Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **472,** 1995–2008 (2017)
Advance Access publication 2017 September 6

doi:10.1093/mnras/stx1884

Dark-ages reionization and galaxy formation simulation–XI. Clustering and halo masses of high redshift galaxies

Jaehong Park,[1,3★] Han-Seek Kim,[1] Chuanwu Liu,[1] Michele Trenti,[1] Alan R. Duffy,[2] Paul M. Geil,[1] Simon J. Mutch,[1] Gregory B. Poole,[1] Andrei Mesinger[3] and J. Stuart B. Wyithe[1]

[1]*School of Physics, The University of Melbourne, Parkville, VIC 3010, Australia*
[2]*Centre for Astrophysics and Supercomputing, Swinburne University of Technology, PO Box 218, Hawthorn, VIC 3122, Australia*
[3]*Scuola Normale Superiore, Piazza dei Cavalieri 7, I-56126 Pisa, Italy*

Accepted 2017 July 21. Received 2017 July 10; in original form 2017 March 15

ABSTRACT

We investigate the clustering properties of Lyman-break galaxies (LBGs) at $z \sim 6-8$. Using the semi-analytical model MERAXES constructed as part of the dark-ages reionization and galaxy-formation observables from numerical simulation (DRAGONS) project, we predict the angular correlation function (ACF) of LBGs at $z \sim 6-8$. Overall, we find that the predicted ACFs are in good agreement with recent measurements at $z \sim 6$ and $z \sim 7.2$ from observations consisting of the Hubble eXtreme Deep Field, the *Hubble Ultra Deep Field* and cosmic sssembly near-infrared deep extragalactic legacy survey field. We confirm the dependence of clustering on luminosity, with more massive dark matter haloes hosting brighter galaxies, remains valid at high redshift. The predicted galaxy bias at fixed luminosity is found to increase with redshift, in agreement with observations. We find that LBGs of magnitude $M_{\rm AB(1600)} < -19.4$ at $6 \lesssim z \lesssim 8$ reside in dark matter haloes of mean mass $\sim 10^{11.0}$–$10^{11.5}$ $M_\odot$, and this dark matter halo mass does not evolve significantly during reionisation.

Key words: galaxies: haloes – galaixes: high-redshift – dark ages, reionization, first stars – large-scale structure of Universe – cosmology: theory.

1 INTRODUCTION

Clustering of galaxies allows us to probe the large-scale structure of the Universe as a biased tracer of the density field, and galaxy formation physics by providing a measure of dark matter halo mass. The formation and evolution of dark matter haloes are described by analytic models (e.g. Mo & White 1996; Cooray & Sheth 2002) and by *N*-body simulations (e.g. Springel et al. 2005). Galaxies are thought to form inside those dark matter haloes (White & Frenk 1991; Cole 1991), but we are still far from having a complete understanding of the galaxy formation process, which involves complicated non-linear physical processes (e.g. Baugh 2006; Benson 2010; Schaye et al. 2015).

Accurate measurements of the clustering of galaxies in the local Universe have determined the clustering dependence on luminosity (e.g. Norberg et al. 2001, 2002; Zehavi et al. 2005, 2011). These observational results provide strong constraints on theoretical predictions for galaxy properties (e.g. Henriques et al. 2015; Schaye et al. 2015; Lacey et al. 2016).

At higher redshift, galaxies selected by the Lyman-break technique (LBGs) are the most extensively studied sources (e.g. Giavalisco 2002). Galaxies at a specific redshift can be selected using a colour selection criteria designed to detect spectral features of star-forming galaxies corresponding to absorption of the rest-frame far-ultraviolate (UV) emission (below 1216 Å) by neutral hydrogen. Since the work of Steidel & Hamilton (1993) and Steidel et al. (1996) at $z \sim 3$, this technique has proved to be an effective and efficient method of discovery. This technique has now been extended to detect galaxies up to $z \sim 10$ (e.g. Bradley et al. 2012; Oesch et al. 2012; McLure et al. 2013; Duncan et al. 2014; Zitrin et al. 2014; Bouwens et al. 2015, 2016), providing the dominant source of information about galaxies during reionisation.

Over the past decade, clustering of LBGs has been measured in the redshift range $z \sim 3$–7 (Ouchi et al. 2005; Cooray & Ouchi 2006; Kashikawa et al. 2006; Lee et al. 2006; Hildebrandt et al. 2009; McLure et al. 2009; Harikane et al. 2016). Recently, Bouwens et al. (2015) identified LBGs in the redshift range $z \sim 4$–10 in a combined survey field consisting of the Hubble eXtreme Deep Field (XDF), the *Hubble Ultra Deep Field* (HUDF) and cosmic assembly near-infrared deep extragalactic legacy survey (CANDELS) field. Using the sample of Bouwens et al. (2015), Barone-Nugent et al.

★ E-mail: jaehong.park@sns.it

(2014) measured the angular correlation function (ACF) of LBGs in the redshift range $z \sim 4$–7.2. This includes the first measurement of LBG clustering at $z \sim 7.2$, using combined samples with $z \sim 7$ LBGs and $z \sim 8$ LBGs. Since Barone-Nugent et al. (2014) measured the ACF in each survey field independently, this measurement gives us an estimate of sample variance by comparing the results from different fields, while the measured ACF from the Hubble XDF allows us to investigate the clustering of the fainter LBGs. More recently, Harikane et al. (2016) measured the ACF of LBGs in the redshift range $z \sim 4$–7 using a combined data set from HUDF and CANDELS fields, and from XMM and GAMA09h fields of Subaru/Hyper Suprime-Cam (HSC) observations. They estimated the dark matter halo mass using Halo Occupation Distribution (HOD) modelling.

