Light-cone distortion of the clustering and abundance of massive galaxies at high redshifts

Mon. Not. R. Astron. Soc. **386,** 2323–2329 (2008) doi:10.1111/j.1365-2966.2008.13210.x

Light-cone distortion of the clustering and abundance of massive galaxies at high redshifts

Joseph A. Muñoz⋆ and Abraham Loeb⋆

Harvard–Smithsonian Center for Astrophysics, 60 Garden Street, MS 10, Cambridge, MA 02138, USA

Accepted 2008 March 7. Received 2008 February 15; in original form 2007 November 15

ABSTRACT

Observational surveys of galaxies are not trivially related to single-epoch snapshots from computer simulations. Observationally, an increase in the distance along the line of sight corresponds to an earlier cosmic time at which the properties of the surveyed galaxy population may change. The effect of observing a survey volume along the light cone must be considered in the regime where the mass function of galaxies varies exponentially with redshift. This occurs when the haloes under consideration are rare, that is either when they are very massive or observed at high redshift. While the effect of the light cone is negligible for narrow-band surveys of Lyα emitters, it can be significant for dropout surveys of Lyman-break galaxies (LBGs) where the selection functions of the photometric bands are broad. Since there are exponentially more haloes at the low-redshift end of the survey, the low-redshift tail of the selection function contains a disproportionate fraction of the galaxies observed in the survey. This leads to a redshift probability distribution for the dropout LBGs with a mean less than that of the photometric selection function (PHSF) by an amount of order the standard deviation of the PHSF. The inferred mass function of galaxies is then shallower than the true mass function at a single redshift with the abundance at the high-mass end being twice or more as large as expected. Moreover, the statistical moments of the count of galaxies calculated ignoring the light-cone effect deviate from the actual values.

Key words: galaxies: high-redshift – cosmology: theory.

1 INTRODUCTION

It has become common practice to analyse and interpret the observed abundance and distribution of high-redshift galaxies by approximating a limited survey volume to a single-epoch snapshot of the Universe (Ellis 2008). The observed data is then compared to theoretical predictions which were calculated for an idealized snapshot of this nature. However, in actual observations, an increase in the distance along the line of sight corresponds to an earlier cosmic time at which the properties of the surveyed galaxy population may change.

The 'snapshot approximation' is adequate for galaxy surveys at low redshifts, when galaxy haloes are common and their mass function is not evolving rapidly with cosmic time. At these low redshifts, a relatively small region of space spanning a narrow redshift range can still be sufficiently large to contain an adequate sample of these abundant objects. However, the validity of the approximation should be carefully examined at high redshifts when massive galaxies are rare and their abundance varies exponentially with redshift.

To illustrate the situation at high redshifts, let us consider two regions of the same shape centred at different redshifts and containing the same number of galaxy haloes of a particular mass. The region centred at the higher redshift will span a larger range in redshift for two reasons. First, since haloes of a given mass are rarer at a higher redshift, the higher redshift region must have a larger comoving size than the one at smaller redshift for each to contain the same number of haloes. Secondly, the same comoving distance corresponds to a larger redshift interval at higher redshift than at lower redshift {i.e. $\mathrm{d}z = [H(z)/c]\mathrm{d}\chi$, where χ is the comoving length and the Hubble parameter, $H(z)$, is an increasing function of z}. The difference between snapshot analysis (on a space-like hypersurface) and that along the light cone is becoming increasingly relevant with purported discoveries of very massive galaxies near $z = 6$ (Mobasher, Dickinson & Ferguson 2005) and as new surveys probe redshifts up to $z = 10$ (Bouwens et al. 2006; Iye et al. 2006; Stark et al. 2007a). Even at high redshifts, narrow-band surveys of Lyα emitters (LAEs) span such a small range of redshifts that they are unaffected by the exponential change in the mass function of haloes with redshift.

⋆E-mail: jamunoz@cfa.harvard.edu (JAM); aloeb@cfa.harvard.edu (AL)

However, the photometric selection functions (PHSFs) of the bands used in dropout surveys of Lyman-break galaxies (LBGs) can be fairly broad in redshift space (Bouwens & Illingworth 2006).

In this paper, we examine the significance of light-cone distortions on the inferred abundance and clustering properties of high-redshift galaxies in dropout surveys of LBGs. First, we describe the PHSFs used in dropout surveys in Section 2. Subsequently, we derive analytic formulae for the first and second statistical moments of the count of haloes in a given survey volume (Section 3) and consider the two-point correlation function of haloes on the light cone (Section 4). In Section 5, we review a simple model for high-redshift star-forming galaxies by Stark, Loeb & Ellis (2007b), which gives the luminosity of LBGs and LAEs contained in a halo of a given mass. We then use this model to calculate the quantitative difference between our light-cone formalism and the standard snapshot approach for various survey volumes (Section 6), exploring the dependence on cosmological parameters. Finally, we discuss the significance of our results in Section 7.

Unless otherwise stated, we assume a flat, ΛCDM cosmology with cosmological parameters $(\Omega_{\rm m},\ \Omega_\Lambda,\ \Omega_{\rm b},\ h,\ \sigma_8,\ \alpha,\ r) = (0.268,\ 0.732,\ 0.042,\ 0.704,\ 0.776,\ 0.947,\ 0.000)$ (Spergel, Bean & Doré 2007). All distance scales are comoving.

