment measurement to be evaluated by using the 3D *N*th nearest-neighbor method. On the other hand, observationally, the local density is often estimated by using a projected method because the measured redshift is a combination of both the cosmological distance and the peculiar velocity.

Here we define two sets of galaxy samples: the primary and the secondary. The primary sample contains galaxies brighter than a particular magnitude (m_i^p), and is used when presenting the results. The secondary sample refers to galaxies used in the search for neighbors, and is restricted to those galaxies that are brighter than a particular magnitude (m_i^s). The limiting magnitude of the secondary sample is particularly important because it sets the galaxy number densities in the calculation of density field. Figure 4 shows the median distances to the 3D *N*th nearest neighbor with various choices of m_i^s. The colored areas show the range between the minimum and maximum distances to the 3D *N*th nearest neighbor. There are several parameters that should be considered in the *N*th nearest-neighbor method: (1) the choice of the *N*th neighbor, which represents the scale of the environment, (2) the magnitude limit of the primary sample (m_i^p), (3) the magnitude limit of the secondary sample (m_i^s), and (4) the velocity window (V_{cut}) that defines the redshift boundaries of the neighbors considered in the 2D projected method. These parameters should be adjusted according to different science goals and galaxy samples. One of the goals of this work is to provide an empirical framework that determines these parameters by calculating galaxy densities with different combinations of parameters in order to understand their influence.

3.1. 2D Projected Nth Nearest-neighbor Galaxy Environment

In the 2D projected method, the local density of each galaxy is computed as the surface density averaged over the area enclosed by the *N*th closest galaxy within the velocity interval V_{cut}:

$$\Sigma_n = \frac{n+1}{\pi r_n^2}, \quad (1)$$

where n is the *N*th closest galaxy for each reference galaxy and r_n is the distance from the reference galaxy to the *N*th closest galaxy on a 2D surface. There is no simple way to determine the choice of velocity window $V_{\rm cut}$ in the 2D nearest-neighbor method. In principle, it is not meaningful to adopt a $V_{\rm cut}$ that is too small compared to the redshift uncertainty of the data. Conversely, adopting a large velocity cut enlarges the projection effect, which leads to greater errors in the density measurement. For instance, two galaxies that are close in the projected plane may actually be widely separated in the third dimension, and vice versa (Muldrew et al. 2012). Previous studies have utilized the velocity interval $V_{\rm cut}$ of a value close to the distance uncertainties in the line-of-sight direction, because they found that the density estimate does not vary significantly when changing $V_{\rm cut}$ around this value. We will further test this approach using the mock catalog in Section 4.1.

3.2. 3D Nth Nearest-neighbor Galaxy Environment

The galaxy environment defined by the 3D *N*th nearest-neighbor method is similar to that defined by the 2D projected *N*th nearest-neighbor method except that the projected circular area is replaced by the enclosed spherical volume. The volume density of galaxies is evaluated using the 3D *N*th nearest-neighbor method as

$$\rho_n = \frac{n+1}{(4/3)\pi r_n^3}, \tag{2}$$

where n is the *N*th closest galaxy of each reference galaxy and r_n is the distance from the reference galaxy to the *N*th closest galaxy in the three-dimensional space. We compute the real-space density using the 3D *N*th nearest-neighbor method of the simulation where information about the three-dimensional positions of galaxies is known. We treat the 3D density as the "true" density to be compared with the 2D density to quantify how well the real-space density can be recovered by the 2D projected density under various conditions. We note that in practice the 3D density is rarely used, even in a spectroscopic redshift sample, because the observed "redshift" includes contributions from both the Hubble flow and the peculiar velocity of galaxies, which it is not possible to differentiate observationally.

Finally, in order to contrast the most dense environments with the least dense environments, we convert the initial primordial density into an overdensity. The overdensity is conventionally defined as the initial primordial density divided by the median density as follows:

$$1 + \delta_n = \frac{D_i}{D_{Mdn}}, \tag{3}$$

where D_i is the measured density of a galaxy (i.e., $D_i \in \{\rho_n, \Sigma_n\}$), and D_{Mdn} is the median density computed by counting galaxies within a bin of $\Delta z = 0.04$. The term $1 + \delta_n$ is the so-called overdensity, and δ_n can be $\delta_n^{\rm 3D}$ or $\delta_n^{\rm 2D}$, depending on whether $D_i = \rho_n$ or Σ_n.

4. QUANTIFYING DIFFERENCES BETWEEN ENVIRONMENT MEASUREMENTS

In this section, we compare the 2D projected density obtained under various conditions to the 3D real-space density. To quantify their differences, we adopt Spearman's rank correlation coefficient, r_s, which is commonly used to measure the strength of a relationship between two ranked variables (Spearman 1904; Curran 2014). Spearman's rank correlation coefficient is defined as

$$r_s = 1 - \frac{6\sum d_i^2}{s(s^2-1)}, \tag{4}$$

where d_i is the difference between the ranks of the two variables and s is the sample size. A coefficient $r_s = 0$ corresponds to no correlation between two variables, while $r_s = 1$ (-1) corresponds to a perfect-positive (perfect-negative) correlation. In our analysis, we first measure the 2D projected density Σ_n as well as the 3D real-space density ρ_n for each galaxy, and then convert all the density measurements into an overdensity as defined in the previous section. After ranking the 3D real-space overdensity and 2D projected overdensity, we compute the difference d_i between the ranks of the two overdensities, and then use Equation (4) to calculate Spearman's rank correlation coefficient r_s.

Figure 5 is an example showing the correlation between the 3D and 2D measurements using $N_{\rm 2D} = 6$, 30, 60, and 90, and how the r_s coefficient changes with different choices of $N_{\rm 2D}$.

4.1. 3D Galaxy Environment Versus 2D Projected Galaxy Environment with Different Parameters

We first probe the effect of the velocity interval $V_{\rm cut}$ on the 2D projected density measurement. We restrict the sample to $m_i^p < 25$ and $m_i^s < 25$ when calculating 2D projected overdensity, $1 + \delta_6^{\rm 2D}$, and 3D real-space overdensity, $1 + \delta_6^{\rm 3D}$. Figure 6 shows the scatter plots of the 2D projected overdensity versus the real-space overdensity on a log scale using the sixth-nearest neighbor. Here we consider the following four different choices of $V_{\rm cut}$ in the 2D projected measurement for a galaxy sample with photo-z error $= 0.04(1 + z)$: $\pm 0.005(1 + z)$, $\pm 0.02(1 + z)$, $\pm 0.04(1 + z)$, and $\pm 0.06(1 + z)$. It can be seen that the difference in r_s among the four cases of $V_{\rm cut}$ is not significant when the size of velocity interval is close to the photo-z error. For example, r_s of 0.434 is obtained in the case of $V_{\rm cut} = \pm 0.02(1 + z)$ (upper-right panel of Figure 6), while the value of r_s increases to 0.473 in the case of $V_{\rm cut} = \pm 0.06(1 + z)$ (lower-right panel of Figure 6). The difference in r_s is small ($\sim$0.039) between these two cases even though $V_{\rm cut}$ differs by a factor of 3, which confirms the finding in previous studies that the 2D projected measurement is not sensitive to $V_{\rm cut}$ when $V_{\rm cut}$ is comparable to the photo-z uncertainty (Gallazzi et al. 2009; Muldrew et al. 2012). We further repeat similar exercises using samples with a larger redshift uncertainty up to $0.08(1 + z)$ and at higher redshifts ($0.6 < z < 0.8$), and we find that this conclusion still holds. Therefore throughout this work we set $V_{\rm cut}$ to be the typical photo-z uncertainty of the galaxy sample.

Next we consider the effect of the secondary magnitude limit employed on the galaxy sample when searching for neighbors. A brighter (fainter) m_i^s probes a larger (smaller) scale of environment for a fixed *N*th nearest neighbor. It is therefore expected that the correlation between the 2D projected and 3D real-space environments could depend on the choice of m_i^s. Again we select the sample with photo-z error $= 0.02(1 + z)$ and set $V_{\rm cut} = \pm 0.02(1 + z)$ for the reason given above. Figure 7 shows the scatter plots of the 2D projected overdensity versus real-space overdensity using the sixth-nearest neighbor,

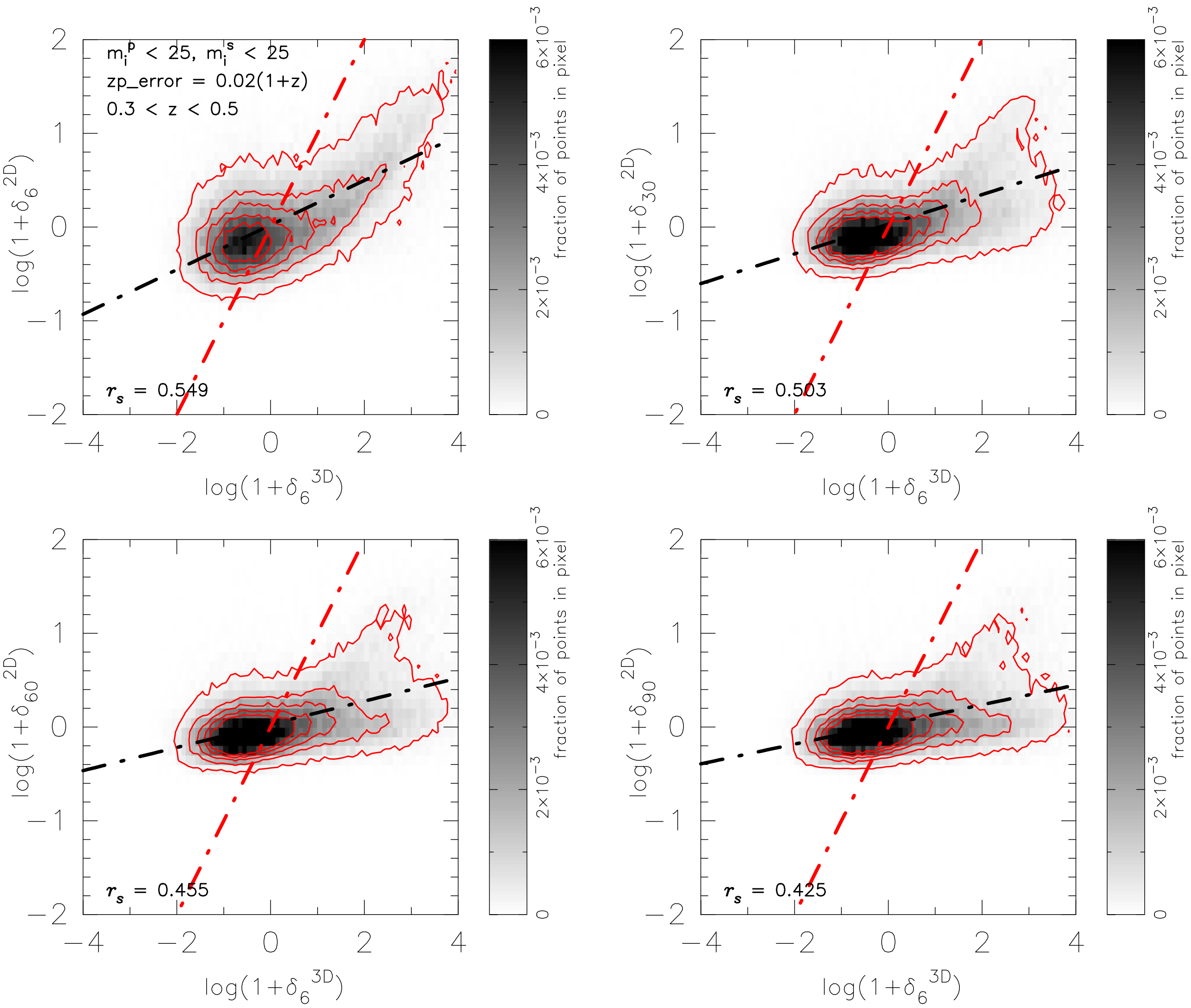

Figure 5. Scatter plot of the 3D real-space overdensity, $1 + \delta_6^{3D}$, vs. 2D projected overdensity: $1 + \delta_6^{2D}$ (upper-left panel), $1 + \delta_{30}^{2D}$ (upper-right panel), $1 + \delta_{60}^{2D}$ (lower-left panel), and $1 + \delta_{90}^{2D}$ (lower-right panel) on a log scale. The numbers in the bottom left of each panel indicate the r_s coefficient. The black dashed–dotted lines represent the best fit to the data points, the red dashed–dotted lines represent the one-to-one relation, and the contours show the regions of constant galaxy number.

and both 2D and 3D environments are measured using $m_i^s < 21$, $m_i^s < 23$, and $m_i^s < 25$. As can be seen, the largest r_s is obtained when m_i^s in the 2D measurement is equal to m_i^s in the 3D measurement. Furthermore, for identical m_i^s used in the 2D and 3D measurements, the environments measured by using fainter m_i^s (and hence smaller scales) have better correlation than those measured using a brighter m_i^s. This is consistent with the results from Shattow et al. (2013), which also shows that for samples with photo-z error, the 2D projected environments have a weaker correlation with 3D real-space environments on larger scales.

Finally, Figure 8 shows the 2D versus 3D scatter plots in the density measurement to understand how the galaxy environment is affected by the photo-z errors. Here we consider $m_i^p < 25$ and $m_i^s < 25$, and vary the photo-z error from 0, $0.02(1 + z)$, $0.04(1 + z)$, to $0.06(1 + z)$. V_{cut} is correspondingly set to be $\pm 0.001(1 + z)$, $\pm 0.02(1 + z)$, $\pm 0.04(1 + z)$, and $\pm 0.06(1 + z)$ respectively. It is worth noting that even in the perfect situation where the redshift error is zero, the correlation is still not perfect owing to the projection effect. The correlation between 3D real-space and 2D projected environments becomes gradually worse when the photo-z error increases. However, there still exists some correlation especially for galaxies located in high-density regions, while the environment in lower and intermediate densities is less distinguishable. This is consistent with the result from Capak et al. (2007), who also showed that the galaxy environment is difficult to measure for galaxies located in regions of low density when a redshift error is present.

4.2. Optimizing 2D Environment Parameters

So far we have compared 3D real-space environments with various 2D projected environments to show their correlation and we have adopted r_s to quantify the goodness of the correlation. We now expand this to construct an optimization scheme to determine the value of N_{2D} that gives the best correlation (largest r_s) between the 2D and 3D environments. Figure 9 shows r_s calculated by fitting 2D projected and 3D real-space environments with various choices of N_{2D} and N_{3D}.