Recently, based on the development of observations, semi-analytical models (e.g. Samui, Srianand & Subramanian 2007; Samui 2014; Jose, Srianand & Subramanian 2014; Lacey et al. 2016; Liu et al. 2016; Jose et al. 2017) and hydrodynamic simulations (e.g. Ocvirk et al. 2016; Waters et al. 2016) predict UV luminosity functions of LBGs up to $z \sim 6$–10. Although the predicted luminosity functions provide the abundance of galaxies, they do not show the spatial distribution of galaxies. Comparing the model prediction of clustering with observations therefore provides additional insight into high redshift galaxies. Previous studies have been restricted to comparison of the predictions from theoretical models with observational measurements (e.g. Kashikawa et al. 2006; Jose et al. 2013) at redshifts less than $z \sim 5$ due to the insufficient number of LBGs observed at $z > 5$. Recently, Waters et al. (2016) predicted the galaxy bias from the real-space correlation function at higher redshifts of $z = 8$–10 but did not compare their result with measured ACFs.

In this paper, we investigate the clustering properties of LBGs at $z \sim 6$–8. We use the semi-analytical galaxy formation model, MERAXES (Mutch et al. 2016), which is designed to study galaxy formation during the Epoch of Reionisation (EoR). MERAXES is able to describe the luminosity function, stellar mass function and their evolution (Liu et al. 2016). We predict the ACF of LBGs selected from the model, providing model predictions for clustering of LBGs up to $z \sim 8$. We compare these model predictions with the clustering measured from Barone-Nugent et al. (2014), and also compare with the results from Harikane et al. (2016).

We begin in Section 2 by briefly describing MERAXES. In Section 3, we present the methodology used to select LBGs, show a resulting luminosity function, and describe how to compute ACFs in the model. We present the predictions for the clustering properties and compare the model predictions with observations in Section 4. Then, we conclude in Section 5. Throughout the paper we use apparent magnitudes in the observers frame in the AB system. Where we refer to the UV magnitude, this corresponds to the rest-frame 1600 Å AB magnitude. We employ a standard spatially flat Λ cold dark matter (CDM) cosmology based on Planck 2015 result (Planck Collaboration et al. 2016): (h, $\Omega_{\rm m}$, $\Omega_{\rm b}$, Ω_Λ, σ_8, $n_{\rm s}$)=(0.678, 0.308, 0.0484, 0.692, 0.815, 0.968).

2 THE MODEL

In this section we summarize the model used in this study. In Section 2.1, we briefly introduce the model MERAXES. Then we describe how to compute magnitudes of galaxies in Section 2.2. For further details, interested readers are referred to Mutch et al. (2016) for MERAXES and to Liu et al. (2016) for the UV luminosity function.

2.1 MERAXES

MERAXES is the semi-analytical model (Mutch et al. 2016) constructed as part of the dark-ages reionisation and galaxy-formation observables from numerical simulations[1] (DRAGONS) project. DRAGONS integrates a semi-analytical model with a semi-numerical model (Mesinger, Furlanetto & Cen 2011) in order to self-consistently simulate the reionisation and galaxy formation processes.

MERAXES computes the formation and evolution of galaxy properties in the redshift range $z \geq 5$. MERAXES has been developed based on the model of Croton et al. (2006) and extended in Guo et al. (2011). We implement MERAXES within the *N*-body dark matter simulation *Tiamat* (see Poole et al. 2016 for more details of *Tiamat*). In the fiducial *Tiamat* simulation the particle mass is $3.89 \times 10^6\,{\rm M}_\odot$, with a side length of 100 Mpc. *Tiamat* provides 100 output snapshots between $z = 35$–5 with a cadence of 11.1 Myr. This high cadence is necessary because the galaxy dynamical time at $z \gtrsim 6$ becomes comparable to the lifetime of massive stars. This allows us to compute time-resolved supernova feedback.

We describe basic differences of MERAXES to traditional semi-analytical models below (see Mutch et al. 2016 for more details of MERAXES).

(i) Merger trees: the merger trees are constructed 'horizontally'. During the EoR, ionizing photons from galaxies tens of Mpc away can ionize the intergalactic medium (IGM), affecting subsequent galaxy formation processes. MERAXES takes into account radiation from galaxies which are spatially associated with each other at each snapshot.

(ii) Delayed supernova feedback: MERAXES follows the parametrization of Guo et al. (2013) for the efficiency of the supernova feedback. However, the *N*-body simulation (*Tiamat*) on which MERAXES is run has one snapshot per ∼11.1 Myr, providing much higher time resolution than many semi-analytical models. This high time resolution is required at high redshift where the dynamical time in galaxies becomes smaller than the lifetime of massive stars. The lifetime of the massive stars corresponds to ∼4 snapshots in this model, and MERAXES adopts a delayed supernova feedback scheme. The total amount of supernova energy released by a galaxy at a snapshot is computed by tracking the total mass of stars formed in each galaxy for the last 40 Myr.

(iii) Reionisation: to model the reionisation process, MERAXES uses a modified 21CMFAST (Mesinger et al. 2011; Sobacchi & Mesinger 2013) algorithm. The criterion to find ionized regions can be written as

$$\xi \frac{m_*(r)}{M_{\rm tot}} \geq 1, \tag{1}$$

where $m_*(r)$ is the integrated stellar mass within radius r, $M_{\rm tot}$ is the total mass within r and ξ is an H II ionizing efficiency.