2 THE DROPOUT SELECTION FUNCTION

Dropout surveys at high redshifts ($z \gtrsim 6$) select LBGs by measuring a drop in flux shortward of the Lyα wavelength (due to absorption by intergalactic hydrogen). This requires comparing the observed flux in different photometric bands. The filter for each band is described by a profile that indicates how much light is transmitted at each wavelength. This transmission profile provides a probability distribution for the wavelength of a given photon that has passed through the filter. Since the edge of the Lyα absorption trough appears at a wavelength corresponding to the redshift of the observed galaxy, the filter profile can be expressed in redshift space as the PHSF for a given photometric band, which gives the distribution of the surveyed galaxies over redshift (Bouwens & Illingworth 2006). The volume of the survey is an integral over this function (Steidel et al. 1999). In this paper we focus on dropout selections in the *i*, *z* and *J* bands corresponding to the standard *Hubble Space Telescope* (*HST*) filters F775W, F850LP and F110W, respectively. The PHSFs for each band depend on the specific selection criteria chosen, but are roughly approximated by Gaussians (Bouwens, personal communication). We take the mean redshifts of the *i*-, *z*- and *J*-band PHSFs to be $\mu_z = 6.5,\ 7.4$ and 10, respectively, and their standard deviations to be $\sigma_z = 0.5,\ 0.5$ and 1.0. We also ignore, for simplicity, possible interlopers at lower redshifts whose spectra mimic those of LBGs at higher redshift as a result of dust attenuation.

Due to the evolution of the mass function of galaxy haloes within the survey volume, the redshift probability distribution (RPD) of a galaxy in the survey is not the same as the PHSF. Even though the contribution from galaxies in the Gaussian tail of the PHSFs is exponentially suppressed, the density of the rare haloes that contain the observed galaxies is theoretically expected to be exponentially higher towards the low-redshift end of the survey. The volume per redshift interval also changes within the survey since the area of the survey perpendicular to the line of sight and the comoving distance per redshift interval along the line of sight are both redshift dependent, but this is a small correction.

Ignoring the variation of the survey volume per redshift interval, the RPD for LBGs with luminosity at a wavelength of 1500 Å that is greater than L_{1500} within the volume observed in a given dropout band is given by

$$P_g(z) = \frac{n(z, >L_{1500}(M_{\rm halo}))\, {\rm e}^{-(z-\mu_z)^2/(2\sigma_z^2)}}{\int_0^\infty n(z, >L_{1500}(M_{\rm halo}))\, {\rm e}^{-(z-\mu_z)^2/(2\sigma_z^2)}\, {\rm d}z}, \quad (1)$$

where μ_z and σ_z are the mean and standard deviation of the given band, $n(z, M)$ is the mass function of haloes, and $L_{1500}(M_{\rm halo})$ is a relation for the luminosity of an LBG contained in a halo of mass $M_{\rm halo}$, which we describe in Section 5. Fig. 1 shows how the PHSF is multiplied by the mass function to generate the real RPD for *J* dropouts. In Section 6, we calculate the moments of this true distribution.

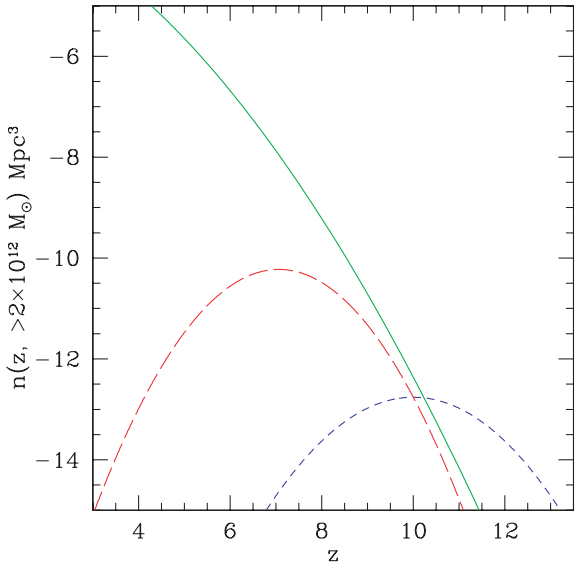

Figure 1. An example of how the PHSF (dashed) is multiplied by the mass function (solid) to yield the RPD (long-dashed). The PHSF shown is for the *J* band, and the Gaussian is normalized to the mass function at $z = 10$ for easy viewing. The mass function and resulting RPD were calculated for galaxy haloes with $M_{\rm halo} > 2 \times 10^{12}\ {\rm M}_\odot$.

3 MOMENTS OF A COUNT OF OBJECTS

In this section we derive and discuss the formulae for the statistical moments of a count of haloes above a given minimum mass, $M_{\rm min}$, in a given survey volume taking into account the variations of the mass and correlation functions along the light cone. We limit the discussion to the first and second moments, $\langle N\rangle$ and $\langle (N - \langle N\rangle)^2\rangle$.