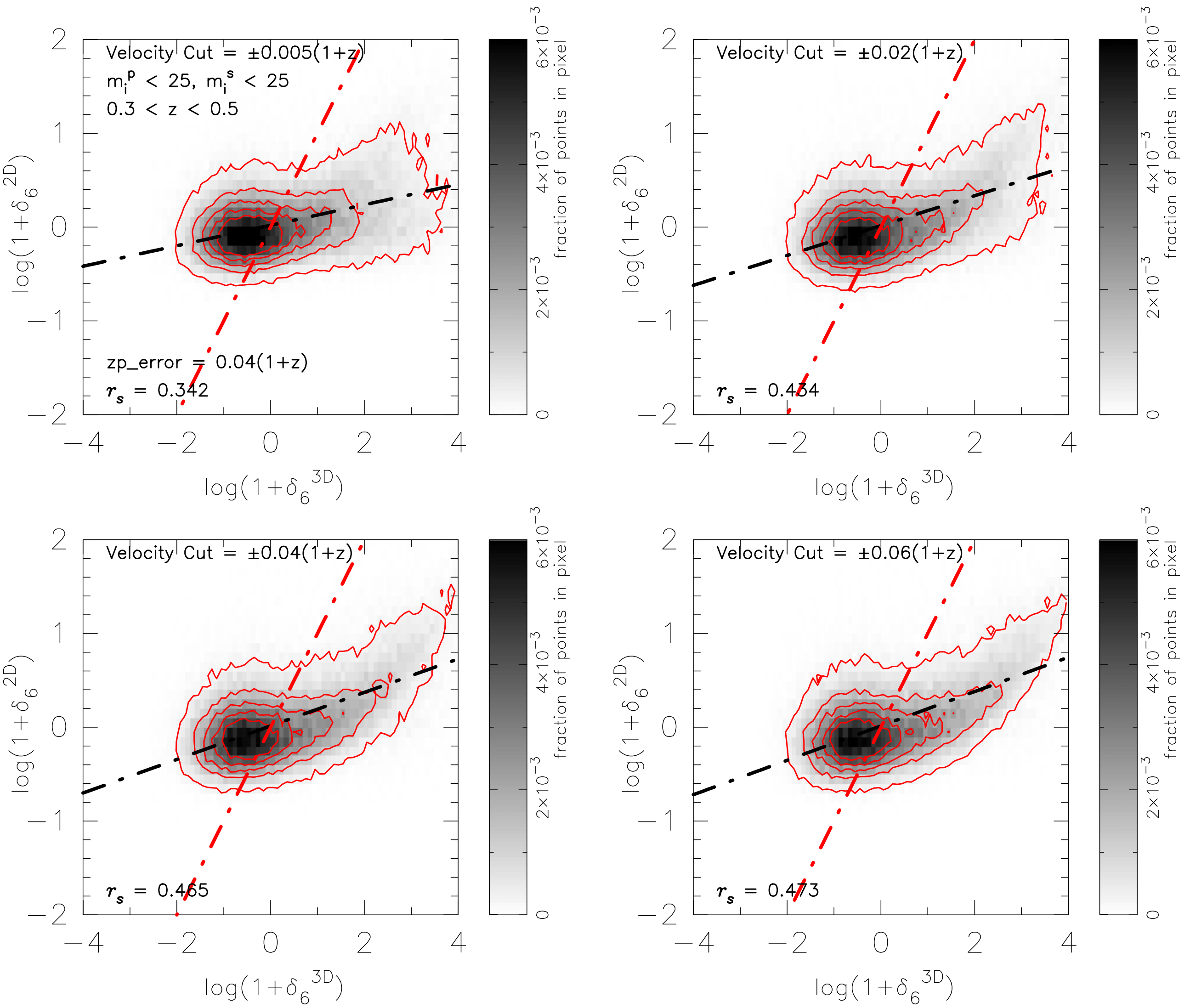

Figure 6. Scatter plot of the 3D real-space overdensity, $1 + \delta_6^{3D}$, vs. 2D projected overdensity, $1 + \delta_6^{2D}$, with various velocity cuts: $V_{cut} = \pm 0.005(1 + z)$, $\pm 0.02(1 + z)$, $\pm 0.04(1 + z)$, and $\pm 0.06(1 + z)$ over the redshift interval $0.3 < z < 0.5$. All cases are considered by using galaxy samples with photo-z error = 0.04 $(1 + z)$, $m_i^s < 25$, and $m_i^p < 25$. The numbers in the bottom left of each panel indicate the r_s coefficient. The black dashed–dotted lines represent the best fit to the data points, the red dashed–dotted lines represent the one-to-one relation, and the contours show the regions of constant galaxy number.

The red, green, blue, cyan, and magenta dots are for the cases where the 2D projected environments are calculated using $N_{2D} = 6$, 30, 60, 90, and N_{3D} respectively. We also mark the four cases of Figure 5, $r_s = 0.425$, 0.455, 0.503, and 0.549 respectively, to demonstrate how r_s varies with different choices of N_{2D}. In the following analysis, we show only the optimized choice of N_{2D} corresponding to the case that yields the largest r_s, as a function of N_{3D}.

Figure 10 shows the largest r_s (upper panel) and corresponding choice of N_{2D} (lower panel) that yields the best correlation between 2D projected and 3D real-space environments for different choices of N_{2D}, as a function of N_{3D} from mock galaxy catalogs. The red, green, blue, and cyan dots are for samples with different photo-z errors: 0.00, $0.02(1 + z)$, $0.04(1 + z)$, $0.06(1 + z)$ respectively. As expected, the densities are relatively easier to recover for samples with lower photo-z errors than for those with higher photo-z errors. The r_s obtained for the case without photo-z error (red dots) are greater than 0.9, meaning that the optimized 2D projected environments are strongly correlated with the 3D real-space environments. However, for the samples contaminated by photo-z errors (green, blue, and cyan dots), the performance of recovery becomes gradually worse as the photo-z error increases. In addition, the correlation between 2D projected and 3D real-space environments depends not only on the redshift accuracy but also on the scale. For example, at a given photo-z error, r_s decreases with N_{3D}, which suggests that small-scale environments are easier to recover. A possible explanation is that the 2D projected environments calculated by using photo-z samples might include more contaminations when we probe the larger scale of environments.

One interesting feature in the lower panel of Figure 10 is that the best choices of N_{2D} are in general equal not to N_{3D} but to only half of N_{3D}, for producing the largest r_s except for the case that is error-free. However, we note that the relation between the optimized N_{2D} and N_{3D} depends on the choices of redshift interval. Detailed discussions are given in the Appendix.

As the value of density measurements also depends on the choice of the secondary magnitude limit, it is interesting to see how the correlation between 2D and 3D measurements changes

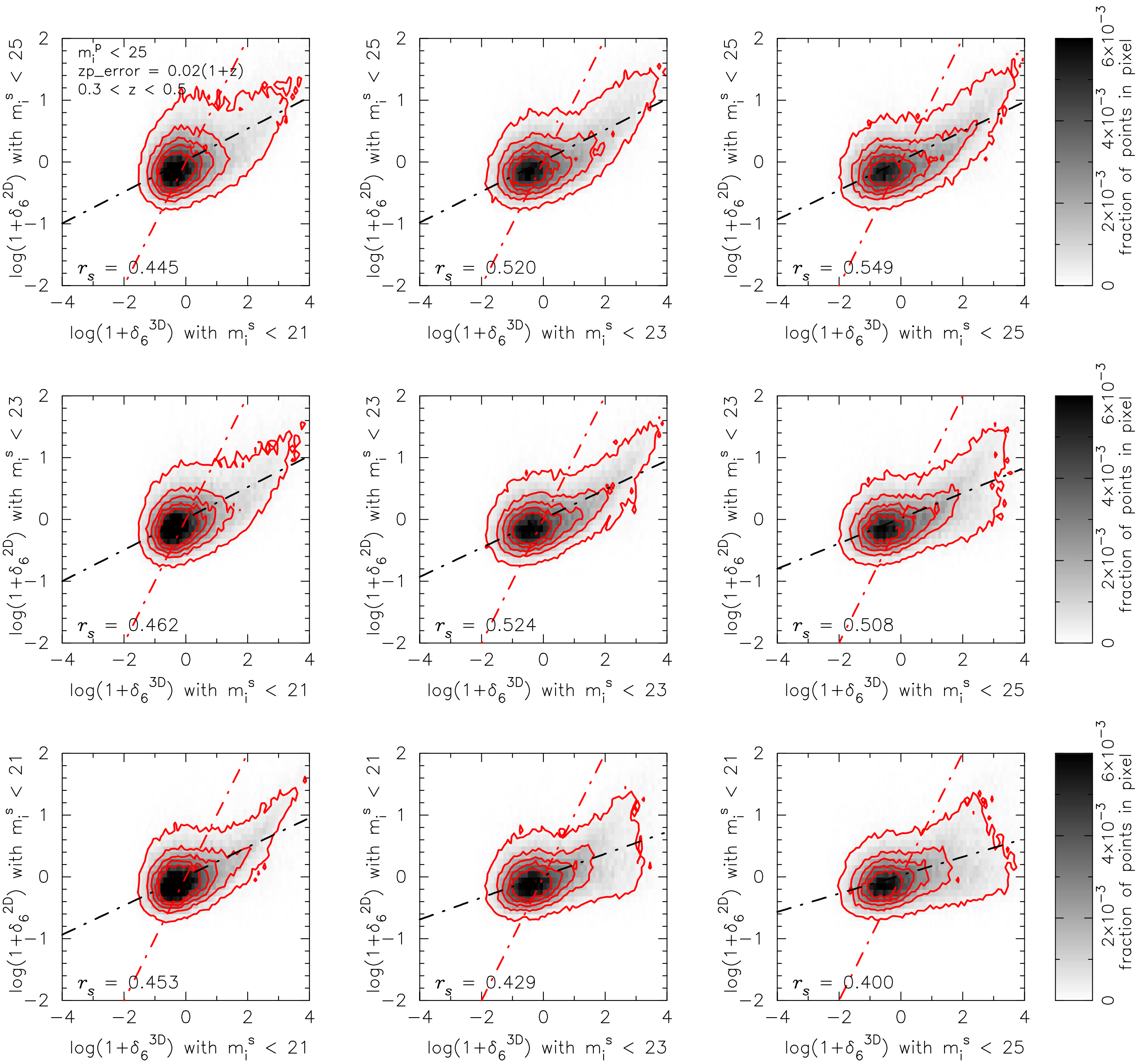

Figure 7. Scatter plot of the 3D real-space overdensity, $1 + \delta_6^{3D}$, vs. 2D projected overdensity, $1 + \delta_6^{2D}$, with various secondary magnitude limits in 2D measurements (panels in a row, from left to right: $m_i^s < 21$, $m_i^s < 23$, and $m_i^s < 25$) and 3D measurements (panels in a column, from bottom to top: $m_i^s < 21$, $m_i^s < 23$, and $m_i^s < 25$) over the redshift interval $0.3 < z < 0.5$. All cases are considered by using galaxy samples with photo-z error $= 0.02(1 + z)$, $V_{\rm cut} = \pm 0.02(1 + z)$, and $m_i^p < 25$. The numbers in the bottom left of each panel indicate the r_s coefficient. The black dashed–dotted lines represent the best fit to the data points, the red dashed–dotted lines represent the one-to-one relation, and the contours show the regions of constant galaxy number.

by varying the secondary magnitude limits in both 3D and 2D environment measurements. Figures 11–13 show the largest r_s as a function of $N_{\rm 3D}$ for different 3D secondary magnitude limits. In each figure, the red, green, and blue dots are for samples with secondary magnitude limits corresponding to $m_i^s < 21$, $m_i^s < 23$, and $m_i^s < 25$ in 2D environment measurement respectively, at a fixed secondary magnitude limit in 3D local density, and at a fixed photo-z error $= 0.02(1 + z)$. Our results show that when the 3D environment is defined using brighter secondary magnitude limits, there is no significant difference in the performance of recovery among different choices of 2D secondary magnitude limit that are fainter than the 3D secondary magnitude limit. On the other hand, adopting a 2D secondary magnitude limit that is brighter than the 3D one results in a poorer recovery of the environment. This means that a deeper sample is favored when constructing the 2D density field.

5. THE CORRELATION BETWEEN ENVIRONMENT AND GALAXY PROPERTIES

In Section 4 we optimized the choice of $N_{\rm 2D}$ for the 2D projected density measurement to yield the best correlation with the 3D real-space environments by using the r_s metric. In

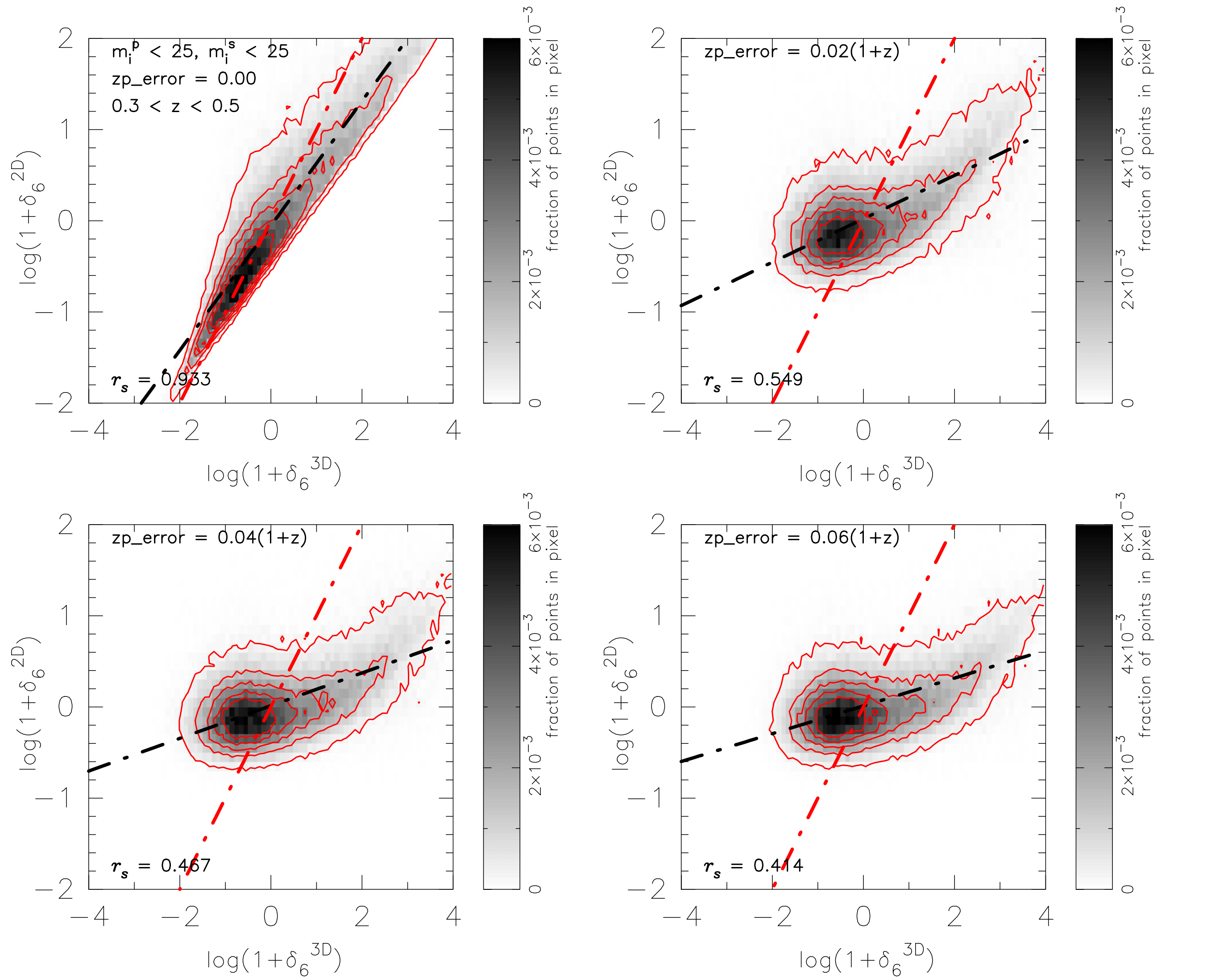

Figure 8. Scatter plot of the 3D real-space overdensity, $1 + \delta_6^{3D}$, vs. 2D projected overdensity, $1 + \delta_6^{2D}$, with different photo-z errors over the redshift interval $0.3 < z < 0.5$. All cases are considered by using galaxy samples with photo-z error = 0.00, 0.02(1 + z), 0.04(1 + z), and 0.06(1 + z) and $V_{\rm cut} = \pm 0.001(1 + z)$, $\pm 0.02(1 + z)$, $\pm 0.04(1 + z)$, and $\pm 0.06(1 + z)$ respectively. The primary and secondary magnitude limits are $m_i^p < 25$ and $m_i^s < 25$. The numbers in the bottom left of each panel indicate the r_s coefficient. The black dashed–dotted lines represent the best fit to the data points, the red dashed–dotted lines represent the one-to-one relation, and the contours show the regions of constant galaxy number.

this section, we use these optimized results to study how the color–density relation in the simulation changes when varying the photo-z uncertainties and outlier rates. Although the density–color and/or halo mass–color relations seen in the simulations may not fully represent the observed universe, this provides us with a guideline to understand how reliably we can study the dependence of galaxy properties on environment using the photo-z samples.

5.1. Environment versus Galaxy Color

To explore the relation between galaxy color and environment, we compare the apparent magnitude i versus $g - i$ colors of galaxies located in the 20% most dense and the 20% least dense galaxy environments. We note that although conventionally the color–magnitude relation is defined in the rest frame when studying the color–density relation, here we look only at the observed quantity since the redshift range is very small and our main purpose is to see whether the density dependence of color distributions can still be revealed in the photometric redshift sample, rather than quantifying the "color–density relation" itself. Galaxies are first classified to be red or blue according to their locations in the observed color–magnitude diagram (CMD). We use $g - i = 1.5$ and 1.75 as dividing lines to separate blue and red galaxies at $0.3 < z < 0.5$ and $0.6 < z < 0.8$, respectively. Next we bin the galaxies according to their i-band apparent magnitude and then compute the percentage of red galaxies defined as

$$f_{\rm red} = \frac{N_{\rm red}}{N_{\rm bin}}, \quad (5)$$

where $N_{\rm bin}$ is the total number of galaxies and $N_{\rm red}$ is the number of red galaxies in each bin.