From equation (1), MERAXES computes a global ionization structure for the simulation volume. Then, MERAXES computes the value of the baryon fraction modifier, $f_{\rm mod}$, (Sobacchi & Mesinger 2013)

$$f_{\rm mod} = 2^{-M_{\rm filt}/M_{\rm vir}}, \tag{2}$$

where $M_{\rm filt}$ is the filtering mass at $f_{\rm mod} = 0.5$, given by

$$M_{\rm filt} = M_0 J_{21}^a \left(\frac{1+z}{10}\right)^b \left[1 - \left(\frac{1+z}{1+z_{\rm ion}}\right)^c\right]^d, \tag{3}$$

[1] http://dragons.ph.unimelb.edu.au

where J_{21} is the local ionizing intensity and $z_{\rm ion}$ is the redshift at which a halo was first ionized. The values of parameters are (M_0, a, b, c, d) = (2.8 × 10^9 $M_\odot$, 0.17, −2.1, 2.0, 2.5) as found by Sobacchi & Mesinger (2013). Based on the ionized structure and $z_{\rm ion}$, each halo uses the baryon fraction modifier, $f_{\rm mod}$, to compute the infalling baryonic mass.

(iv) Baryonic infall: ionizing UV background radiation heats the IGM and raises the local Jeans mass, reducing the fraction of the baryonic infall, $f_{\rm b}$ (Dijkstra et al. 2004). MERAXES parameterises this reduction using a baryon fraction modifier, $f_{\rm mod}$. The infalling baryonic mass into the dark matter haloes is

$$m_{\rm infall} = f_{\rm mod} f_{\rm b} M_{\rm vir} - \sum_{i=0}^{N_{\rm gal}-1} m^i_* + m^i_{\rm cold} + m^i_{\rm hot} + m^i_{\rm ejected}, \quad (4)$$

where $M_{\rm vir}$ is the mass of the halo, $N_{\rm gal}$ is the number of galaxies in the dark matter halo and the baryon fraction modifier has a range, $0 \leq f_{\rm mod} \leq 1$. m_*, $m_{\rm cold}$, $m_{\rm hot}$ and $m_{\rm ejected}$ are the stellar mass, cold gas mass, hot gas mass and the ejected gas mass from the dark matter halo, respectively.

2.2 Prediction properties of model galaxies

2.2.1 *Lyman α absorption*

To compute the intrinsic luminosity of each model galaxy, we start by building a star-formation history as a function of time by tracing all progenitors of a star, and calculating the intrinsic stellar luminosity. We use STARBURST99 (Leitherer et al. 1999, 2014) to model stellar energy distributions (SED), following the same methodology of Liu et al. (2016). The UV radiation from galaxies is absorbed by neutral hydrogen in the IGM. We calculate this attenuation by adopting an effective Ly α absorption optical depth from Fan et al. (2006). Fan et al. (2006) found that the effective optical depth evolves rapidly, $\tau_{\rm eff} \propto (1+z)^{10.9}$, at $z = 5.5$–6.3. For simplicity, we use this relation for all redshifts at $z \geq 5.5$. Since the observed Ly α flux vanishes at $z > 6$, this assumption does not affect LBG selections (see Liu et al. 2016 for more details).

2.2.2 *Dust attenuation*

Following the discussion in Liu et al. (2016) the rest-frame UV continuum for a galaxy can be written as

$$f_\lambda \propto \lambda^\beta, \quad (5)$$

where f_λ is the flux density per wavelength interval and β is the UV continuum slope. Since the amount of dust attenuation increases with shorter wavelengths, the dust-attenuation makes the continuum slope steepen. UV flux attenuated by dust grains within galaxies can be parametrized as

$$F_{\rm o}(\lambda) = F_{\rm i} 10^{-0.4\, A_\lambda}, \quad (6)$$

where $F_{\rm o}$ and $F_{\rm i}$ are the observed and intrinsic continuum flux densities, and A_λ is the change in magnitude at rest-frame wavelength λ. To compute the dust-attenuated UV continuum slope, β, we use the relation between the observed UV continuum, β, and the UV dust attenuation (Meurer, Heckman & Calzetti 1999)

$$A_{1600} = 4.43 + 1.99\,\beta, \quad (7)$$

where A_{1600} is the dust attenuation at 1600 Å. The values of β can be obtained from observations. Bouwens et al. (2014) found a piece-wise linear relation between the mean of β and $M_{\rm AB,1600}$ at $z \sim 4$–6

$$\beta = \begin{cases} \frac{{\rm d}\beta}{{\rm d}M_{\rm AB,1600}}(M_{\rm AB,1600} + 18.8) + \beta_{M_{\rm AB,1600}=-18.8}, \\ \qquad\qquad M_{\rm AB,1600} \leq 18.8, \\ -0.08(M_{\rm AB,1600} + 18.8) + \beta_{M_{\rm AB,1600}=-18.8}, \\ \qquad\qquad M_{\rm AB,1600} > 18.8, \end{cases} \quad (8)$$

where ${\rm d}\beta/{\rm d}M_{\rm AB,1600}$ and $\beta_{M_{\rm AB,1600}=-18.8}$ are taken from table 4 of Bouwens et al. (2014). For galaxies at $z \sim 7$–8, they also found a linear relation

$$\beta = \frac{{\rm d}\beta}{{\rm d}M_{\rm AB,1600}}(M_{\rm AB,1600} + 19.5) + \beta_{M_{\rm AB,1600}=-19.5}, \quad (9)$$

where ${\rm d}\beta/{\rm d}M_{\rm AB,1600}$ and $\beta_{M_{\rm AB,1600}=-19.5}$ are taken from table 3 of Bouwens et al. (2014). We assume the values of β are distributed with Gaussian standard deviation of 0.35 (Bouwens et al. 2014). Using equation (7) we obtain the relation between the mean dust attenuation, $\langle A_{1600} \rangle$, and the intrinsic UV luminosity at 1600 Å. These luminosities are converted to the intrinsic rest-frame magnitude, $M^i_{\rm AB,1600}$, using SED with tophat bands of 100 Å bandwidth.