We generally follow Peebles (1980) in calculating moments of a count of objects in a box but augment the derivation by allowing the mass and correlation functions to vary with redshift. We begin by dividing the box into infinitesimal units of size dV. The *i*th unit in the box contains N_i objects, so that the total number of objects in the box is $N = \sum_i N_i$. If every part of the box is viewed at the same cosmic time, then the average number of objects in each unit is obtained through the average number density of objects in the box: $\langle N_i\rangle = n{\rm d}V$. However, if the box is viewed at some distance from the observer, then each unit sits at a particular redshift, and the average number of objects in units at redshift z is related to the average number density of objects in the box at that redshift, $\langle N_{i,z}\rangle = n(z){\rm d}V$. We assume that d$V$ is small enough so that $N_i = \{1, 0\}$ and the statistics of galaxies within each unit is Poisson distributed, i.e. $\langle N_i^k\rangle = \langle N_i\rangle$ at each redshift for every k in $\mathbb{Z}$.

We can now calculate the moments of a count of objects, N, in the box. The first moment is given by

$$\langle N\rangle = \left\langle \sum_i N_i \right\rangle = \int n(z)\,\mathrm{d}V, \tag{2}$$

where the comoving volume element $\mathrm{d}V$ depends on the survey geometry. In the specific case of haloes above a given mass threshold $M_{\rm min}$, n is the mass function of haloes above $M_{\rm min}$. Equation (2) is precisely what one would expect from simply integrating the mass function over the survey volume as in Naoz, Noter & Barkana (2006).

The second moment is

$$\langle N^2\rangle = \left\langle \left(\sum_i N_i\right)^2 \right\rangle = \sum_i \left\langle N_i^2\right\rangle + \sum_i \sum_j \langle N_i\, N_j\rangle. \tag{3}$$

If there is an object in each of the disjoint units i and j, the product N_iN_j is equal to unity. Otherwise, the product is equal to zero. The probability of both units containing an object is

$$\mathrm{d}P = n_i n_j\,\mathrm{d}V_i\,\mathrm{d}V_j\,(1+\xi_{ij}), \tag{4}$$

where ξ_{ij} is the correlation between objects in units i and j. Thus, equation (3) reduces to

$$\langle N^2\rangle = \langle N\rangle + \langle N\rangle^2 + I_2, \tag{5}$$

where

$$I_2 = \int\int n(z_1)\,n(z_2)\,\xi(z_1, z_2, r_{1,2})\,\mathrm{d}V(z_1)\,\mathrm{d}V(z_2). \tag{6}$$

The variance of the count in the box can be expressed as

$$\frac{\langle N^2\rangle - \langle N\rangle^2}{\langle N\rangle^2} = \frac{1}{\langle N\rangle} + \frac{I_2}{\langle N\rangle^2} = \sigma_{\rm P} + \sigma_{\rm c}, \tag{7}$$

which can be thought of as the sum of a Poisson part, $\sigma_{\rm P}$, and a clustering part, $\sigma_{\rm c}$. In the specific case of haloes above a given mass threshold, $\xi(z_1, z_2, r_{1,2})$ in equation (6) is the correlation function between haloes with masses (possibly different) above $M_{\rm min}$ at different redshifts, and n is the mass function of such haloes.

4 THE LIGHT-CONE CORRELATION FUNCTION

While several analytic prescriptions exist for calculating the correlation function of haloes at different redshifts (Mo & White 1996; Porciani et al. 1998; Scannapieco & Barkana 2002), the correlation function measured in observations of LBGs is an average along the light cone over the survey volume. The observed correlation between each halo with mass greater than $M_{\rm min}$ and each other halo with mass greater than $M_{\rm min}$ is (Matarrese et al. 1997)

$$\xi^{\rm hh}_{LC,M}(>M_{\rm min}, r) = \langle N\rangle^{-2} \int\int \mathrm{d}V(z_1)\,\mathrm{d}V(z_2)\,\xi^{mm}(\overline{z}, r) \times n(z_1)\,n(z_2)\,b_{\rm eff}(z_1)\,b_{\rm eff}(z_2), \tag{8}$$

where

$$b_{\rm eff}(z) = \frac{\int_{M_{\rm min}}^{\infty} \mathrm{d}M\,\frac{\mathrm{d}n}{\mathrm{d}M}(z, M)\,b(z, M)}{n(z, >M_{\rm min})} \tag{9}$$

is the effective bias used to include haloes at all masses above $M_{\rm min}$, ξ^{mm} is the mass autocorrelation function, n is the halo mass function, b is the bias factor, r is the comoving distance between haloes and $\overline{z} = \overline{z}_M \equiv (z_1 + z_2)/2$.

We use the reformulation of the correlation function on the light cone given by Yamamoto & Suto (1999) which involves only a single integral:

$$\xi^{\rm hh}_{\rm LC,YS}(>M_{\rm min}, r) = \frac{\int_{z_{\rm min}}^{z_{\rm max}} n^2(z)\,\xi^{mm}(z, r)\,\mathrm{d}V(z)}{\int_{z_{\rm min}}^{z_{\rm max}} n^2(z)\,\mathrm{d}V(z)}. \tag{10}$$

While the mass function evolves extremely rapidly over the range of the PHSF, the evolution is minimal between two points separated by a distance, r, small enough to produce a non-negligible correlation. This is the key approximation made by Yamamoto & Suto (1999) and hold well even in our regime.