We determine the choice of $N_{\rm 2D}$ that yields the largest r_s for photo-z samples with different photo-z uncertainties using the methodology described in Section 4.2. In the case where we study the color–density relation for the environment scale corresponding to the sixth-nearest neighbor in the 3D space,

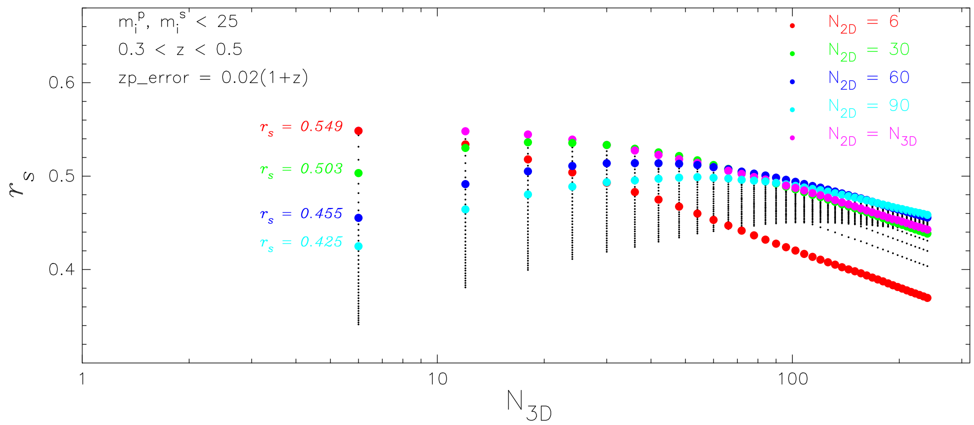

Figure 9. The black dots show the r_s coefficients calculated by fitting the real-space and 2D projected environments for different choices of N_{2D} as a function of N_{3D} from mock galaxy catalogs. The r_s for the choice of $N_{2D} = 6$, 30, 60, 90, and N_{3D} are shown as red, green, blue, cyan, and magenta dots, respectively. The values of r_s corresponding to the four cases (0.549, 0.503, 0.455, 0.425) shown in Figure 5 are also marked.

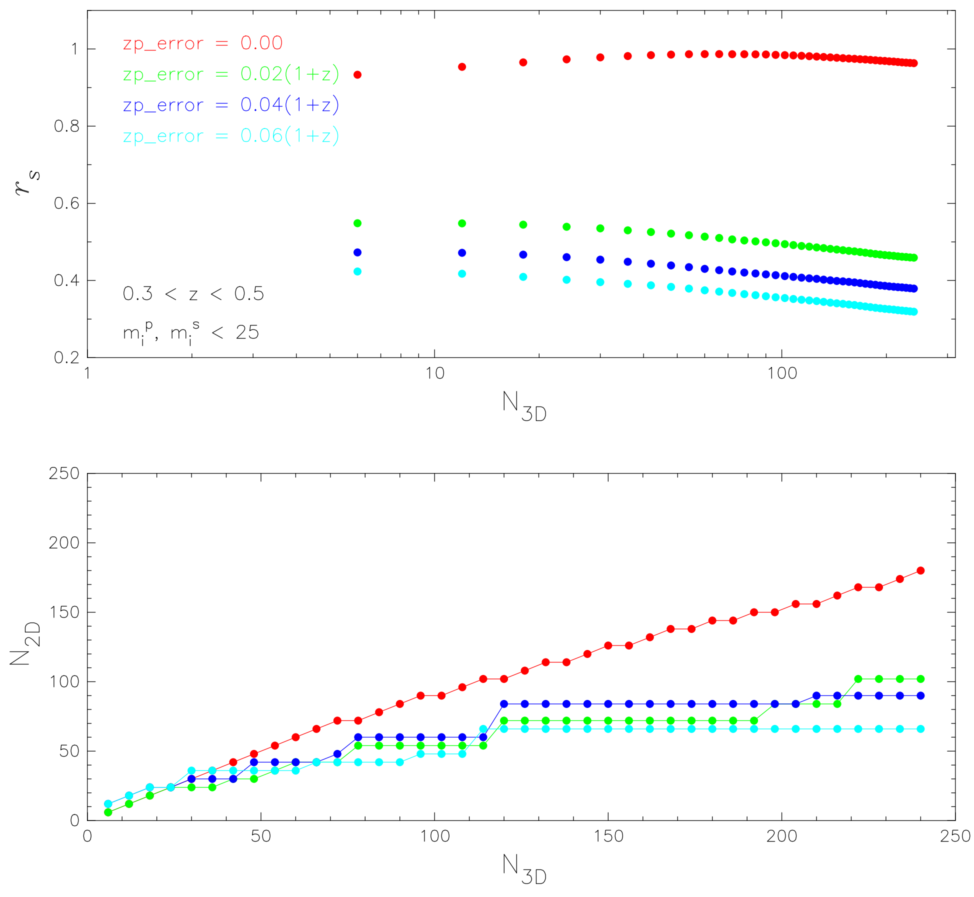

Figure 10. The largest r_s (upper panel) obtained by varying N_{2D} and the corresponding choice of N_{2D} (lower panel) that yields the best correlation between the real-space and 2D projected environments for different choices of N_{2D} as a function of N_{3D} from mock galaxy catalogs. Dots of different color correspond to samples with different photo-z errors: 0.0, $0.02(1 + z)$, $0.04(1 + z)$, and $0.06(1 + z)$.

i.e., ρ_6, it is found that the optimized $N_{2D} = 6$, 6, 6, 12 is for the cases with photo-z error = 0.00, $0.02(1 + z)$, $0.04(1 + z)$, and $0.06(1 + z)$ respectively (see the lower panel of Figure 10). We note that in the case of photo-z error = $0.06(1 + z)$, although $N_{2D} = 12$ is the best choice for optimization, we still adopt $N_{2D} = 6$ to show its CMD for convenience because there is almost no significant difference between using $N_{2D} = 6$ and $N_{2D} = 12$ for the optimization.

The upper panels of Figure 14 show the CMD for the 20% most dense environments (red contours) and the 20% least dense environments (blue contours) for galaxy samples with different photo-z errors. Here we use the 2D projected measurements Σ_6 and set $V_{cut} = \pm 0.001(1 + z)$, $\pm 0.02(1 + z)$, $\pm 0.04(1 + z)$, and $\pm 0.06(1 + z)$ for the cases with photo-z error = 0.00, $0.02(1 + z)$, $0.04(1 + z)$, and $0.06(1 + z)$ respectively. All cases are considered for $m_i^p < 25$ and $m_i^s < 25$. Here the contours connect points with equal pixel density in the CMD. The lower panels of Figure 14 show the red fraction, f_{red}, as a function of i-band apparent magnitude for local densities corresponding to the 20% most dense (red), 60%–80% densest

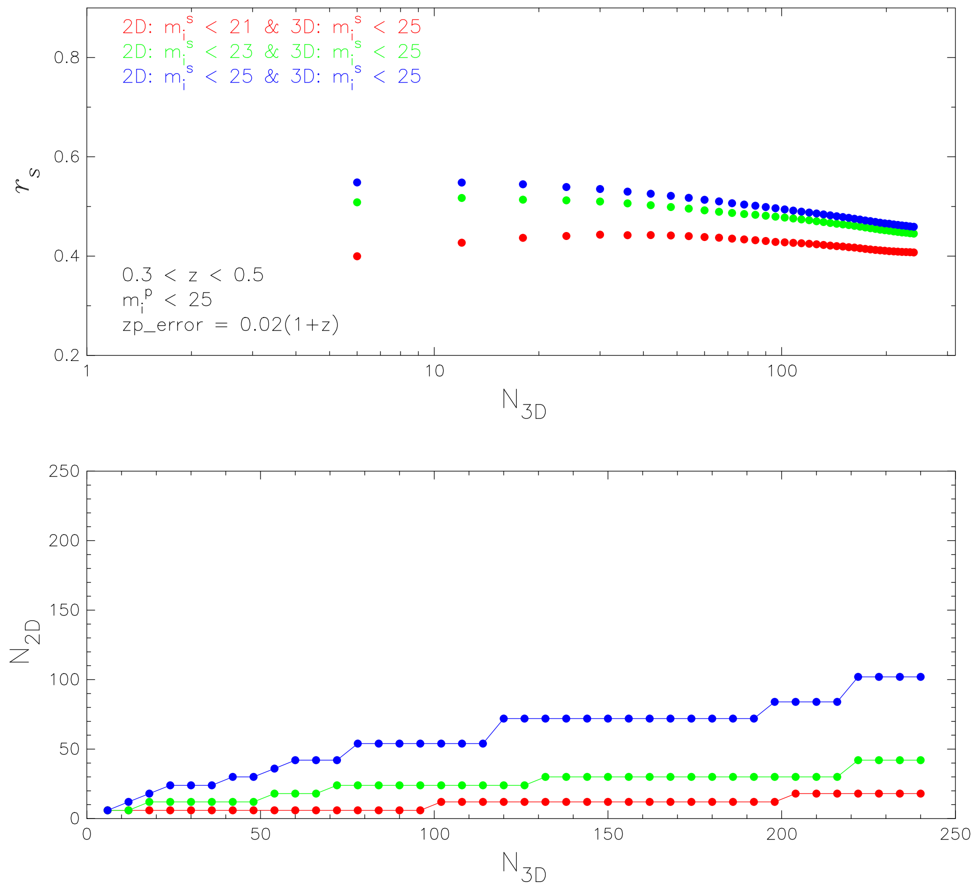

Figure 11. The largest r_s (upper panel) obtained by varying N_{2D} and the corresponding choice of N_{2D} (lower panel) that yields the best correlation between the real-space and 2D projected environments for different choices of N_{2D} as a function of N_{3D} from mock galaxy catalogs. Here the secondary magnitude limit in 3D measurement is set to $m_i^s = 25$. Dots of different color correspond to samples with different secondary magnitude limits in 2D measurements: $m_i^s < 21$, $m_i^s < 23$, and $m_i^s < 25$.

(orange), 40%–60% densest (yellow), 20%–40% densest (green), and 20% least dense (blue). The error bars show the 1σ Poisson uncertainty in each bin.

It is clear that in the simulation, galaxy colors are strongly correlated with environment, being redder in denser environments. In the case where there is no photo-z error, the red and blue contours occupy distinct regions in the CMD. This trend is in good agreement with observational results (Balogh et al. 2004; Cooper et al. 2006, 2007; Cassata et al. 2007). As the photo-z error increases (from left to right), red and blue contours begin to overlap. To further quantify the influence of the photo-z uncertainty on the CMD, we plot the red fraction as a function of i-band apparent magnitude for galaxies located in five different density percentiles (lower panels of Figure 14). The difference in the red fraction becomes gradually smaller with increasing photo-z error. Considering the case without photo-z error and galaxies $m_i > 20$, the difference in red fraction between the 20% most dense and 20% least dense environments ranges from 0.5 to 0.6, but decreases to 0.3–0.4 in the case of photo-z error = 0.06(1 + z). Nevertheless, it is still encouraging that even in the worst case (photo-z error = 0.06(1 + z)), the dependence of the red fraction on galaxy environment can still be seen.

Figure 15 presents similar information to Figure 14, but now for the cases with various secondary magnitude limits. Here we consider the cases with $m_i^s < 25$, $m_i^s < 23$, and $m_i^s < 21$. All the three cases are considered using $m_i^p < 25$. To remove additional uncertainties due to projection effects, the density ranking is based on the 3D real-space environments, ρ_6. Our results show that the environment measured with fainter secondary magnitude limits yields a better correlation with galaxy properties. This result is somewhat expected because the scales of the environment defined by different m_i^s are different for a given neighbor N, being greater for brighter m_i^s. If the color–density relation is scale-dependent, it can lead to a dependence on the adopted m_i^s. The degraded color–density relation with a brighter sample in Figure 15 could be due to environmental effects on color being weaker with increasing scale of the environment. Furthermore, a fainter sample is spatially denser and therefore contains more information about the environment than a sparse sample does. Including more galaxies in the density estimate thus also helps to characterize the environments. We will investigate the correspondence between galaxy environment and dark matter halo mass in the next section.

5.2. *Environments versus Dark Matter Halo Mass*

Observationally it is found that galaxies located in massive halos, such as groups and clusters, are in general formed earlier and hence are more evolved than galaxies in the field (Capak et al. 2007). Semi-analytic models of galaxy formation have also successfully reproduced the observed trend (Lemson & Kauffmann 1999; Benson et al. 2000; Haas et al. 2012;

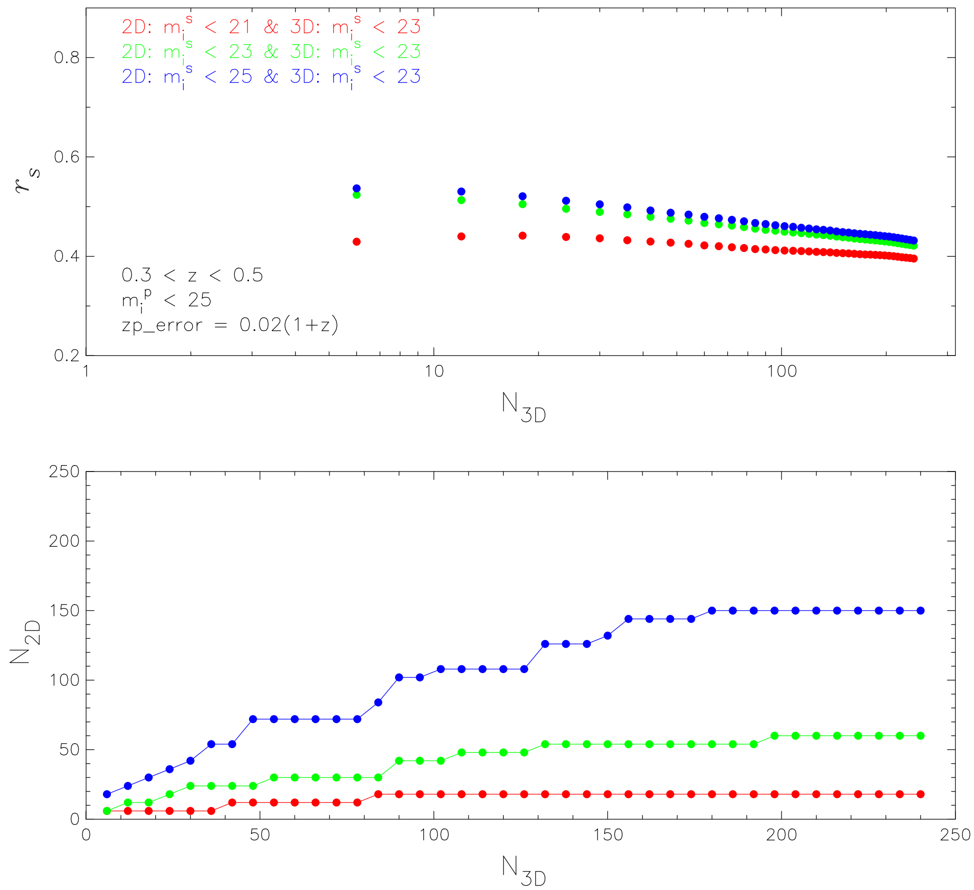

Figure 12. Similar to Figure 11 but the secondary magnitude limit in 3D measurement is set to $m_i^s = 23$.

Muldrew et al. 2012). Figure 16 shows the red fraction as a function of dark matter halo mass in our mock catalog based on the model of Lagos et al. (2012). As can be seen, the red fraction increases rapidly toward massive halos, in good agreement with observation.