At other wavelengths the dust attenuation can be written as

$$A_\lambda = E(B-V)k(\lambda), \quad (10)$$

where $E(B-V)$ is the colour excess. The dust reddening curve $k(\lambda)$ is given by Calzetti et al. (2000)

$$\beta = \begin{cases} 2.659\left(-2.156 + \frac{1.509}{\lambda} - \frac{0.198}{\lambda^2} + \frac{0.011}{\lambda^3}\right) + R_V, \\ \qquad\qquad 0.12\ \mu{\rm m} \leq \lambda < 0.63\ \mu{\rm m}, \\ 2.659\left(-1.857 + \frac{1.040}{\lambda}\right) + R_V, \\ \qquad\qquad 0.63\ \mu{\rm m} \leq \lambda < 2.20\ \mu{\rm m}, \end{cases} \quad (11)$$

where λ is the rest-frame wavelength in units of μm and $R_V = 4.05 \pm 0.80$ is the effective obscuration. For wavelength $\lambda < 0.12$ μm, we extrapolate the reddening curve.

3 LBGS IN THE MODEL

In this section we describe how we select model LBGs. Then, we compare the predicted results with observations.

3.1 Selecting LBGs

To select model LBGs we use a similar method to that described in Liu et al. (2016). They applied the colour selection criteria from Bouwens et al. (2015) to galaxies generated from MERAXES. The selection criteria for LBGs at $z \sim 6$ are

$$(i_{775} - z_{850} > 1.0) \wedge (Y_{105} - H_{160} < 1.0) \wedge$$
$$(i_{775} - z_{850}) > 0.78(Y_{105} - H_{160}) + 1.0,$$
$$\text{(not in } z \sim 7 \text{ selection)}, \quad (12)$$

for LBGs at $z \sim 7$ are

$$(z_{850} - Y_{105} > 0.7) \wedge (J_{125} - H_{160} < 0.45) \wedge$$
$$(z_{850} - Y_{105} > 0.8(J_{125} - H_{160}) + 0.7),$$
$$\text{(not in } z \sim 8 \text{ selection)}, \quad (13)$$

and for LBGs at $z \sim 8$ are

$$(Y_{125} - J_{125} > 0.45) \wedge (J_{125} - H_{160} < 0.5) \wedge$$
$$(Y_{105} - J_{125} > 0.75(J_{125} - H_{160}) + 0.525), \quad (14)$$

Table 1. Flux limits and areas of the individual survey fields. Each magnitude limit is quoted as a 5σ depth (Bouwens et al. 2015). The units of area is arcsec^2. The last two columns represent the number of LBGs at each redshift (Barone-Nugent et al. 2014).

Field	Area	i_{775}	z_{850}	Y_{105}	J_{125}	H_{160}	$z\sim6$	$z\sim7.2$
XDF	4.7	29.8	29.2	29.7	29.3	29.4	104	149
HUDF09-1	4.7	28.5	28.4	28.3	28.5	28.3	38	52
HUDF09-2	4.7	28.8	28.8	28.6	28.9	28.7	36	54
GS-Deep	64.5	27.5	27.3	27.5	27.8	27.5	203	134
GS-Wide	34.2	27.5	27.1	27.0	27.1	26.8	41	–
GS-ERS	40.2	27.2	27.1	27.0	27.6	27.4	62	64
GN-Deep	62.9	27.3	27.3	27.3	27.7	27.5	197	220
GN-Wide2	60.9	27.2	27.2	26.2	26.8	26.7	51	–