For the correlation function at different redshifts that appears in equation (6), we use

$$\xi(z_1, z_2, r_{1,2}) = \xi^{\rm hh}(>M_{\rm min}, z_1, z_2, r) = \xi^{mm}(r, z=0)\,D(z_1)\,D(z_2) \times b_{\rm eff}(z_2, M_{\rm min})\,b_{\rm eff}(z_1, M_{\rm min}), \tag{11}$$

where $D(z)$ is the linear growth factor. Using instead the measured correlation function along the light cone given in equation (10) [$\xi(z_1, z_2, r_{1,2}) = \xi^{\rm hh}_{\rm LC,YS}(>M_{\rm min}, r)$], results in a value of $\sigma_{\rm c}$ different by less than a per cent.

5 A MODEL FOR HIGH-REDSHIFT STAR-FORMING GALAXIES

To compare our results for the statistics of haloes to those observed, we need a way of equating the mass of a halo, $M_{\rm halo}$, with the luminosity of the observable galaxy it contains. In this section, we review a model given by Stark et al. (2007b) that prescribes such a transformation for LBGs and LAEs.

The SLE07 model associates LBGs and LAEs with merger-activated star formation in dark matter haloes. The ratio of baryonic to dark matter mass in these haloes is equal to the cosmic value, $\Omega_{\rm b}/\Omega_{\rm m}$. The efficiency with which the baryons are converted into stars, denoted by f, is a constant, $f = f_\star$, for haloes more massive than a critical value $M_{\rm halo,crit}$. However, for haloes below this mass, the feedback from supernovae suppresses star formation such that $f = f_\star (M_{\rm halo}/M_{\rm halo,crit})^{2/3}$. Modelling and low-redshift observations suggest that $M_{\rm halo,crit}$ corresponds to a velocity in the halo of $\sim$100 km s^{-1} (Dekel & Woo 2003; Kauffmann et al. 2003). We express the time-scale for star formation at z as the cosmic time, $t_{\rm H}(z)$, times the star formation duty cycle, $\epsilon_{\rm DC}$. $\epsilon_{\rm DC}$ gives the fraction of the Hubble time during which the star formation occurs. The average star formation rate is then

$$\dot{M}_\star(M_{\rm halo}) = \frac{f\,(\Omega_{\rm b}/\Omega_{\rm m})\,M_{\rm halo}}{t_{\rm H}(z)\,\epsilon_{\rm DC}}. \tag{12}$$

For LBGs, the luminosity per unit frequency at a wavelength of 1500 Å is given by

$$L_{1500} = 8.0 \times 10^{27}\,(\dot{M}_\star/{\rm M}_\odot\,{\rm yr}^{-1})\,{\rm erg\,s^{-1}\,Hz^{-1}}, \tag{13}$$

assuming a Salpeter initial mass function (IMF) of stars.

For LAEs with a low metallicity (1/20 solar) and a Salpeter IMF one gets $N_{\rm ip} = 4 \times 10^{53}$ ionizing photons emitted per M$_\odot$ of star formation per year. A fraction $1 - f_{\rm ip}$ of these photons do not escape from the galaxy and produce ions, two-thirds of the resulting

recombinations each produce a Lyα photon with energy $h\nu_{\rm Ly\alpha}$, and only a fraction $T_{\rm Ly\alpha}$ these photons escape into and pass through the intergalactic medium to be observed. The Lyα luminosity is then

$$L_{\rm Ly\alpha} = \frac{2}{3}\, h\, \nu_{\rm Ly\alpha}\, T_{\rm Ly\alpha}\, (1 - f_{\rm ip})\, N_{\rm ip}\, \dot{M}_{\star}, \tag{14}$$

where $\nu_{\rm Ly\alpha}$ is the frequency of the Lyα transition.

SLE07 fit the free parameters in their model ($f_\star$ and $\epsilon_{\rm DC}$) to observations at $z \sim 6$. For LBGs at $z \simeq 6$, the best-fitting values and 1σ errors are $f_\star = 0.16^{+0.06}_{-0.03}$ and $\epsilon_{\rm DC} = 0.25^{+0.38}_{-0.09}$. For LAEs at $z = 6.6$, $f_\star\ T_{\rm Ly\alpha} = 0.063^{+0.004}_{-0.018}$ and $\epsilon_{\rm DC} = 1.0^{+0.0}_{-0.5}$. These fit parameters are then used to determine the model at higher redshifts. We adopt this simple model with a fixed choice of its free parameters only as an illustrative example for relating the statistics of dark matter haloes to observed galaxies. All of the plots given as functions of halo mass in the subsequent sections can be easily related to galaxy luminosities in the context of any more complicated models for galaxy formation and evolution.

6 RESULTS

Next, we present the moments of the true RPD for *i*-, *z*- and *J*-dropout LGBs in Section 6.1. Having derived expressions for the moments of haloes counts in a survey volume and the correlation function for such haloes along the light cone in Section 6.2, we may compare these expressions quantitatively with those derived using a snapshot approach for various dropout surveys of LBGs in haloes of different masses. The fractional variation does not depend on the survey field of view but only on the variation along the line of sight. Thus, our results apply to a wide variety of surveys for LBGs including the Great Observatories Origins Deep Survey (GOODS), the *Hubble Ultra-Deep Field* (HUDF), and future surveys using the Subaru Multi-Object Infrared Camera and Spectrograph (MOIRCS) and the *HST* Wide Field Camera 3 (WFC3).

For our calculations, we use the mass function given by Sheth & Tormen (1999) and the bias factor in Sheth, Mo & Tormen (2001). Our results without the light-cone effect are produced by assuming that the entire volume exists at the mean redshift of the PHSF.