We now proceed to show the correlation between host halo mass and overdensity in order to understand why the galaxy environment calculated using a fainter secondary magnitude limit has a better correlation with galaxy color. We adopt a method similar to the one used in Muldrew et al. (2012), which is to plot the relationship between host halo mass and overdensity. The upper panels of Figure 17 show the correlations between host halo mass and the 3D overdensity for the galaxy samples with different secondary magnitude limits. In general we find that, even in the case of zero redshift error, the 3D environment measures are a poor tracer of mass for individual objects as revealed by the large scatters between 3D density field and halo mass. Similar to what is found by Muldrew et al. (2012), for low N_{3D} a galaxy found at high 3D density is actually more likely to be in a low-mass halo than in a high-mass one. Furthermore, our results show that for a fixed N_{3D}, the low-density region probed using a brighter m_i^s has a wider spread in halo mass. In contrast, the low-density region measured using a fainter m_i^s is dominated by galaxies located in small halos ($<10^{12}\,M_\odot$) where the red fraction drops significantly with decreasing halo mass (see Figure 16). This is because the scale of the density defined by a fainter magnitude limit for a fixed N_{3D} is typically smaller and less contaminated by the two-halo term, and therefore is a better tracer of the halo mass, except for very massive halos ($>10^{14}\,M_\odot$).

This point is further illustrated in the lower panels of Figure 17, where we show that a smaller N_{3D} yields a stronger correlation between overdensity and host halo mass than using a larger N_{3D}. Therefore the tighter relationship between the galaxy colors and densities computed using a fainter secondary magnitude limit seen in Figure 15 can be attributed to the fact that the environment defined using a fainter sample traces the host halo masses more closely.

Figure 18 shows $1 + \delta_6^{2D}$ versus halo mass for the galaxies with photo-z errors varying from 0.0 to $0.06(1 + z)$. The overdensity $1 + \delta_6^{2D}$ increases with host halo mass but with a large scatter, as seen in Muldrew et al. (2012), which is based on different simulations. Nevertheless, the correlation becomes progressively weaker when photo-z uncertainty increases. In the case of photo-z error $= 0.06(1 + z)$, the correlation is almost flat, suggesting that the 2D projected density is no longer a good tracer of halo mass when photo-z errors are non-negligible.

5.3. Comparison with Spectroscopic Observation

In previous sections, we have presented how the photo-z uncertainty can have an impact on the measurement of local density and discussed how well the 2D projected density traces the real-space density when adopting different choices of the size of velocity (redshift) window, magnitude limit, and the Nth nearest neighbor. Furthermore, we have also studied how the color–density relation is affected by the presence of photo-z errors and as a function of the magnitude limit that is applied to the secondary sample. In this section, we further investigate the difference in the local density measurements between the photo-z and spectral-z samples using mock galaxies, by taking

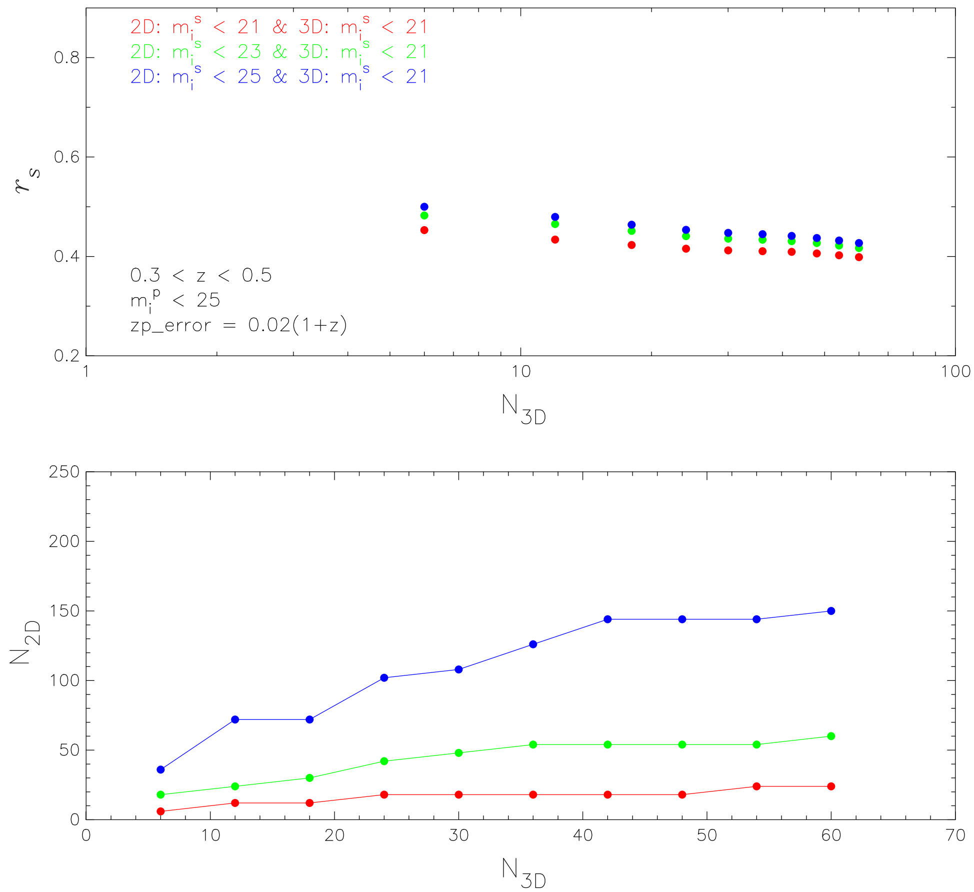

Figure 13. Similar to Figure 11 but the secondary magnitude limit in 3D measurement is set to $m_i^s = 21$.

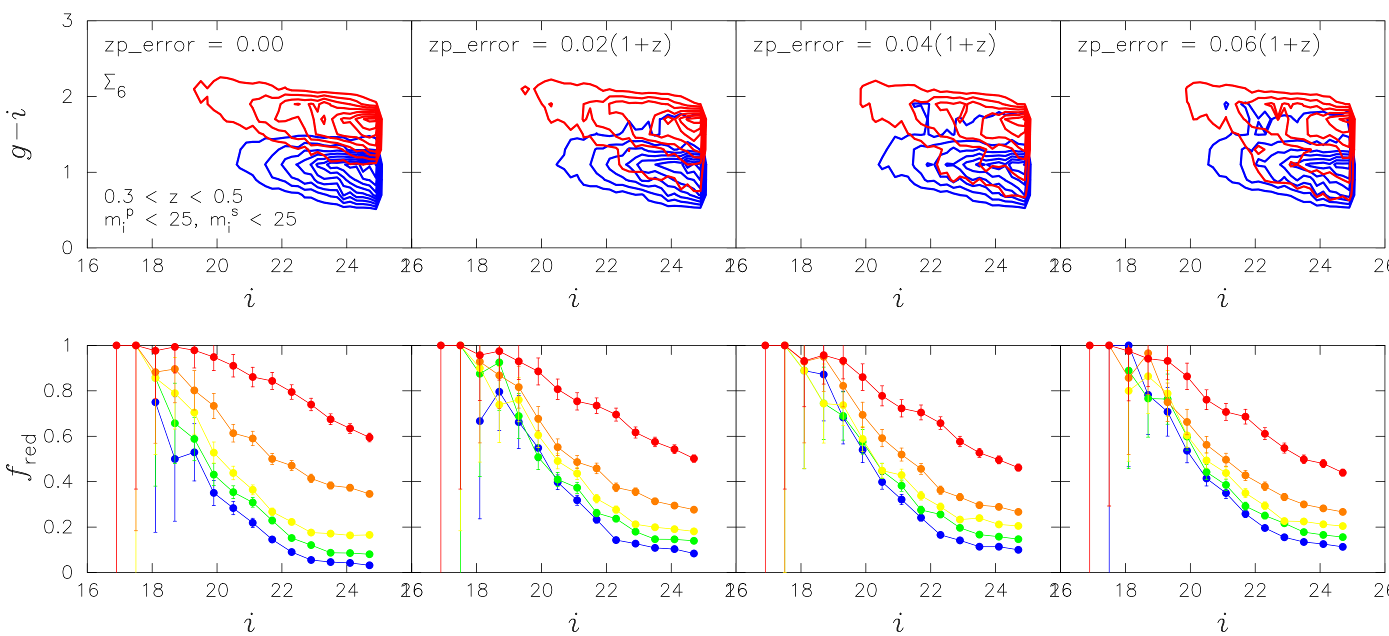

Figure 14. Upper panels: color–magnitude diagrams for galaxies in the 20% most dense (red contours) and 20% least dense (blue contours) environments with different photo-z errors (from left to right: 0, $0.02(1 + z)$, $0.04(1 + z)$, and $0.06(1 + z)$). Lower panels: the red fraction, f_{red}, as a function of i-band apparent magnitude with different percentages of density: the 20% most dense (red), 60%–80% densest (orange), 40%–60% densest (yellow), 20%–40% densest (green), and 20% least dense (blue). The error bars are given by Poisson statistics, and the contours show the regions of constant galaxy number.

into account more realistic situations including the incompleteness of the spectroscopic sample.

As we discussed in Section 4, the photo-z uncertainty has a strong impact on the correlation between the 2D and 3D densities. Ideally, the density measurement based on the spectral-z sample is more reliable. However, spectroscopic observations of galaxies are time-consuming and hence are normally limited to a small sample size, brighter galaxies, and a

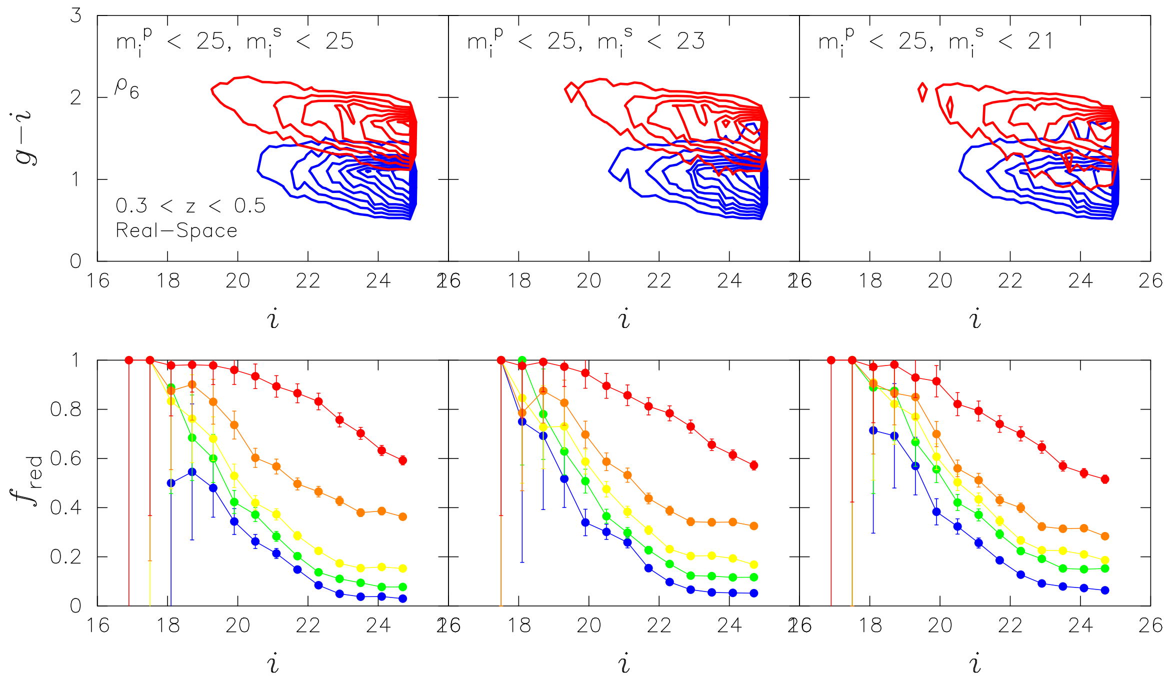

Figure 15. Upper panels: color–magnitude diagrams for galaxies in the 20% most dense (red contours) and 20% least dense (blue contours) real-space environments with different secondary magnitude limits (from left to right: $m_i^s < 25$, $m_i^s < 23$, and $m_i^s < 21$). Lower panels: the red fraction, $f_{\rm red}$, as a function of i-band apparent magnitude with different percentages of density: the 20% most dense (red), 60%–80% densest (orange), 40%–60% densest (yellow), 20%–40% densest (green), and 20% least dense (blue). The error bars are given by Poisson statistics, and the contours show the regions of constant galaxy number.

lower redshift range. Moreover they often suffer from incompleteness due to limited observing time as well as fiber and/or slit collisions. In contrast, the photo-z can be relatively easily obtained down to fainter galaxies and out to higher redshifts with a much larger sample size, but with the drawback that the redshift resolution is substantially poorer than for spectral-z. Nevertheless, both spectral-z and photo-z samples have been used in environment studies (Cooper et al. 2006; Cassata et al. 2007; Elbaz et al. 2007; Quadri et al. 2012). It is thus interesting to investigate to what extent the local density from photo-z samples can be compared to that from spectral-z samples. To do so, we randomly choose parts of the entire spectral-z sample from the mock catalogs to simulate spectral-z samples with different percentages of completeness. We then apply our optimized scheme as introduced in Section 4.2 to both incomplete spectral-z samples and photo-z samples to compare their r_s.

Figures 19–21 show r_s as a function of $N_{\rm 3D}$ for spectral-z samples with different completeness values of 10%, 20%, 40%, 60%, 80%, and 100% (denoted by different colors). For comparison, we also overplot the results of photo-z samples with different photo-z uncertainties presented in different line styles in the top-left panel. To fairly compare r_s on the same physical distance scales among various cases, r_s is calculated using the $N_{\rm 3D}$ that corresponds to the same Nth nearest neighbor in the case of a 100% complete spectroscopic sample. An enlarged version on small scales is shown in the top-right panel. The difference among the three Figures (19–21) is the secondary magnitude limit applied. For example, in Figure 19 we consider the case where galaxy environments are calculated by using $m_i^s < 25.0$. It shows that the galaxy environments calculated by using an incomplete spectral-z sample with this deep magnitude selection are more reliable than those calculated by using a complete photo-z sample.

However, in general it is difficult to obtain spectroscopic redshifts for a large sample of very faint galaxies. For example, the DEEP2 survey (Newman et al. 2013) is limited to $R < 24.1$ and the zCOSMOS bright sample (Lilly et al. 2007) is limited to $i = 22.5$. Next we vary the secondary magnitude limits of the spectroscopic sample to see how the trend changes. Here we consider the two optimized results, $m_i^s < 24.1$ and $m_i^s < 22.5$, to roughly mimic the results of DEEP2 sky survey and zCOSMOS bright survey, respectively. Strictly speaking, the DEEP2 is limited in the R-band instead of the i-band; however, here we simply adopt the i-band in order to reveal the trend more clearly. Our results show that in the case of $m_i^s < 24.1$, the performance of density recovery with photo-z error as low as $\sim 0.02(1 + z)$ is always worse than that of the spectral-z samples. However, when the magnitude limit of the spectral-z samples decreases to $m_i^s < 22.5$, the performance becomes comparable to that of the spectral-z sample with 10% completeness. In other words, as the spectral-z sample gets brighter, the r_s coefficients between the spectral-z and photo-z samples become closer. However, the difference between their r_s coefficients gradually becomes larger when we probe a larger scale of environment, as described in Section 4.2. That is, a deeper photo-z sample can yield a performance as good as an incomplete, shallower spectral-z sample, but this is restricted to small-scale environments.