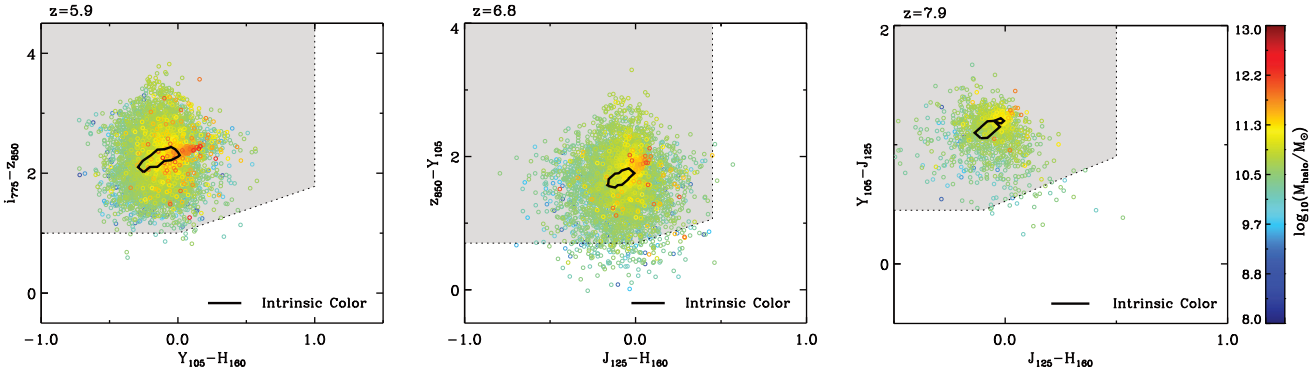

Figure 1. Predicted colour–colour diagrams showing colours of model galaxies at mean redshifts of 5.9 (left), 6.8 (central) and 7.9 (right), corresponding to the colour selection criteria. Different colour symbols correspond to the host dark matter halo mass as indicated by the colour bar. In each panel, thick contours show the distribution of intrinsic colours, i.e. without photometric errors, at the mean redshifts. All contours enclose 99.4 per cent of model galaxies. Grey shaded regions and dotted lines represent the colour selection criteria adopted from Bouwens et al. (2015). The panels show colours of model galaxies brighter than the XDF flux limit. Compared to the predicted distribution of intrinsic colours, we can see the noticeable scatter for the colours of model galaxies.

where $\wedge$ represents the logical AND symbol. i_{775}, z_{850}, Y_{105}, J_{125} and H_{160} correspond to the magnitudes of F775W, F850LP, F105W, F125W and F160W bands in ACS and WFC3/IR, respectively. Together with the above criteria, Bouwens et al. (2015) adopted additional criteria to exclude low redshift interlopers. We checked that the additional criteria do not change results of selecting model LBGs.

Before using the colour selection criteria, we consider two photometric conditions to mimic the observations. First, we take into account the photometric error for a predicted magnitude, which may affect the clustering signal for faint galaxies (Park et al. 2016). We obtain the apparent magnitude including the photometric error using

$$m' = -2.5\log_{10}(10^{-0.4\times m} + \text{noise}), \quad (15)$$

where m is an intrinsic apparent magnitude predicted from MERAXES and the noise denotes a random Gaussian flux uncertainty with a mean value of zero. We obtain a 1σ noise magnitude from the 5σ flux limits (Bouwens et al. 2015) listed in Table 1, using $\text{noise}_{1\sigma} = 10^{(-0.4\times m_{5\sigma})/5}$, where $m_{5\sigma}$ indicates the 5σ flux limit.

Secondly, we take into account the observational flux limits from Bouwens et al. (2015). Bouwens et al. (2015) observed LBGs in combined survey fields consisting of the Hubble XDF and CANDELS survey. We apply the colour selection criteria to model galaxies brighter than the 5σ flux limits and use the different flux limits for the individual survey fields listed in Table 1. In cases of non-detection in the drop-out band (i_{775} for $z\sim6$ and z_{850} for $z\sim7$), Bouwens et al. (2015) set the magnitudes to be equal to the 1σ flux limit to measure a colour. We use the same substitution to model galaxies that are detected redward of the Lyman break.

Fig. 1 shows the colours of model galaxies with the colour selection regions for LBGs at $z\sim6$, 7 and 8, respectively. We plot the predicted colours of galaxies at mean redshifts $z=5.9$, 6.8 and 7.9 (for LBGs at $z\sim6$, 7 and 8) estimated from observations (Bouwens et al. 2015). We find that the predicted colour distributions of model galaxies are broadened by photometric errors compared with the distribution of intrinsic colours. This is because the noise term in equation (15) causes colours to be scattered, especially for faint galaxies, which are mainly hosted by low mass haloes.

The redshift distribution can be written as

$$N(z) = n(z)\frac{\mathrm{d}^2V}{\mathrm{d}z\,\mathrm{d}\Omega}, \quad (16)$$

where n(z) is the comoving number density of galaxies and $\mathrm{d}^2V/\mathrm{d}z\,\mathrm{d}\Omega$ is the comoving volume per solid angle per redshift interval. Observationally, the redshift distribution is measured using an estimate of completeness. The completeness is defined as the ratio of selected galaxies to all galaxies at a given magnitude and redshift, and is estimated using the probability of recovering artificial LBGs (e.g. Yoshida et al. 2006; Bouwens et al. 2015). Similarly, we compute the redshift distribution using the ratio of model galaxies selected as LBGs to all model galaxies at a given magnitude bin in each snapshot. Fig. 2 shows the predicted redshift distribution of model LBGs corresponding to individual survey fields at $z\sim6$, 7 and 8, respectively. We find that the predicted redshift distribution using the ratio between model LBGs and all model galaxies shows better agreement with the measured distribution (Bouwens