6.1 The redshift probability distribution

Since the haloes that host LBGs are rare at high redshifts and their abundance varies exponentially with redshift over the width of the PHSFs for dropout surveys, the RPD of LBGs in such a survey is biased towards lower redshifts. The low-redshift tail of a Gaussian PHSF exaggerates this effect. The true RPD, $P_g(z)$, is given by equation (1). The mean, variance and skewness of the RPD are given by

$$\mu_{z,g} = \int_0^\infty z\, P_g(z)\, {\rm d}z, \tag{15}$$

$$\sigma^2_{z,g} = \int_0^\infty (z - \mu_{z\,g})^2\, P_g(z)\, {\rm d}z \tag{16}$$

and

$$\gamma_{z,g} = \frac{\int_0^\infty (z - \mu_{z\,g})^3\, P_g(z)\, {\rm d}z}{\sigma^{3/2}_{z\,g}}. \tag{17}$$

The moments of the RPD for *i*-, *z*- and *J*-band dropouts are shown in Figs 2–4. The plots show their dependence on halo mass, the broadness of the Gaussian shapes assumed for the PHSFs, and the cosmological parameter σ_8.

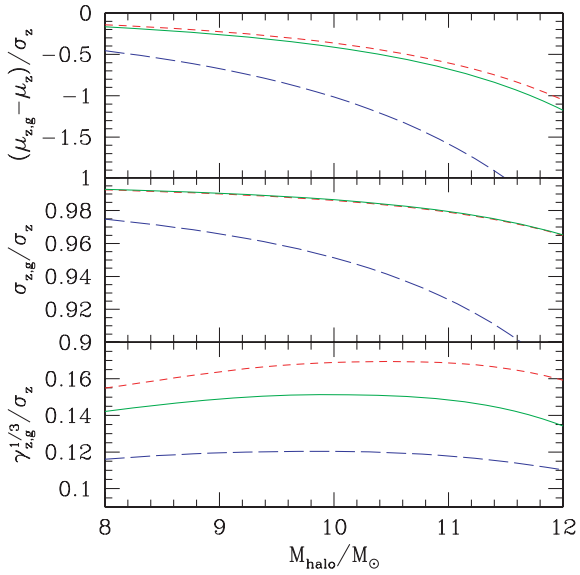

Figure 2. Moments of the RPD for the *i* (dashed), *z* (solid) and *J* (long-dashed) bands normalized by the standard deviation of the PHSF as a function of host halo mass. The upper panel shows the difference between the mean of the RPD and the PHSF, while the centre and lower panels plot the rms variation and asymmetry in the RPD, respectively.

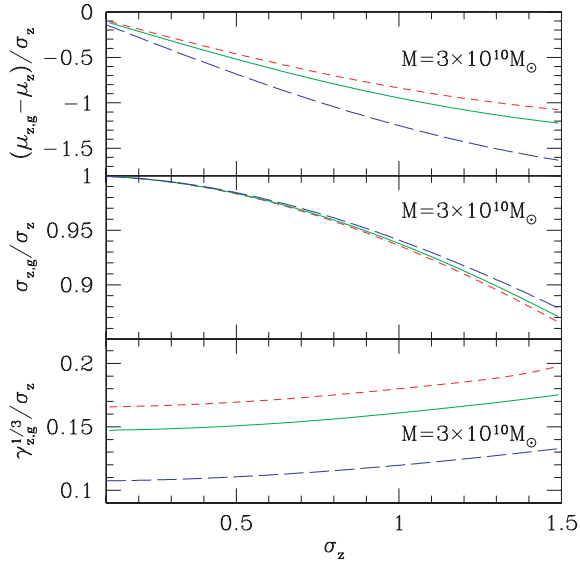

Figure 3. Same as Fig. 2 except that the moments are plotted as functions of the standard deviation assumed for the PHSF for a fiducial minimum halo mass of 3×10^{10} $M_\odot$.

The most important difference between the PHSF and the RPD for a given band is in their means. The exponentially varying mass function biases the survey towards lower redshifts. The mean of the RPD is offset from that of the PHSF by an amount on the order of the PHSF's standard deviation. Thus, in a *J*-dropout survey, the LBGs can be clustered around a redshift of as low as $z \sim 8.5$, depending on the luminosity of the galaxies considered, instead of being at $z \sim 10$.

Another important point to extract from Fig. 2 is that the LBGs are segregated by their halo mass (luminosity) through the survey. Since the mean of the RPD decreases monotonically with halo mass

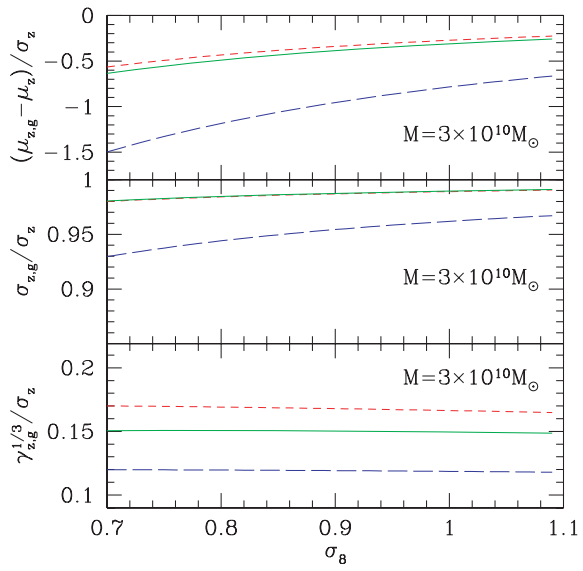

Figure 4. Same as Fig. 2 except that the moments are plotted as functions of the cosmological parameter σ_8 for a fiducial minimum halo mass of 3×10^{10} $M_\odot$.