5.4. Effect of Outliers

So far the studies on the effect of the photo-z uncertainty on the density measurement have been carried out by perturbing the redshifts of mock galaxies with a Gaussian function. However, this method does not totally mimic the realistic case because the photo-z errors may not exactly follow the Gaussian distribution. For example, in the cases where there are

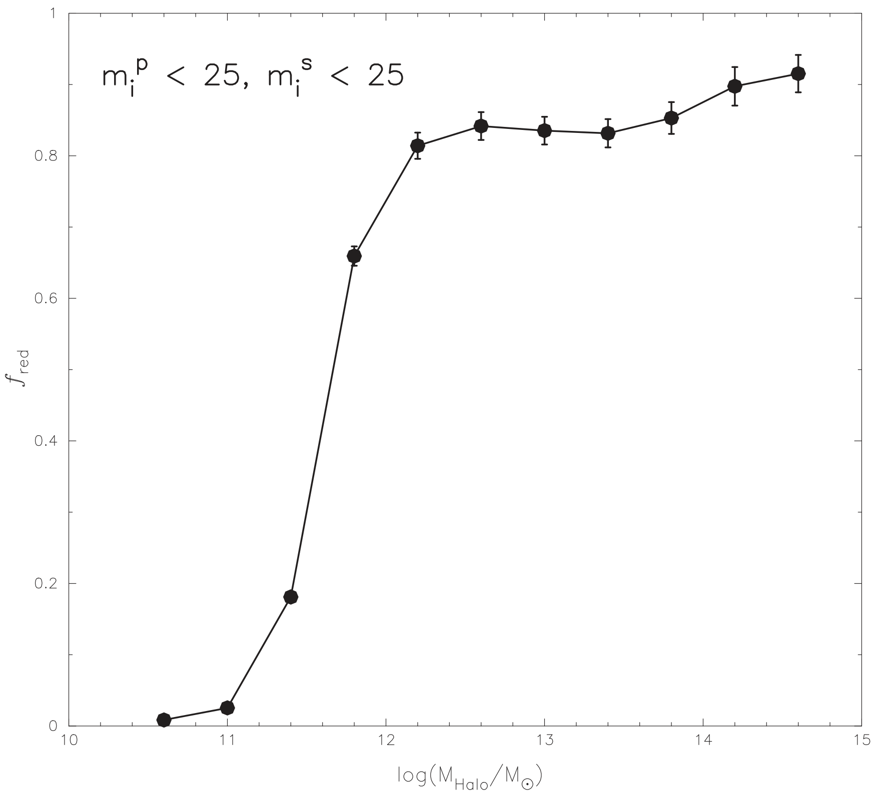

Figure 16. Red fraction, f_{red}, as a function of dark matter halo mass from mock galaxy catalogs. The error bars are given by Poisson statistics.

insufficient numbers and/or wavelength coverage of bandpasses, the feature of the Lyman break (~912 Å) can be misidentified as the Balmer break (~4000 Å) and vice versa, leading to a catastrophic failure in the photo-z estimation, the so-called "redshift outliers" (Brough et al. 2013). Next we study how the outlier rate influences the correlation between the 2D and 3D local density measurements.

To simulate galaxy samples with outliers, we randomly choose part of the entire simulation samples according to the desired outlier rate, and assign them a new redshift randomly between 0 and 2.0. Figure 22 shows the results using samples with photo-$z = 0.02(1 + z)$ and four different percentages of outliers: 5%, 10%, 15%, and 20%. Similar to Figures 19–21, we also mark the results for different photo-z uncertainties in the case of 0% outliers with different line styles for comparison. From this figure, we can see that r_s also depends strongly on the outlier fraction, becoming worse as the outlier fraction increases. The effect is similar to the degradation of photo-z uncertainty and the completeness of the sample. For example, for the sample with photo-z error $= 0.02(1 + z)$ and outlier rate = 10% (green dots), we find that its optimized result is similar to the result of the outlier-free sample with photo-z error $= 0.04(1 + z)$.

Figure 23 shows the CMD for the samples with photo-z error $= 0.02(1 + z)$ and four different percentages of outliers. The environments are measured by using Σ_6, $m_i^p < 25.0$, and $m_i^s < 25.0$. As can be seen, in the case with 5% outliers, it is comparable to a non-outlier case with photo-z error between $\sim 0.02(1 + z)$ and $\sim 0.04(1 + z)$, and the color–density relation can still be revealed. However, as the outlier rate goes up to 10%, the density measurements for underdense environments (for example, the 20% least dense (blue) and 20%–40% densest (green) environments) are no longer distinguishable, resulting in a weaker color–density relation. This is in contrast to the situations with pure photo-z errors, for which the curves of lowest density remain distinguishable. The outliers have larger effects in lower density environments because the change in the density measurements is proportionally larger in those regions when some fraction of galaxies are scattered inside or outside the relevant redshift window.

6. COLOR–DENSITY RELATION FOR PAN-STARRS1 DATA

The main purpose of this work is to understand the systematics in the 2D density measurement and its limitation, with the ultimate goal of applying it to the ongoing and future large photometric surveys. So far we have explored various aspects of the density measurements by using mock galaxy catalogs for which the real-space density is known. We have considered several factors such as photo-z uncertainty, magnitude limit, completeness, and outlier rate that make simulation data as similar to realistic samples as possible. However, these factors are still not sufficient to imitate realistic samples. An alternative is to compare the results of the overlapping samples directly between photo-z and spectral-z surveys. We adopt this approach by using the PS1 MD07 photometric redshift catalog (Lin et al. 2014) because it covers the well-known EGS field, which has the spectroscopic redshifts from DEEP2, allowing for a direct comparison of the density measurements.

We first compute the environments using galaxies with spectroscopic redshifts and compare these environments with those calculated by using photometric redshifts from Pan-STARRS1 for the same galaxies. For comparison, we also utilize the mock galaxy catalogs described in Section 2.1 and

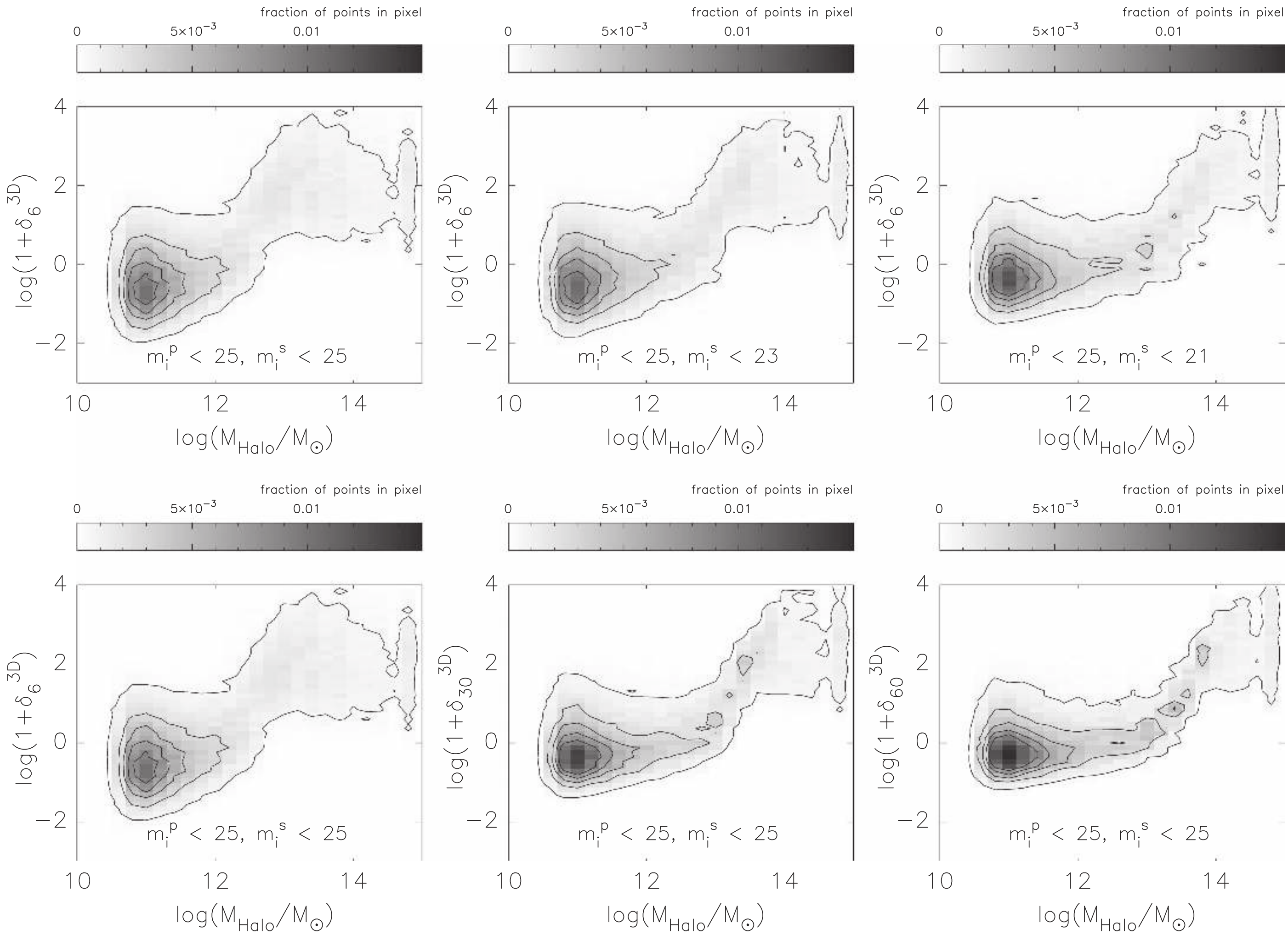

Figure 17. Relationship between host halo mass and galaxy overdensity, $1 + \delta_n^{3D}$, assuming no redshift error, with three different secondary magnitude limits (top panels, from left to right: $m_i^s < 25$, $m_i^s < 23$, and $m_i^s < 21$) and choices of N_{3D} (bottom panels, from left to right: $N_{3D} = 6$, $N_{3D} = 30$, and $N_{3D} = 60$). The contours show the regions of constant galaxy number.

perturb their redshifts to simulate the photo-z conditions of Pan-STARRS1, where the typical error is $\sim 0.06(1 + z)$ and the outlier rate is ~6%. Figure 24 shows the scatter plot for galaxies in the EGS field (left panel) and in the simulations (right panel). In the left panel, the 2D and 3D environments are evaluated by using samples from Pan-STARRS1 (photo-z) and DEEP2 (spectral-z) respectively. As can be seen, while we compare the 2D projected and 3D environments in the realistic case, the slope of the scatter plot is similar to the simulation result but r_s is smaller than the results of simulation. While this might be explained by the intrinsic difference in the spatial distribution of galaxies between the real universe and simulations, it is also noticed that the galaxies in the MD07/EGS field span a narrower range in 3D overdensity because of the smaller field size, such that the extreme environments are not well sampled. As a result, the simulated sample includes very dense environments that are more discernible and easier to recover than intermediate environments. Nevertheless, there still exists a weak correlation in comparison with real data even though their r_s coefficients are smaller than those of the simulated data.

To know whether the color–density relation can still be revealed in the realistic photo-z sample, in Figure 25 we plot the CMD (upper panels) and the red fractions versus i_{p1} magnitude (lower panels) for galaxies located in the 20% most dense (red contours) and 20% least dense (blue contours) galaxy environments. The red fraction is defined as the ratio of the number of galaxies with $g_{p1} - i_{p1}$ color redder than 1.5 to that of the full sample. For each galaxy in the right panel of PS1, we compute Σ_6 with $m_i^s < 24$, using the photometric redshifts derived in Lin et al. (2014) with the redshift range $0.3 < z < 0.5$. The left panel of Figure 25, which is for comparison, shows the DEEP2 result using galaxy densities computed by Cooper et al. (2005). Their density measurements have been corrected for several effects such as the survey edges, redshift precision, redshift-space distortion, and target selection as described in Cooper et al. (2005). Therefore, their measurements can be regarded as the "true" answer in this comparison. The middle panel shows the result calculated with PS1 photometric redshifts only for galaxies located in the region overlapping with the EGS field, which is $\sim 0.5\ \mathrm{deg}^2$ (~1500 galaxies), while the right panel shows the result based on the entire PS1/MD07 field of $\sim 5\ \mathrm{deg}^2$ (~25,000 galaxies). To minimize the impact of the edge effects, we also exclude galaxies near the survey boundaries when showing the color–density relation. Among all the three samples, the color–density relation is

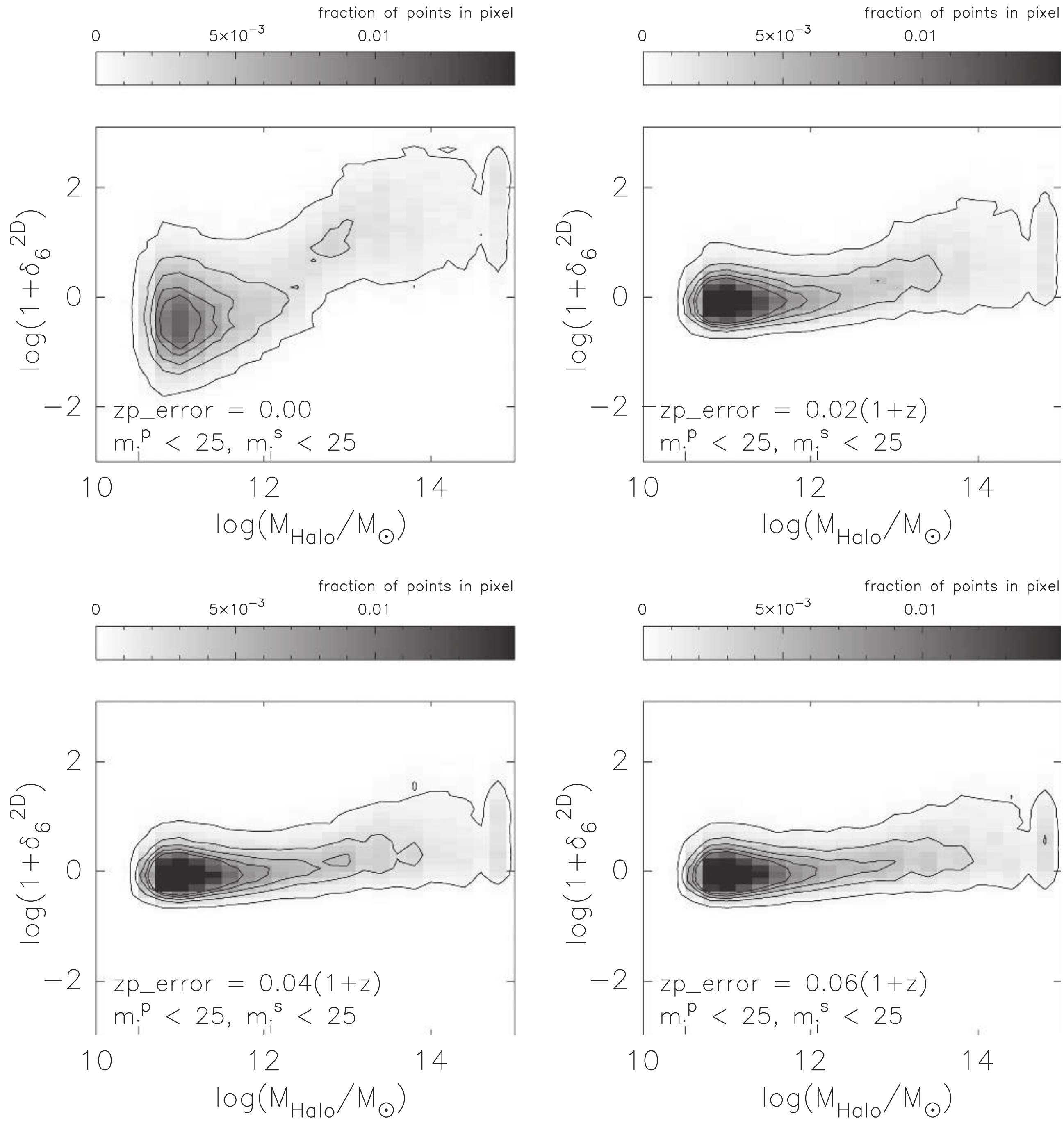

Figure 18. Relationship between host halo mass and galaxy overdensity, $1 + \delta_6^{2D}$, with four different photo-z errors: 0.00, $0.02(1 + z)$, $0.04(1 + z)$, and $0.06(1 + z)$. The contours show the regions of constant galaxy number.

significantly detected ($>3\sigma$) in only the $\sim 5\,\deg^2$ PS1 photo-z sample. This is because, although the photo-z uncertainty in general contaminates the density measurement, which leads to some systematics in the color–density relation, the random errors can be largely improved given the large volumes probed by a photometric survey. In other words, the reduced errors due to the larger sample are sensitive enough to allow for the detection of a "degraded" relation between red fraction and environment.

Furthermore, we also extend our study from a low redshift range ($0.3 < z < 0.5$) to a higher redshift range ($0.6 < z < 0.8$). In this redshift bin, $g_{p1} - i_{p1} = 2.0$ is used to separate blue and red galaxies. Similarly we first show the color–density relation for the redshift range $0.6 < z < 0.8$ using our simulated data set (Figure 26) and PS1 samples (Figure 27, ∼45,000 galaxies). As shown in the right panel of Figure 27, the difference in the red fraction between two extreme environments is still detectable at $\sim 2\sigma$–3σ level. The Kolmogorov–Smirnov test (Andrade et al. 2001) on the color distributions for galaxies with $21 < i_{p1} < 23$ located in the 20% most dense and 20% least dense percentiles returns a value of $p \ll 0.1\%$, rejecting the null hypothesis that the color distributions of galaxies are drawn from the same population.