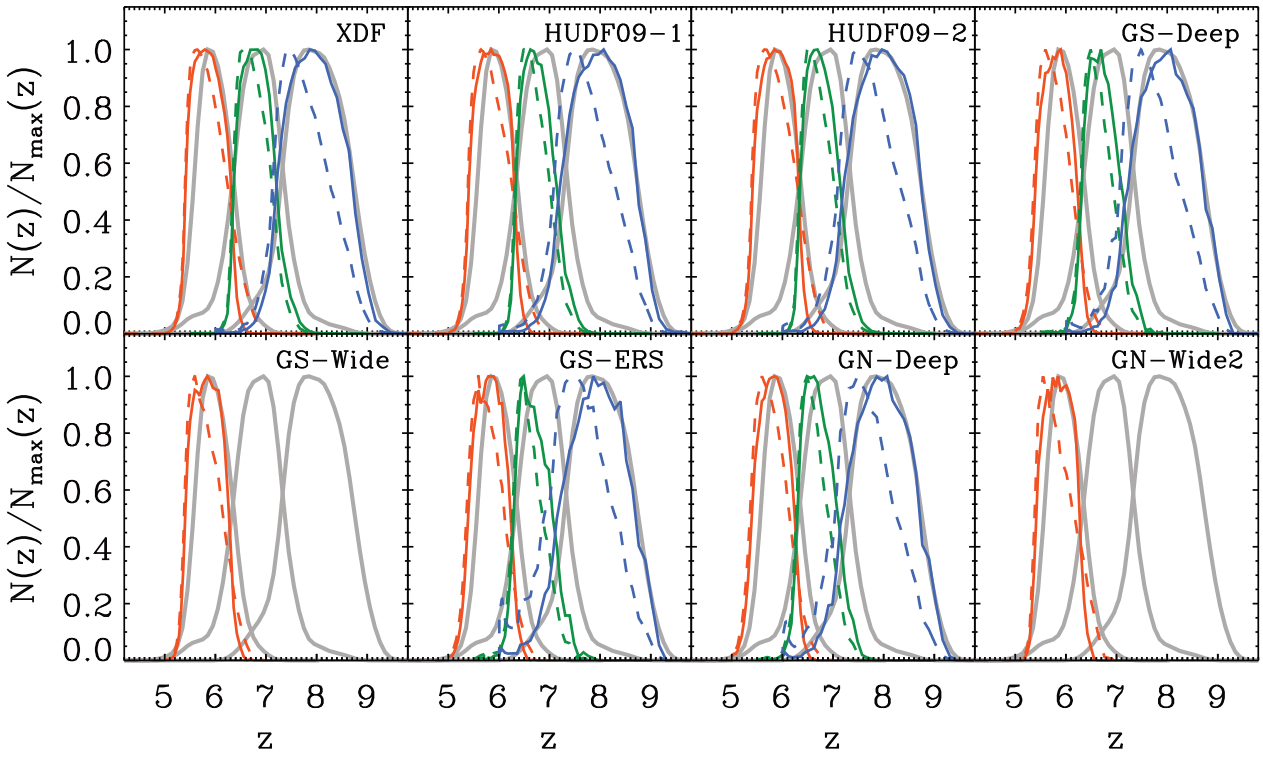

Figure 2. Redshift distributions of selected LBGs for different observational flux limits of the XDF and CANDELS survey fields, as labelled in each panel. The redshift distributions are normalized to have a maximum value of unity. Solid lines represent the redshift distribution using the ratio of model galaxies selected as LBGs to all model galaxies at a given magnitude bin in each snapshot and dashed lines represent the redshift distribution using equation (16). Grey lines represent the redshift distribution of observed LBGs estimated by Bouwens et al. (2015) using Monte Carlo simulations. Note that we do not plot the predicted distribution at $z \sim 7$ and 8, since LBGs are not identified in GS-Wide and GN-Wide2 fields at $z \sim 7.2$ (Barone-Nugent et al. 2014).

et al. 2015) than results based on equation (16). However, we note that the difference in clustering amplitude caused by the different determinations of $N(z)$ is between a few and 10 per cent. Relative to the observational uncertainties, this difference is not significant, but it may be important to include when modelling future observations.

Overall, the colour selection criteria of Bouwens et al. (2015) successfully exclude model galaxies outside the target redshift. We note that the colour selection criteria are designed to exclude low redshift interlopers as well as to select galaxies in the specific redshift range. Therefore, we demonstrate only that the selection criteria effectively select model galaxies in intended redshift ranges.

3.2 Luminosity function

Liu et al. (2016) showed that the MERAXES model successfully predicts the observed rest-frame UV luminosity function in the redshift range $z = 5$–10. To evaluate the influence of flux and redshift uncertainty on model LBG selection, we predict the rest-frame UV luminosity functions at $z \sim 6$, 7 and 8 using all selected model LBGs over the redshift distributions shown in Fig. 2. The observed luminosity function is measured using an effective volume taking into account the completeness (e.g. Yoshida et al. 2006; Bouwens et al. 2015). To mimic observations we define the effective volume in each UV magnitude bin as (Park et al. 2016)

$$V_{\rm eff} = \sum_{i}^{N} V_{\rm sim}\ p(m, z_i), \tag{17}$$

where N is the snapshot number at which model LBGs are selected, $V_{\rm sim}$ is the simulation volume and $p(m, z_i)$ is a ratio of the number of selected LBGs to the total number of galaxies in a magnitude bin at snapshot z_i.

Figure 3 shows the predicted rest-frame UV luminosity function together with the observed luminosity function from Bouwens et al. (2015). For predicted luminosity functions at $z \sim 6$, 7 and 8, the number of selected snapshots are 37, 40 and 40, and their redshift spans are $5.00 \leq z \leq 6.91$, $5.50 \leq z \leq 8.28$ and $6.00 \leq z \leq 9.61$, respectively. We find that the predicted luminosity functions using LBGs selected over the full photometric redshift distributions are in good agreement with the predicted luminosity functions using all galaxies at the target redshifts. This implies that the predicted luminosity function at each target redshift is representative of the predicted luminosity function over the redshift distribution (see also Park et al. 2016). Note that we use snapshots at $z = 5.9$, 6.8 and 7.9 as target redshifts for LBGs at $z \sim 6$ and $z \sim 7$, which are the estimated mean redshifts for LBG samples at $z \sim 6$, 7 and 8 from Bouwens et al. (2015). Overall, the predicted luminosity functions at $z \sim 6$ and 7 are consistent with the measured luminosity functions from Bouwens et al. (2015).