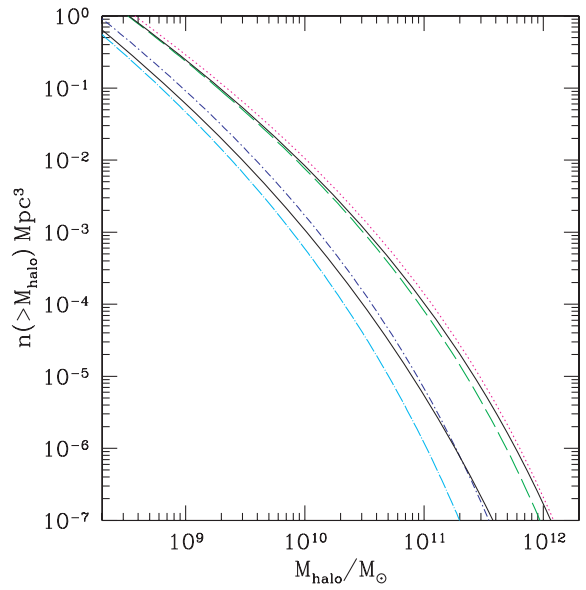

Figure 5. The mass function of haloes extracted from z-dropout (top solid) and J-dropout (bottom solid) surveys centred at $z = 7.4$ and 10, respectively. The underlying Sheth–Tormen mass functions at $z = 7$ (dotted), 7.4 (long-dashed), 9 (dot–dashed) and 10 (dot–long-dashed) are plotted for comparison.

(luminosity), the RPD for more luminous galaxies is shifted to lower redshifts than that of less luminous ones. While this series of RPDs overlaps, it is interesting to note that, for J dropouts, galaxies in haloes with $M \sim 10^8$ and 10^{11} $M_\odot$ are almost completely unmixed since the means of their RPDs are about separated by about one rms variation. This has important implications for trying to reconstruct the mass function from observations since different masses are being observed at different redshifts. Consider haloes in mass bins at two different masses, $M_{\rm HM}$ and $M_{\rm LM}$, where $M_{\rm HM} > M_{\rm LM}$, which are observed at two different redshifts, $z_{\rm HM}$ and $z_{\rm LM}$, such that $z_{\rm HM} < z_{\rm LM}$. The number of haloes with $M = M_{\rm HM}$ is higher than it would be if these haloes were seen at $z_{\rm LM}$. Thus, the resulting inferred mass function is shallower than the true one determined using haloes that are all at the same redshift. The effect is lessened by suppression due to the PHSF, however. The density of haloes of a given mass is less than the underlying density of such haloes where they are seen at $\mu_{z,g}$. Yet, their density is still higher than the underlying density at μ_z. The amplitudes of the mass function and the RPD at various redshifts can be compared for J-band haloes with $M_{\rm halo} > 2 \times 10^{12}$ $M_\odot$ in Fig. 1. Fig. 5 compares the extracted mass function with the underlying Sheth–Tormen mass function used to compute it.

The RPDs are slightly (by $\lesssim$10 per cent) narrower than the PHSFs. The normalized rms variation in the RPD as a function of the standard deviation of the PHSF is nearly identical for each band as are the equivalent plots as a function of halo mass for the i and z bands. This indicates that the slight narrowing is independent of the assumed standard deviation of the Gaussian PHSFs and depends only on the target redshift of a given band, i.e. the mean of its PHSF.

The asymmetry of the RPD is plotted in Figs 2–4 as $\gamma^{1/3}_{z,g}/\sigma_z$. The asymmetry is very small, on the order of tens of per cent. This degree of symmetry and the fact that the skewness is positive might seem counterintuitive since the exponentially varying mass function should bias the RPD towards lower redshifts. This bias, however, is manifested in the shift in the mean of the RPD away from that of the PHSF rather than in skewing the RPD. The Press–Schechter mass function with a simplified growth factor, $D(z) = 1/(1+z)$, has two dependencies on redshift, a linearly increasing factor and an exponentially decaying one that dominates at high redshift, $n = A(1+z)e^{-B(1+z)^2}$, where A and B are independent of redshift. The exponential decay factor, however, is simply the tail of a Gaussian, which when multiplied by the Gaussian PHSF yields another symmetric Gaussian. Only the subordinate linear factor contributes to the asymmetry.

Finally, although changing the value of σ_8 has a small effect on the value of the parameters we calculated, the difference as shown in Fig. 4 was not drastic; all of the basic results we have just presented remain unchanged. In fact, the asymmetry of the RPD is virtually unchanged when σ_8 is varied. This is consistent with the asymmetry being due only to the linear dependence on redshift, as discussed above, and not to the exponentially varying factor, which contains most of the mass function's dependence on σ_8.