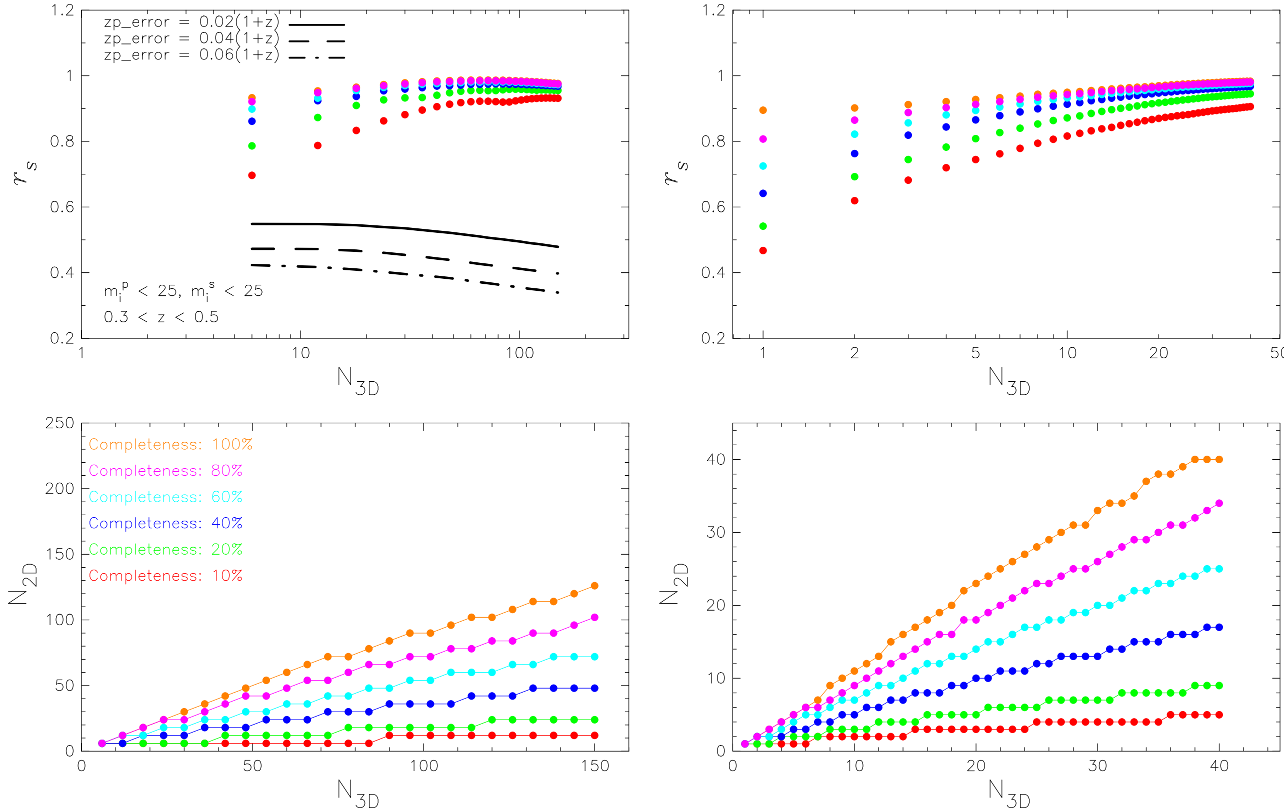

Figure 19. The largest r_s obtained by varying N_{2D}, for a large choice of N_{3D} (top-left panel) and a small choice of N_{3D} (top-right panel), and the corresponding choice of N_{2D} (bottom panels) that yields the best correlation between the real-space and 2D projected environments for different choices of N_{2D} as a function of N_{3D} from mock galaxy catalogs. These cases here are calculated using, $m_i^p < 25.0$ and $m_i^s < 25.0$. Different colors are for samples with different sampling rates (namely, spectroscopic completeness): 10%, 20%, 40%, 60%, 80%, and 100%, and different line styles represent cases with different photo-z uncertainties as shown in Figure 10.

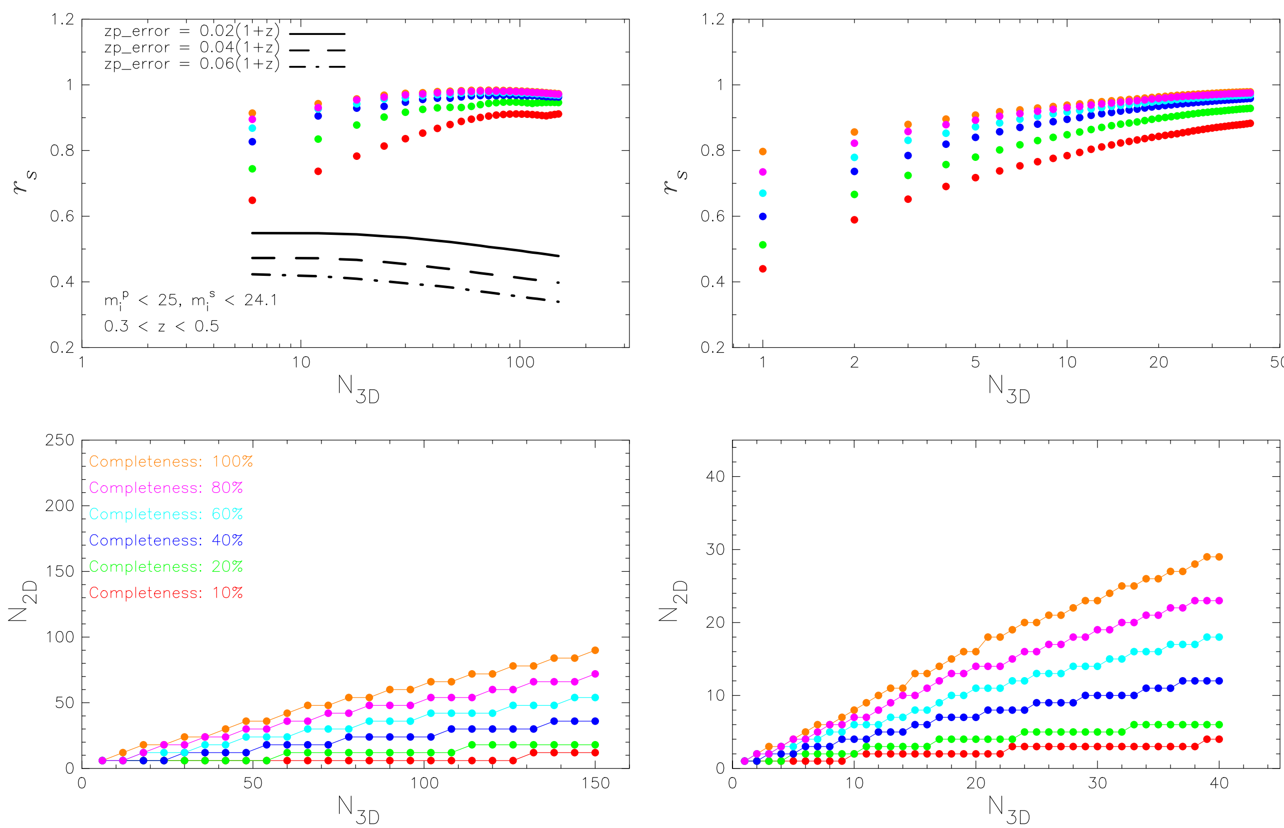

Figure 20. Similar to Figure 19 but with $m_i^s < 24.1$ in the case of the redshift uncertainty equal to zero.

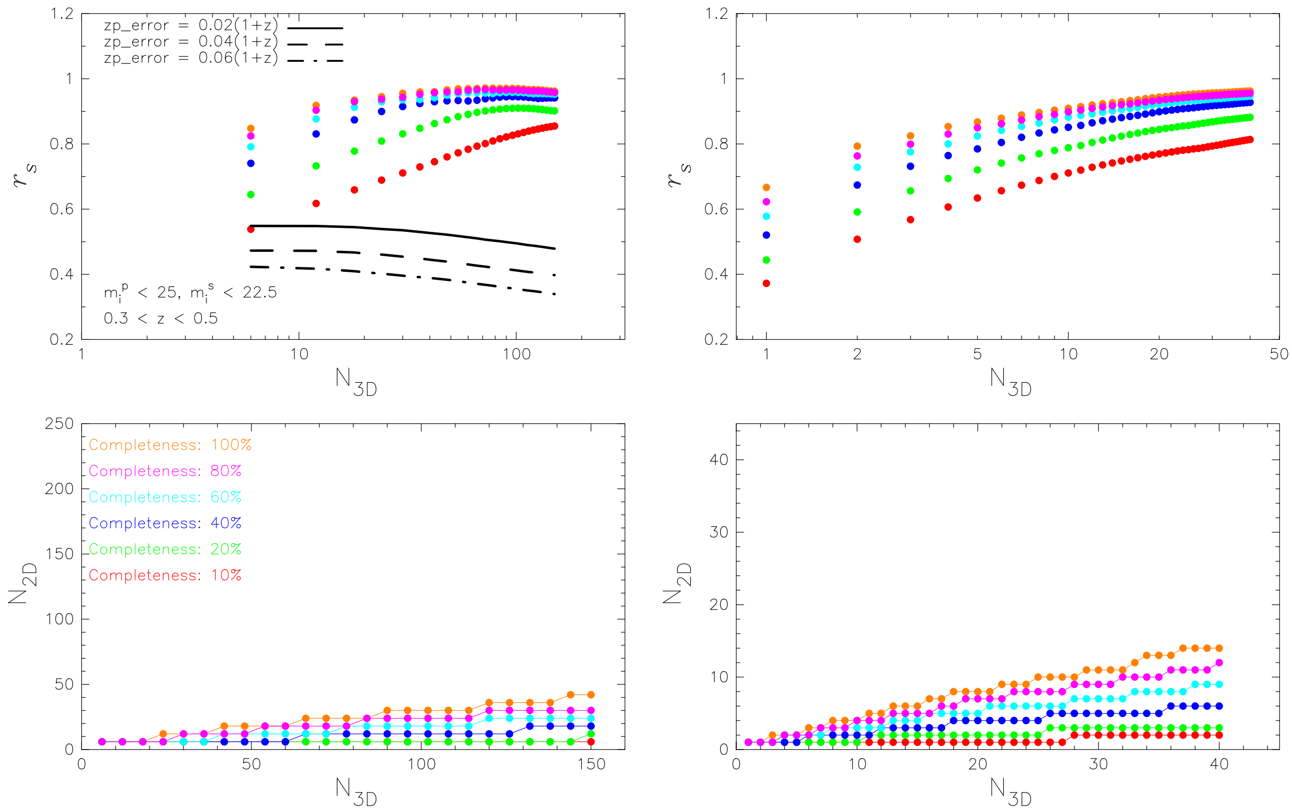

Figure 21. Similar to Figure 19 but with $m_i^s < 22.5$ in the case of the redshift uncertainty equal to zero.

7. CONCLUSIONS

In this work, we have studied how the 2D projected environment correlates with the 3D real-space environment. Using the Durham mock galaxy catalogs, we investigate various parameters in measuring the 2D projected environment and find the best parameters that maximize Spearman's rank correlation coefficient, defined as $r_s = 1 - \frac{6\sum d_i^2}{s(s^2-1)}$, which is a means of quantifying the correlation between 3D real-space and 2D projected overdensities. When applying the Nth nearest-neighbor method to the PS1 photo-z sample, we show that the color–density relation can still be revealed despite the sizable photo-z errors inherent in the data. Our main conclusions are as follows.

i. The correlation between the 2D projected and 3D real-space overdensity is sensitive to the photo-z uncertainty. Smaller N_{3D} is recommended for photo-z samples to achieve a better correlation (larger r_s) between 2D projected and 3D real-space environments.

ii. As the scale of 3D real-space environment increases, the r_s values derived by using spectral-z and photo-z samples show opposite trends: the correlation becomes gradually stronger for the spectral-z samples but worse for the photo-z samples.

iii. The 2D projected environment measurements are less sensitive to the redshift interval (V_{cut}). The redshift interval comparable to the photo-z uncertainty yields a 2D projected overdensity that reasonably traces the real-space density.

iv. The magnitude limit should also be considered when computing local densities of galaxies. The 2D environments measured with fainter magnitude limits yield better correlation with the 3D real-space environments derived from a sample with the same limiting magnitude for a fixed N_{3D}. In addition, the color–density relation is more prominent if the density is measured using fainter magnitude limits for a fixed N_{2D}. This is because the overdensity computed with fainter magnitudes probes smaller scales with the same N_{2D}, and traces the hosting halo mass of galaxies better.

v. Considering the case calculated by using galaxy samples at $0.3 < z < 0.5$ with $m_i^p < 25.0$ and $m_i^s < 25.0$, the performance of recovery of small-scale environments for photometric redshift samples with redshift uncertainty of $0.02(1 + z)$ is roughly comparable to that for shallower $i \sim 22.5$ spectroscopic redshift samples with ~10% completeness. In addition, the effect of catastrophic failures in the photo-z measurements on the density measurement is similar to that of the photo-z errors.

vi. Using Durham mock galaxy catalogs in the redshift range $0.3 < z < 0.5$, we show that the density-dependent red fraction can still be revealed in photometric redshift samples with photo-z uncertainty up to $0.06(1 + z)$. Similarly with the photo-z sample from PS1, we show that the color–density relation is also present in the sample whose photo-z uncertainty is $\sim 0.06(1 + z)$ and the outlier rate is ~6%, but the significance depends strongly on the sample size. Based on the results of PS1 in the two redshift bins ($0.3 < z < 0.5$ and $0.6 < z < 0.8$), we recommend that the survey size should at least exceed $\sim 5\,\mathrm{deg}^2$ in order to yield $>3\sigma$ results. Larger fields will be required in order to reduce the Poisson errors if going to higher redshifts because the color–density relation is less prominent and the number density of galaxies is reduced.

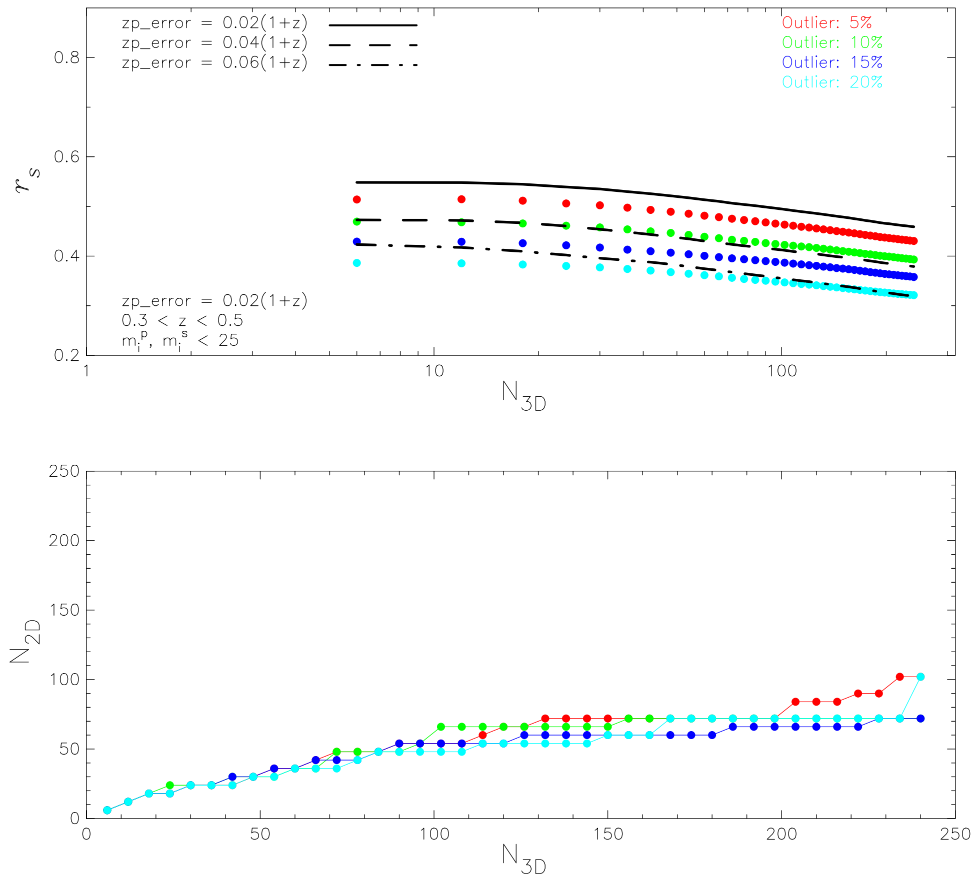

Figure 22. The largest r_s (upper panel) obtained by varying N_{2D} and the corresponding choice of N_{2D} (lower panel) that yields the best correlation between the real-space and 2D projected environments for different choices of N_{2D} as a function of N_{3D} from mock galaxy catalogs. Different colors represent cases with various outlier rates in the case of photo-z uncertainty equal to $0.02(1 + z)$, and different line styles represent cases with different photo-z uncertainties as shown in Figure 10.