3.3 Modelling the ACF

The ACF provides the two-dimensional correlation of galaxies, which are projected along the line of sight. Limber's equation (Limber 1954) describes an integral relation between the real-space correlation function and the ACF. If we assume that the mean number

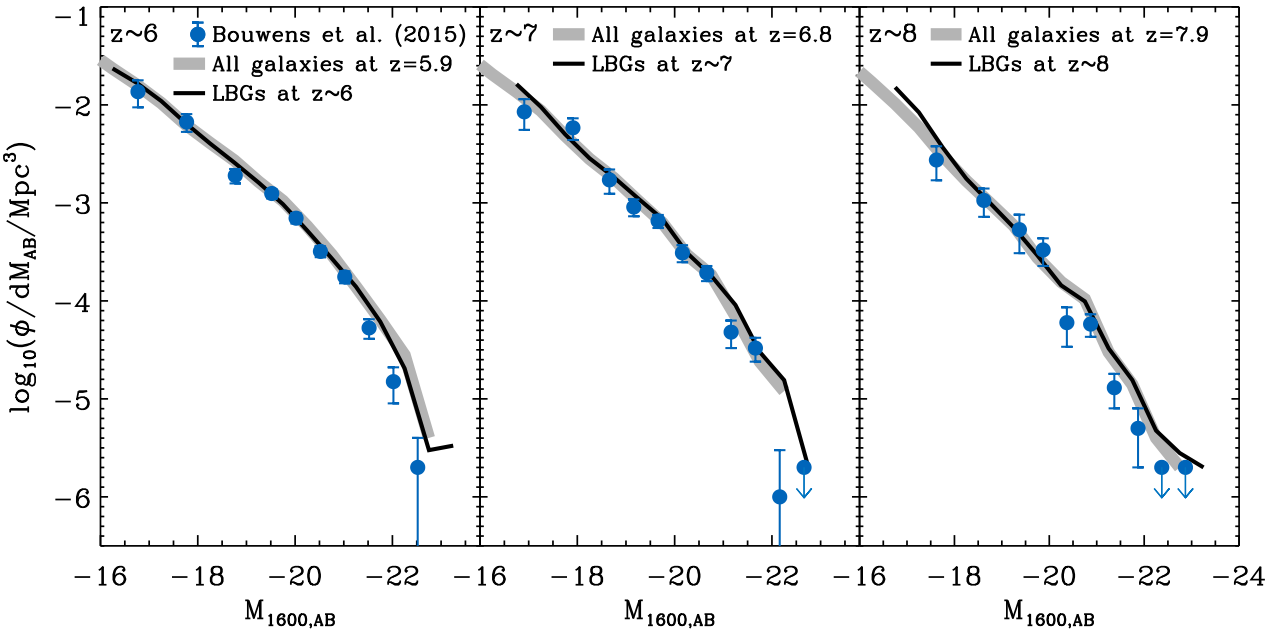

Figure 3. The predicted rest-frame UV luminosity functions with the observed luminosity functions from Bouwens et al. (2015). Thick grey lines represent the predicted luminosity function using all galaxies in the snapshot at $z = 5.9$, 6.8 and 7.9, respectively, which are mean redshifts for LBGs in $z \sim 6$, $z \sim 7$ and $z \sim 8$ samples from Bouwens et al. (2015). Solid lines represent the predicted luminosity functions of selected LBGs in the redshift ranges $5.0 \lesssim z \lesssim 7.5$, $5.5 \lesssim z \lesssim 8.5$ and $6.0 \lesssim z \lesssim 9.5$, respectively.

density does not rapidly vary with redshift and that the small angle approximation, this equation is a good approximation of the ACF. To compute the angular correlation function using snapshots generated at discrete redshifts, we use Limber's equation (Limber 1954),

$$w(\theta) = \frac{2\int_0^\infty [N(z)]^2 / R_{\rm H}(z) \left(\int_0^{2r} {\rm d}u\ \xi(r_{12}, z)\right) {\rm d}z}{\left[\int_0^\infty N(z){\rm d}z\right]^2}, \quad (18)$$

where $N(z)$ is the redshift distribution of galaxies, $R_{\rm H}(z)$ is the Hubble radius and $\xi(r_{12}, z)$ is the two-point correlation function. Using the small angle approximation, we denote $r_{12} = \sqrt{u^2 + r^2\theta^2}$, where $u = r_1 - r_2$ and $r = (r_1 + r_2)/2$ for comoving distances r_1 and r_2 to a pair of galaxies.

Before computing Limber's equation, we first compute the two-point correlation function in each snapshot involved in the redshift distribution. The two-point correlation function is calculated using the excess probability of finding a pair of galaxies at separation r to $r + \delta r$ compared to a random distribution,

$$1 + \xi(r) = \frac{DD}{\bar{n}^2} \frac{1}{V{\rm d}V}, \quad (19)$$

where DD is the number of galaxy pairs, $\bar{n}$ is the mean galaxy number density, V is the simulation volume and ${\rm d}V$ is the differential volume between r and $r + \delta r$. When integrating the two-point correlation function in equation (18) at scales beyond which the model cannot predict due to the finite simulation volume, we use a scaled dark matter two-point correlation function. On these scales the two-point correlation function is $\xi(r, z) = b(z)^2\, \xi_{\rm DM}(r, z)$, where b is the linear galaxy bias and $\xi_{\rm DM}(r, z)$ is calculated by the linear initial dark matter power spectrum. To find the linear galaxy bias we use a mean bias over the range $5\,{\rm Mpc} \leq r \leq 10\,{\rm Mpc}$ (see e.g. Orsi et al. 2008). In this study, we use the predicted two-point correlation function from simulations on scales up to 10 comoving Mpc, corresponding to ~230 arcsec at $z \sim 7$. This length scale is sufficiently large to compare the model predictions with current observations.