6.2 The light-cone effect on moments of counts and the correlation function

As described above, the equations for the moments of the count of objects in the survey and their correlation are different if the light-cone effect is included. This effect is greatly enhanced by the wide PHSFs of the bands used in dropout surveys. For these surveys, the volume element in equations (2), (6) and (10) is replaced via

$$dV(z) \rightarrow \frac{e^{-(z-\mu_z)^2/(2\sigma_z^2)}}{\sqrt{2\pi\sigma_z^2}} \frac{dV(z)}{dz} dz. \quad (18)$$

Fig. 6 shows the results of including this effect on the mean count, clustering variance and correlation function of haloes containing the LGBs in i-, z- and J-band dropout surveys. While the value of each of these statistics varies depending on the field of view of the particular survey under consideration, the fractional effect is independent of field of view since a change in the area of the survey is orthogonal

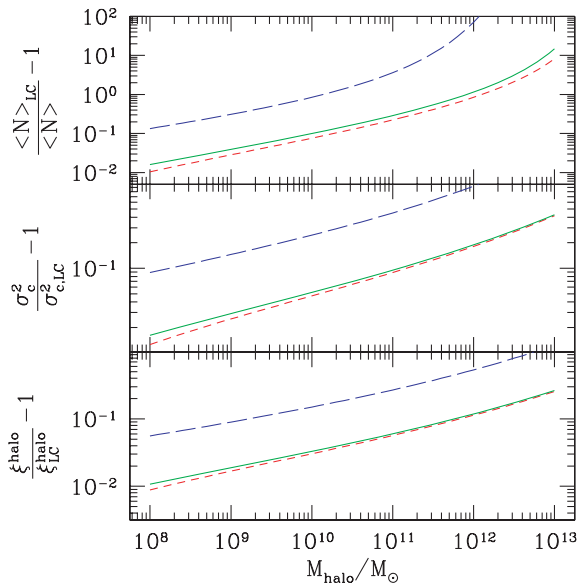

Figure 6. The fractional effect of the light cone on the mean number, clustering variance and correlation function of haloes in a dropout survey for LBGs in the i (dashed), z (solid) and J (long-dashed) bands. The clustering variance plotted here is $\sigma_c = I_2/\langle N\rangle^2$.

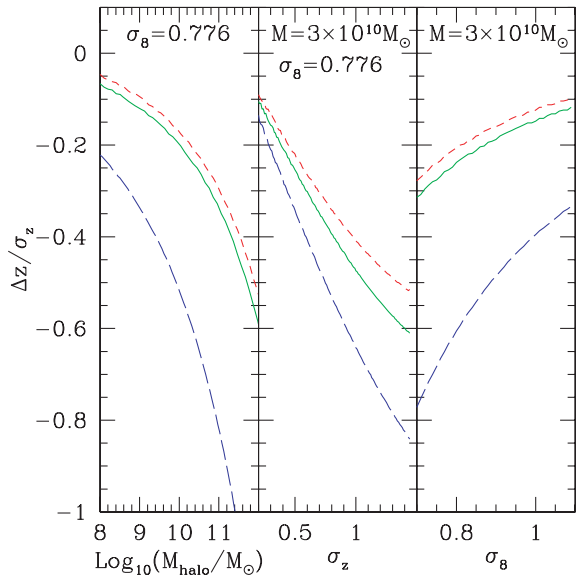

Figure 7. The difference between the redshift at which the underlying mass function gives the same density of haloes as observed in i-band (dashed), z-band (solid) and J-band (long-dashed) dropout surveys and the mean of the PHSF. The left-hand panel plots the difference as a function of minimum halo mass, while the centre and right-hand panels show it as functions of the assumed PHSF standard deviation and σ_8, respectively, for a fiducial minimum halo mass of 3×10^{10} $M_\odot$.

to the variation in the mass and correlation functions along the light cone.

For i and z dropouts, the light-cone effect can make an order unity difference in the mean number of objects and a difference of roughly 10 per cent in the clustering variance of the count and the correlation equation (10) at the high-mass end. The effect for J dropouts is much larger (by almost an order of magnitude) because the mass function has more evolution over the J band both due to the increased steepness of the mass function at higher redshift and because the PHSF is much broader for the J band. The variation in the effect with halo mass is due to the mass segregation discussed in the previous section. In particular, the mass dependence of the effect on $\langle N\rangle$ is manifested in the flattening of the mass function extracted from the survey shown in Fig. 5.

In associating our calculations for the halo correlation function with LBGs, we note that the haloes hosting LBGs do not constitute a fair sample of the entire halo population. Scannapieco & Thacker (2003) show through numerical simulations that these haloes have undergone substantial accretion in their recent past giving them an extra 'temporal' bias. While these simulations were performed at $z = 3$, there is as of yet no analytical method for predicting this extra bias at higher redshift. Since we compute only the fractional difference in the variance and correlation function, we are safe in ignoring this effect as long as the bias of LBGs over haloes does not vary much within the redshift range of the RPD.

Finally, in Section 6.1 we showed that the LBGs in dropout surveys, which are more numerous than would be calculated ignoring the light-cone effect, are distributed at lower redshifts than indicated by the PHSF. However, it is also interesting to ask at what redshift does the 'snapshot' calculation yield the same number of LBGs as the light-cone calculation. Hypothesizing a narrow-band survey for LBGs, which is equivalent to asking at what redshift does this hypothetical survey give the same density of LGBs as a dropout survey. The difference between this redshift and the mean of the PHSF is plotted in Fig. 7 for each band.