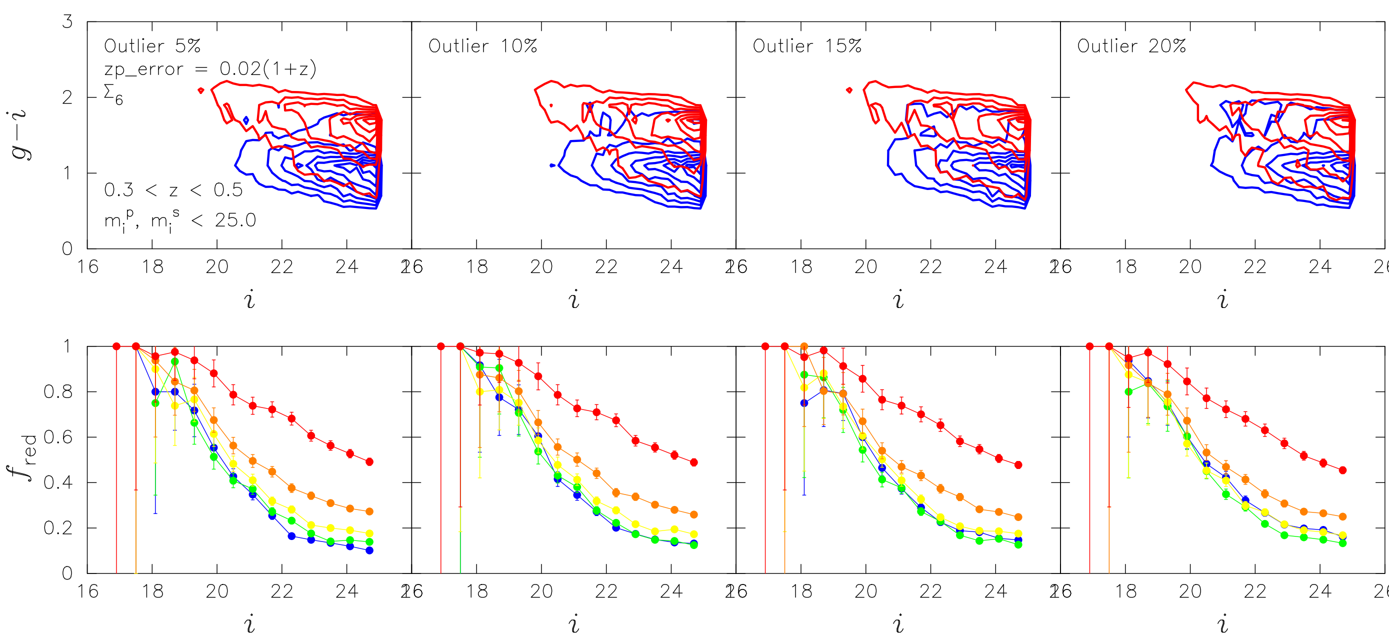

Figure 23. Upper panels: color–magnitude diagrams for galaxies in the 20% most dense (red contours) and 20% least dense (blue contours) environments with different percentages of outliers (from left to right: 5%, 10%, 15%, and 20%). Lower panels: the red fraction, f_{red}, as a function of i-band apparent magnitude with different percentages of density: the 20% most dense (red), 60%–80% densest (orange), 40%–60% densest (yellow), 20%–40% densest (green), and 20% least dense (blue). The error bars are given by Poisson statistics, and the contours show the regions of constant galaxy number.

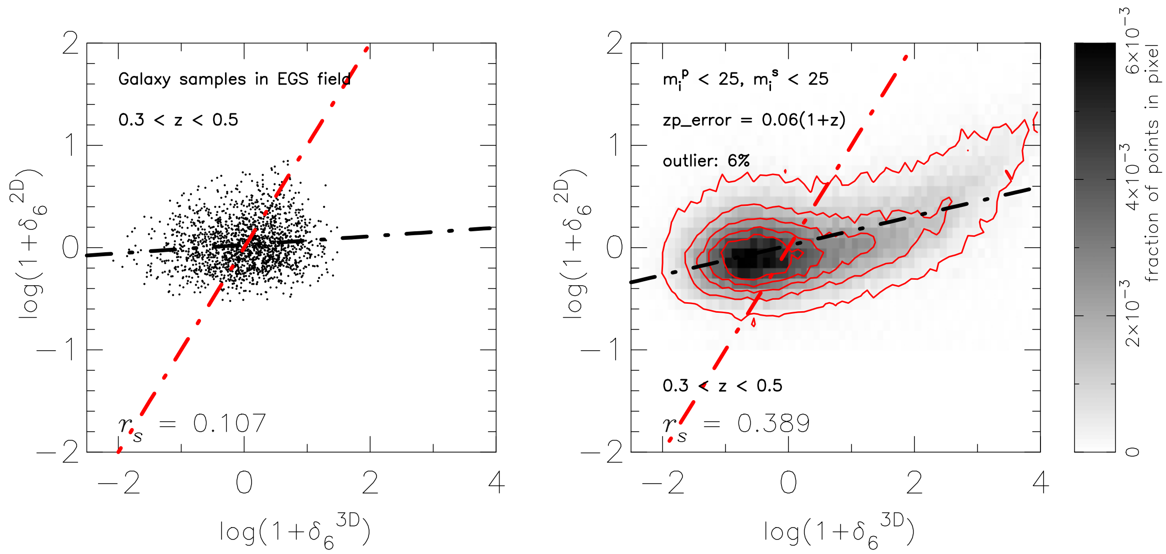

Figure 24. Left panel: scatter plot of the 3D real-space overdensity, $1 + \delta_6^{3D}$, vs. 2D projected overdensity, $1 + \delta_6^{2D}$, with galaxy samples in the EGS field. The 2D and 3D environments are calculated by using redshifts from Pan-STARRS1 and DEEP2 respectively. Right panel: similar to the left panel but using mock galaxy catalogs for the case with real-space overdensity and the case with photo-z error = $0.06(1 + z)$ and 6% outliers. The primary and secondary magnitude limits are considered by using $m_i^p < 25$ and $m_i^s < 25$. The numbers in the bottom left of each panel indicate the r_s coefficient. The black dashed–dotted lines represent the best fit to the data points, the red dashed–dotted lines represent the one-to-one relation, and the contours show the regions of constant galaxy number.

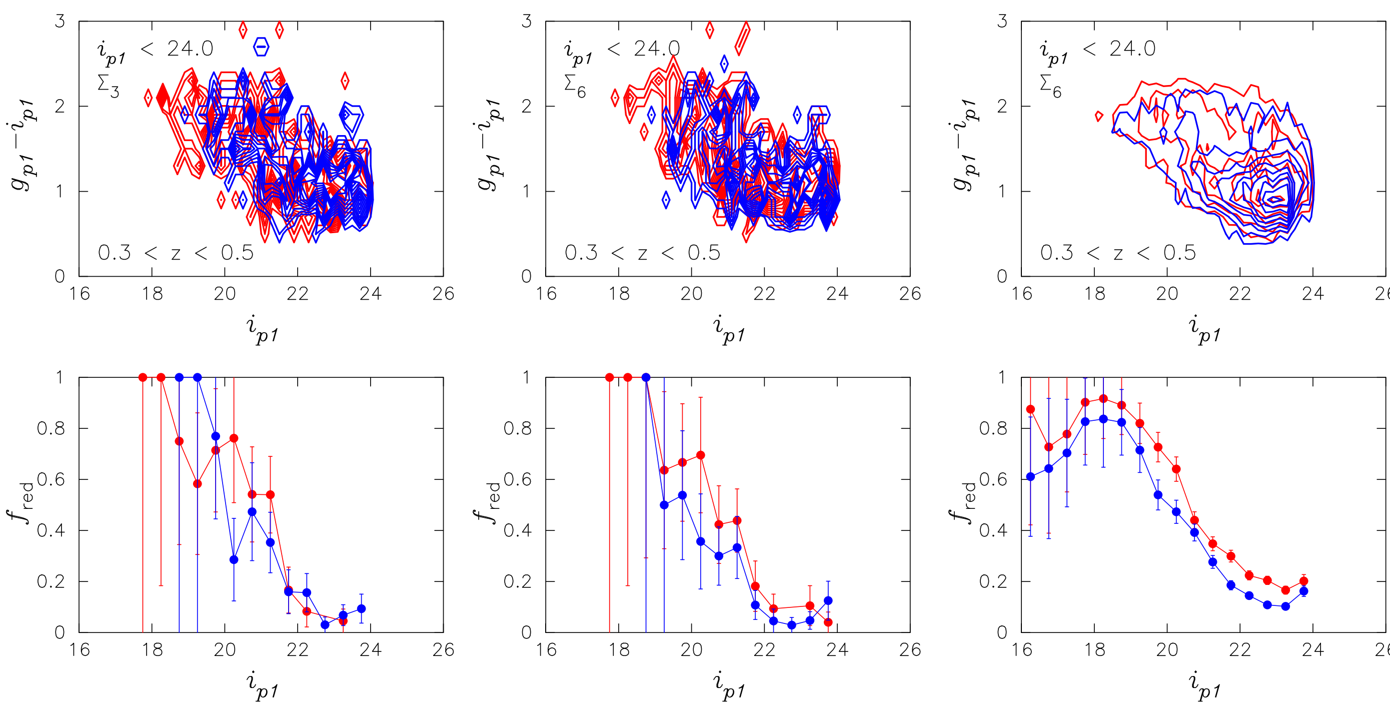

Figure 25. Upper panels: the color–magnitude diagrams for galaxies in the 20% most dense (red contours) and 20% least dense (blue contours) galaxy environments with different data sets over the redshift interval $0.3 < z < 0.5$ (left: spectral-z sample from DEEP2 in the EGS field; middle: PS1 photo-z sample in the overlapping region with the EGS field; right: PS1 photo-z sample with ~5 deg^2). Lower panels: the red fraction, $f_{\rm red}$, as a function of i-band apparent magnitude in the 20% most dense (red) and 20% least dense (blue) galaxy environments. The error bars are given by Poisson statistics, and the contours show the regions of constant galaxy number.

We thank the anonymous referee for very constructive comments, which greatly improved this manuscript. This work is supported by the Ministry of Science and Technology of Taiwan under the grants NSC101-2112-M-001-011-MY2, MOST103-2112-M-001-031-MY3, and MOST102-2112-M-001-001-MY3. H.-Y. Jian acknowledges the support of NSC101-2811-M-002-075. W.-P. Chen acknowledges the support of NSC102-2119-M-008-001. We thank Michael C. Cooper for providing the DEEP2 environment measurements for our work. We also thank the PS1 builders and PS1 operations staff for construction and operation of the PS1 system and access to the data products provided. The Pan-STARRS1 Surveys (PS1) have been made possible through contributions of the Institute for Astronomy, the University of Hawaii, the Pan-STARRS Project

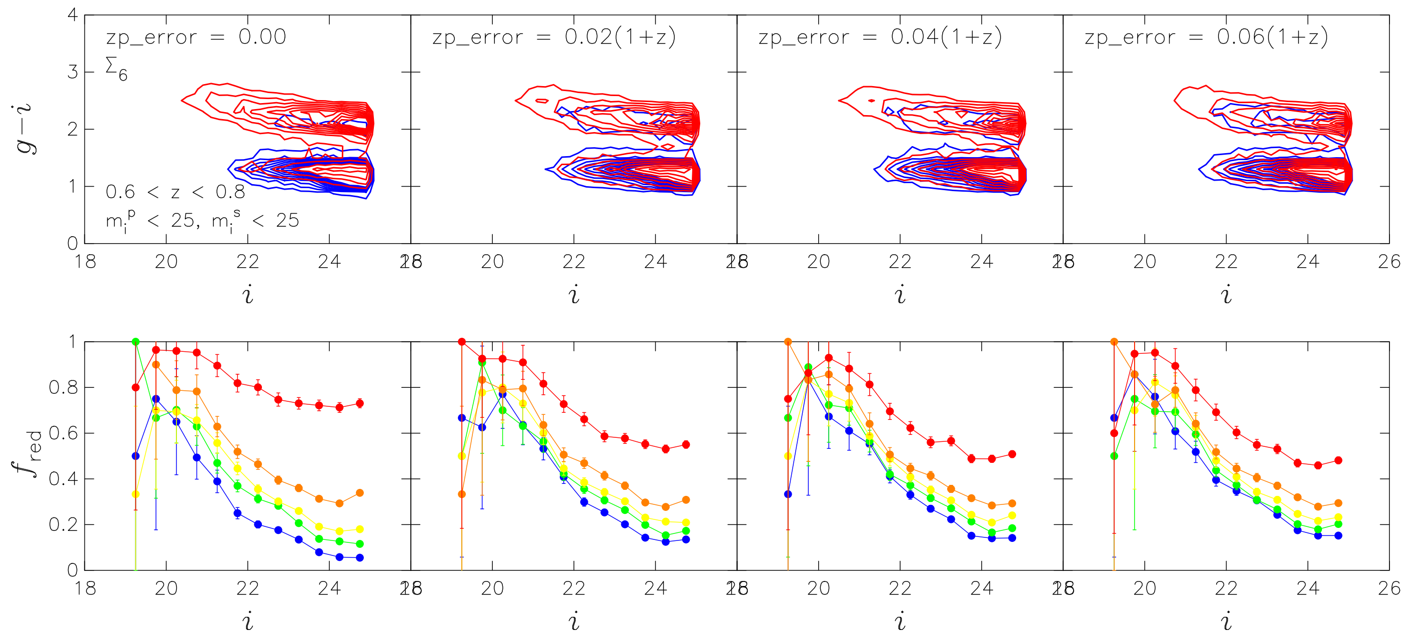

Figure 26. Similar to Figure 14, but for the redshift range $0.6 < z < 0.8$.

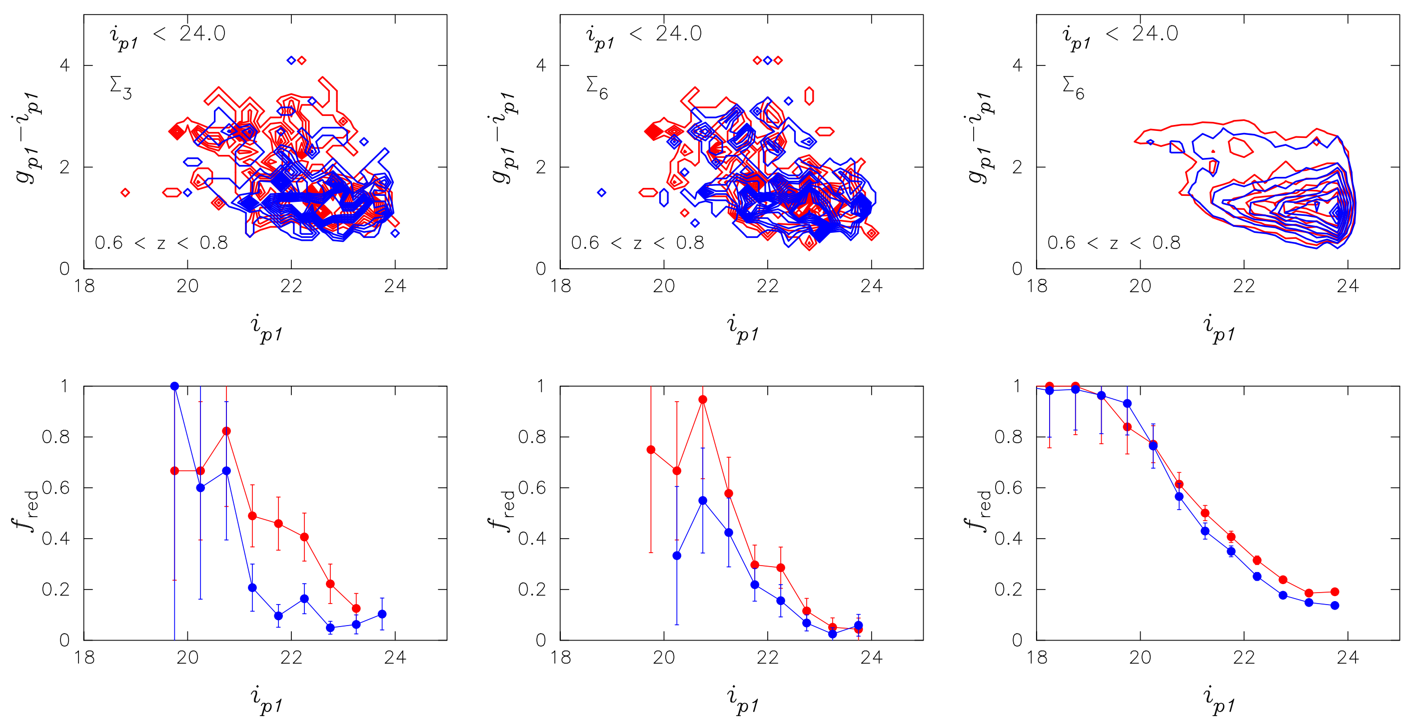

Figure 27. Similar to Figure 25, but for the redshift range $0.6 < z < 0.8$.