7 DISCUSSION

The variation of the mass function of haloes along the light cone within the volume of a dropout survey results in a mass segregation effect that is also manifested in their hosted LBGs. The observed sources are not at the mean of the PHSF but instead are distributed at lower redshifts (see Fig. 1). This effect applies to true LBG dropouts and ignores possible interlopers from lower redshifts whose spectra mimic those of higher redshift LGBs. The mass segregation results in different measured statistics of LBGs from those expected from theory or simulations along a space-like slice (snapshot) through the universe at the 'mean survey redshift'. In particular, the mass/luminosity function extracted from such a survey is shallower than the underlying mass/luminosity function because of the different strengths of the light-cone effect on haloes of different masses (LBGs of different luminosities). This flattening of the mass function is particularly important in the J band because of its larger width in redshift space (see Fig. 5).

Without spectroscopic measurements to confirm the redshifts of a sample of LBGs, the mass segregation effect presents an added complication in continuing efforts to determine the effect of high-redshift LBGs on reionization (Nagamine et al. 2006; Stark et al. 2007b) and the microwave background (Babich & Loeb 2007), or efforts to measure the high-redshift evolution of the star formation rate (Sawicki & Thompson 2006; Ellis 2008). Reionization is a highly inhomogeneous process, and so the light-cone effect on the correlation function is also particularly relevant in that context (Barkana & Loeb 2004; Furlanetto & Loeb 2005). The light-cone effect on the measured correlation function of LBGs would also be important in attempts to use the clustering properties of LBGs to infer the masses of their host haloes at higher redshift.

Dow-Hygelund et al. (2007) made an effort to follow-up spectroscopically on LBGs from i-band dropout surveys near $z \simeq 6$ to measure their redshifts precisely. However, they were only able to

confirm redshifts on six LBGs in their sample, a number insufficient to trace the RPD in a statistically significant way. Ando et al. (2004, 2007) perform similar studies on LBGs near $z \simeq 5$, but the sample size they obtained was also not sufficiently large for this purpose. Ideally, one would like to perform this type of analysis on z- or J-band dropouts with the goal of tracing the RPD, but this would be very difficult both because of the present lack of candidates and because of the high integration times necessary with current technology.

ACKNOWLEDGMENTS

We would like to thank Rychard Bouwens and Dan Stark for useful discussions. JAM acknowledges support from a National Science Foundation Graduate Research Fellowship. This research was supported in part by Harvard University funds.

REFERENCES

Ando M., Ohta K., Iwata I., Watanabe C., Tamura N., Akiyama M., Aoki K., 2004, ApJ, 610, 635
Ando M., Ohta K., Iwata I., Akiyama M., Aoki K., Tamura N., 2007, PASJ, in press (arXiv:0705.1145)
Babich D., Loeb A., 2007, MNRAS, 374, L24
Barkana R., Loeb A., 2004, ApJ, 609, 474
Bouwens R., Illingworth G., 2006, New Astron. Rev., 50, 152
Bouwens R., Illingworth G., Blakeslee J., Franx M., 2006, ApJ, 653, 53
Dekel A., Woo J., 2003, MNRAS, 344, 1131
Dow-Hygelund C. C. et al., 2007, ApJ, 660, 47
Ellis R. S., 2008, in Loeb A., Ferrara A., Ellis R. S., eds, Saas-Fee Advanced Course 36, First Light in the Universe, p. 259
Furlanetto S. R., Loeb A., 2005, ApJ, 634, 1
Iye M. et al., 2006, Nat, 443, 186
Kauffmann G. et al., 2003, MNRAS, 341, 54
Matarrese S., Coles P., Lucchin F., Moscardini L., 1997, MNRAS, 286, 115
Mo H. J., White S. D. M., 1996, MNRAS, 282, 347
Mobasher B., Dickinson M., Ferguson H. C., 2005, ApJ, 635, 832
Nagamine K., Cen R., Furlanetto S. R., Hernquist L., Night C., Ostriker J. P., Ouchi M., 2006, New Astron. Rev., 50, 29
Naoz S., Noter S., Barkana R., 2006, MNRAS, 373, L98
Peebles P. J. E., 1980, The Large-scale Structure of the Universe. Research supported by the National Science Foundation. Princeton Univ. Press, Princeton, NJ, p. 435
Porciani C., 1997, MNRAS, 290, 639
Porciani C., Matarrese S., Lucchin F., Catelan P., 1998, MNRAS, 298, 1097
Sawicki M., Thompson D., 2006, ApJ, 648, 299
Scannapieco E., Barkana R., 2002, ApJ, 571, 585
Scannapieco E., Thacker R. J., 2003, ApJ, 590, L69
Scannapieco E., Thacker R. J., 2005, ApJ, 619, 1
Sheth R. K., Tormen G., 1999, MNRAS, 308, 119
Sheth R. K., Mo H. J., Tormen G., 2001, MNRAS, 323, 1
Spergel D. N., Bean R., Doré O., 2007, ApJS, 170, 377
Stark D. P., Ellis R. S., Richard J., Kneib J.-P., Smith G. P., Santos M. R., 2007a, ApJ, 663, 10
Stark D. P., Loeb A., Ellis R. S., 2007b, preprint (astro-ph/0701882)
Steidel C. C., Adelberger K. L., Giavalisco M., Dickinson M., Pettini M., 1999, ApJ, 519, 1
Yamamoto K., Suto Y., 1999, ApJ, 517, 1

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