Office, the Max-Planck Society and its participating institutes, the Max Planck Institute for Astronomy, Heidelberg and the Max Planck Institute for Extraterrestrial Physics, Garching, The Johns Hopkins University, Durham University, the University of Edinburgh, Queen's University Belfast, the Harvard-Smithsonian Center for Astrophysics, the Las Cumbres Observatory Global Telescope Network Incorporated, the National Central University of Taiwan, the Space Telescope Science Institute, the National Aeronautics and Space Administration under grant No. NNX08AR22G issued through the Planetary Science Division of the NASA Science Mission Directorate, the National Science Foundation under grant No. AST-1238877, the University of Maryland, and Eotvos Lorand University (ELTE).

APPENDIX
N_{3D} VERSUS N_{2D} IN ENVIRONMENT MEASUREMENTS

As mentioned in Section 4.2, the optimized N_{2D} is only half the value of N_{3D}. The ratio of the two quantities in fact depends on the size of the redshift interval when computing the 2D density. Figure 28 shows the optimized scheme for the case with different size of V_{cut}. We consider the sample with photo-z error = 0.04(1 +

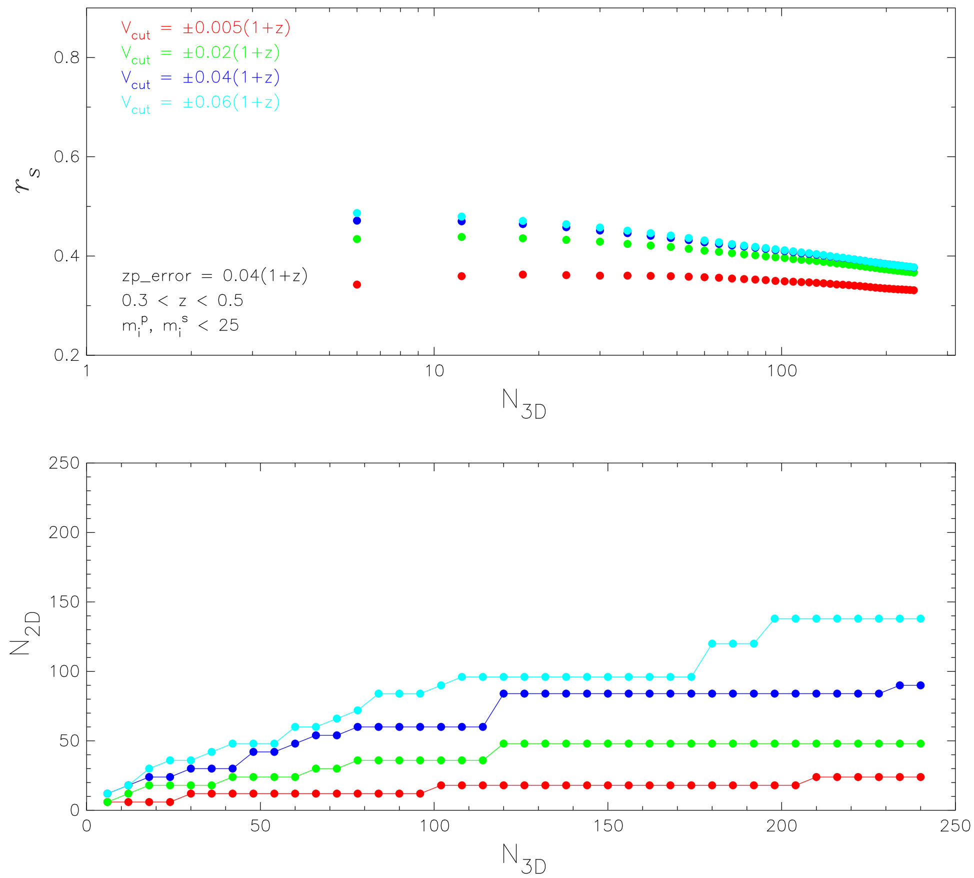

Figure 28. The largest r_s (upper panel) obtained by varying N_{2D}, and the corresponding choice of N_{2D} (lower panel) that yields the best correlation between the real-space and 2D projected environments for different choices of N_{2D} as a function of N_{3D} from mock galaxy catalogs. Dots of different color correspond to samples with different $V_{cut} = \pm 0.005(1 + z)$, $\pm 0.02(1 + z)$, $\pm 0.04(1 + z)$, and $\pm 0.06(1 + z)$.

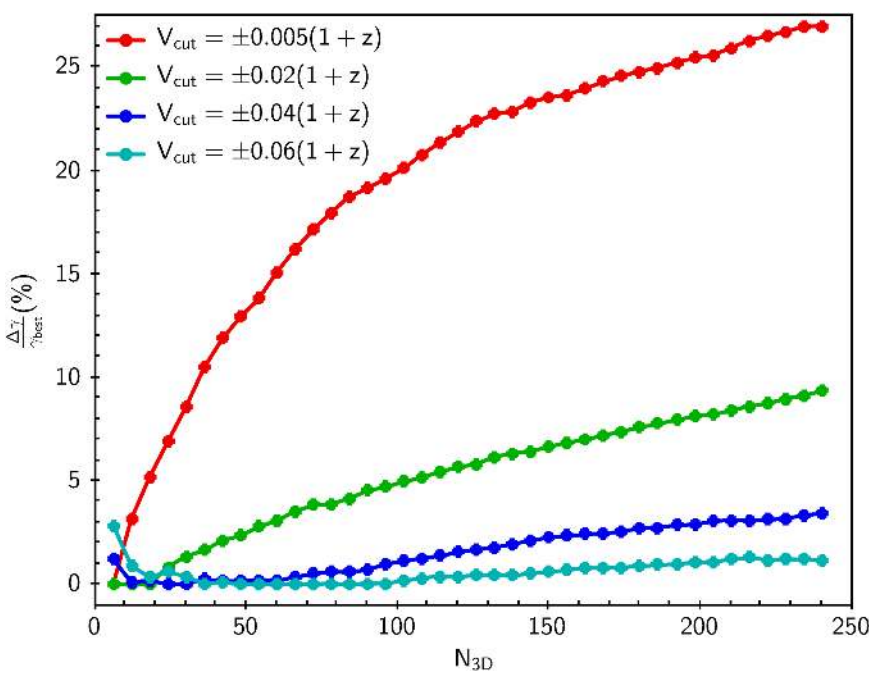

Figure 29. The difference in r_s between the two cases, optimized N_{2D} and $N_{2D} = N_{3D}$, normalized by the former. Red, green, blue, and cyan colors are for $V_{cut} = \pm 0.005(1 + z)$, $\pm 0.02(1 + z)$, $\pm 0.04(1 + z)$, and $\pm 0.06(1 + z)$, respectively.

z) and calculate the environment with different sizes of $V_{cut} = \pm 0.005(1 + z)$, $\pm 0.02(1 + z)$, $\pm 0.04(1 + z)$, and $\pm 0.06(1 + z)$, corresponding to red, green, blue, and cyan dots respectively. As can be seen, for the cases with size of $V_{cut} = \pm 0.02(1 + z)$, $\pm 0.04(1 + z)$, and $\pm 0.06(1 + z)$, their optimized results are quite similar although the best choices of N_{2D} are different. The ratio of the optimized N_{2D} to N_{3D}, roughly, is 1:10, 1:5, 1:2.5, and 1:1.7 for the cases with $V_{cut} = \pm 0.005(1 + z)$, $\pm 0.02(1 + z)$, $\pm 0.04(1 + z)$, and $\pm 0.06(1 + z)$, respectively. N_{2D} in the case of $V_{cut} = \pm 0.06(1 + z)$ is tripled compared to N_{2D} in the case of $V_{cut} = \pm 0.02(1 + z)$. Next we investigate how strong the effect of N_{2D} is on r_s. In Figure 29 we plot the difference between the two r_s, one evaluated by using the best choice of N_{2D} and the other evaluated from $N_{2D} = N_{3D}$, normalized by the former, as a function of N_{3D}. As can be seen, their maximum difference is only ~9%, 3%, and 3% in the cases with $V_{cut} = \pm 0.02(1 + z)$, $\pm 0.04(1 + z)$, and $\pm 0.06(1 + z)$, respectively. This suggests that if the size of V_{cut} is comparable to the photo-z uncertainty, the 2D local density measured with $N_{2D} \sim N_{3D}$ could be as good as that derived with the optimized N_{2D}.

REFERENCES

Albrecht, A., Bernstein, G., Cahn, R., et al. 2006, arXiv:astro-ph/0609591
Andrade, F., Esat, I., & Badi, M. 2001, JSV, 240, 909
Arnalte-Mur, P., Fernández-Soto, A., Martínez, V. J., et al. 2009, MNRAS, 394, 1631
Baldry, I. K., Balogh, M. L., Bower, R. G., et al. 2006, MNRAS, 373, 469
Balogh, M. L., Baldry, I. K., Nichol, R., et al. 2004, ApJL, 615, L101
Balogh, M. L., Navarro, J. F., & Morris, S. L. 2000, ApJ, 540, 113
Benson, A. J., Baugh, C. M., Cole, S., Frenk, C. S., & Lacey, C. G. 2000, MNRAS, 316, 107
Blanton, M. R., Eisenstein, D., Hogg, D. W., & Zehavi, I. 2006, ApJ, 645, 977
Bower, R. G., Benson, A. J., Malbon, R., et al. 2006, MNRAS, 370, 645
Brammer, G. B., van Dokkum, P. G., & Coppi, P. 2008, ApJ, 686, 1503
Brough, S., Croom, S., Sharp, R., et al. 2013, MNRAS, 435, 2903
Capak, P., Abraham, R. G., Ellis, R. S., et al. 2007, ApJS, 172, 284
Casertano, S., & Hut, P. 1985, ApJ, 298, 80
Cassata, P., Guzzo, L., Franceschini, A., et al. 2007, ApJS, 172, 270
Chiang, Y.-K., Overzier, R., & Gebhardt, K. 2014, ApJL, 782, L3
Cole, S., Lacey, C. G., Baugh, C. M., & Frenk, C. S. 2000, MNRAS, 319, 168
Cooper, M. C., Griffith, R. L., Newman, J. A., et al. 2012, MNRAS, 419, 3018
Cooper, M. C., Newman, J. A., Coil, A. L., et al. 2007, MNRAS, 376, 1445
Cooper, M. C., Newman, J. A., Croton, D. J., et al. 2006, MNRAS, 370, 198
Cooper, M. C., Newman, J. A., Madgwick, D. S., et al. 2005, ApJ, 634, 833
Curran, P. A. 2014, arXiv:1411.3816
Darvish, B., Sobral, D., Mobasher, B., et al. 2014, ApJ, 796, 51
Davis, M., Faber, S. M., Newman, J., et al. 2003, Proc. SPIE, 4834, 161
Elbaz, D., Daddi, E., Le Borgne, D., et al. 2007, A&A, 468, 33
Etherington, J., & Thomas, D. 2015, MNRAS, 451, 660
Faber, S. M., Phillips, A. C., Kibrick, R. I., et al. 2003, Proc. SPIE, 4841, 1657
Gallazzi, A., Bell, E. F., Wolf, C., et al. 2009, ApJ, 690, 1883
Gómez, P. L., Nichol, R. C., Miller, C. J., et al. 2003, ApJ, 584, 210
Grützbauch, R., Conselice, C. J., Varela, J., et al. 2011, MNRAS, 411, 929
Gunn, J. E., & Gott, J. R., III 1972, ApJ, 176, 1
Haas, M. R., Schaye, J., & Jeeson-Daniel, A. 2012, MNRAS, 419, 2133
Haines, C. P., La Barbera, F., Mercurio, A., Merluzzi, P., & Busarello, G. 2006, ApJL, 647, L21
Hogg, D. W., Blanton, M. R., Brinchmann, J., et al. 2004, ApJL, 601, L29
Kaiser, N., Burgett, W., Chambers, K., et al. 2010, Proc. SPIE, 7733, 77330E
Lagos, C. d. P., Bayet, E., Baugh, C. M., et al. 2012, MNRAS, 426, 2142
Lagos, C. D. P., Lacey, C. G., Baugh, C. M., Bower, R. G., & Benson, A. J. 2011, MNRAS, 416, 1566
Larson, R. B., Tinsley, B. M., & Caldwell, C. N. 1980, ApJ, 237, 692
Lemson, G., & Kauffmann, G. 1999, MNRAS, 302, 111
Lilly, S. J., Le Fèvre, O., Renzini, A., et al. 2007, ApJS, 172, 70
Lin, L., Capak, P. L., Laigle, C., et al. 2016, ApJ, 817, 97
Lin, L., Jian, H.-Y., Foucaud, S., et al. 2014, ApJ, 782, 33
LSST Science Collaboration, Abell, P. A., Allison, J., et al. 2009, arXiv:0912.0201
Malavasi, N., Pozzetti, L., Cucciati, O., Bardelli, S., & Cimatti, A. 2016, A&A, 585, A116
McCarthy, I. G., Frenk, C. S., Font, A. S., et al. 2008, MNRAS, 383, 593
Merson, A. I., Baugh, C. M., Helly, J. C., et al. 2013, MNRAS, 429, 556
Mihos, J. C., & Hernquist, L. 1994, ApJL, 431, L9
Miyazaki, S., Komiyama, Y., Nakaya, H., et al. 2012, Proc. SPIE, 8446, 84460Z
Moore, B., Katz, N., Lake, G., Dressler, A., & Oemler, A. 1996, Natur, 379, 613
Mostek, N., Coil, A. L., Cooper, M., et al. 2013, ApJ, 767, 89
Muldrew, S. I., Croton, D. J., Skibba, R. A., et al. 2012, MNRAS, 419, 2670
Newman, J. A., Cooper, M. C., Davis, M., et al. 2013, ApJS, 208, 5
Onaka, P., Tonry, J. L., Isani, S., et al. 2008, Proc. SPIE, 7014, 70140D
Quadri, R. F., Williams, R. J., Franx, M., & Hildebrandt, H. 2012, ApJ, 744, 88
Quilis, V., Moore, B., & Bower, R. 2000, Sci, 288, 1617
Schlagenhaufer, H. A., Phleps, S., & Sánchez, A. G. 2012, MNRAS, 425, 2099
Scoville, N., Arnouts, S., Aussel, H., et al. 2013, ApJS, 206, 3
Shattow, G. M., Croton, D. J., Skibba, R. A., et al. 2013, MNRAS, 433, 3314
Spearman, C. 1904, The American Journal of Psychology, 15, 72
Spergel, D. N., Verde, L., Peiris, H. V., et al. 2003, ApJS, 148, 175
Springel, V., White, S. D. M., Jenkins, A., et al. 2005, Natur, 435, 629
Tanaka, M., Goto, T., Okamura, S., Shimasaku, K., & Brinkmann, J. 2004, AJ, 128, 2677
The Dark Energy Survey Collaboration 2005, arXiv:astro-ph/0510346
Wilman, D. J., Zibetti, S., & Budavári, T. 2010, MNRAS, 406, 1701
York, D. G., Adelman, J., Anderson, J. E., Jr., et al. 2000, AJ, 120, 1579