4 COMPARISON WITH OBSERVATIONS

In this section we show the predicted angular clustering and galaxy bias, and compare the resulting clustering properties with observations.

4.1 ACF

Barone-Nugent et al. (2014) measured the ACF of LBGs at $z \sim 4$–7.2 using the samples of Bouwens et al. (2015). They measured ACFs in eight and six individual survey fields for LBGs at $z \sim 6$ and $z \sim 7.2$, respectively. For the ACFs at $z \sim 7.2$, Barone-Nugent et al. combined the $z \sim 7$ and $z \sim 8$ samples together to measure the ACFs. The redshift of 7.2 is the mean redshift estimated from the redshift distribution. In this paper, we do not reproduce the ACFs at $z \sim 7.2$, using combined samples at $z \sim 7$ and 8. Instead, we predict the ACF at $z \sim 7$ and 8 separately and investigate the clustering properties at each redshift. Barone-Nugent et al. (2014) also calculated a combined ACF from the ACFs measured using the individual survey fields (see Barone-Nugent et al. 2014 for more details). We do not reproduce the combined ACF in this study, but we show the combined measurement as a reference.

Fig. 4 shows the predicted ACF of LBGs at $z \sim 6$ selected from the model together with the measured ACFs and the combined ACF from observations. We show the ACFs on a logarithmic scale to accurately depict the clustering on large scales. The predicted ACFs for each individual field are consistent to within $\sim 2.5\sigma$ of the observed ACFs except GS-Deep field on large scale. On large scales it appears as if the model may overpredict the amplitude by up to an order of magnitude. However, we find that there are noticeable

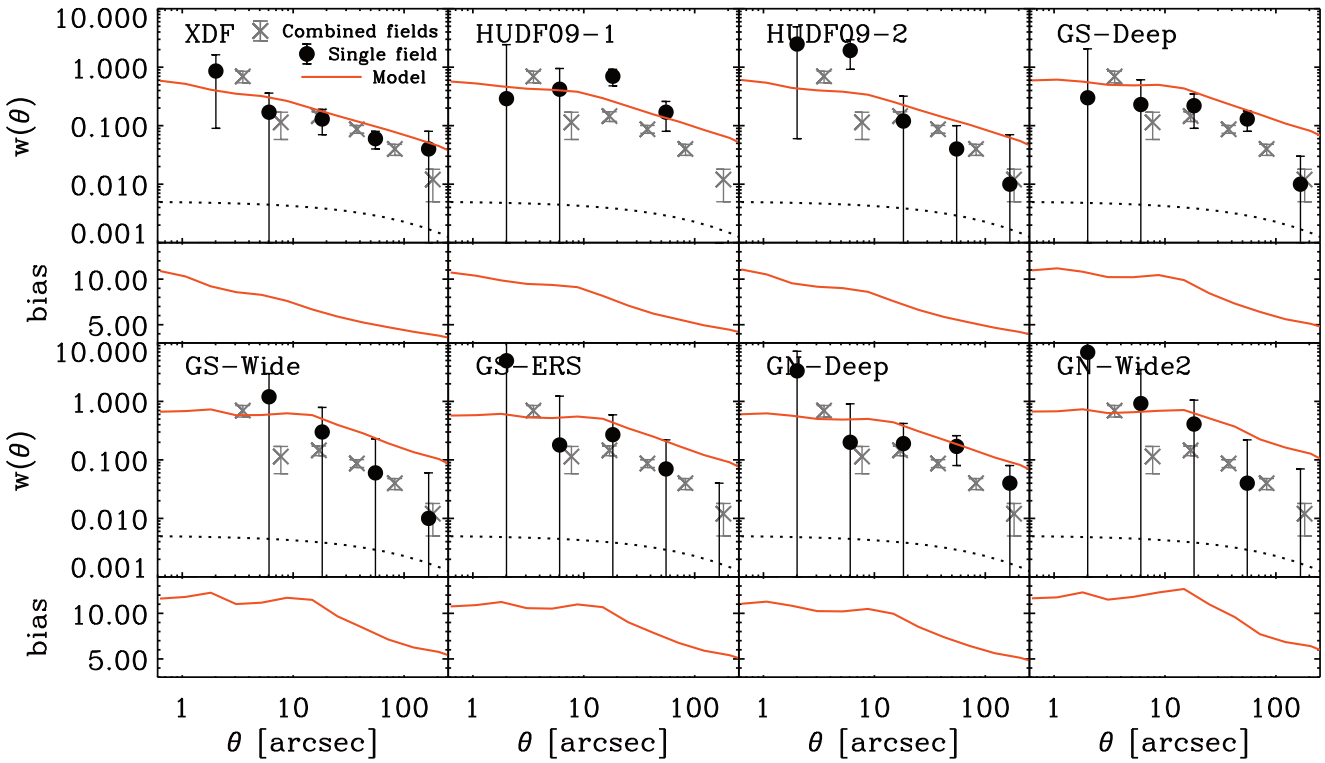

Figure 4. The predicted angular clustering of LBGs at $z \sim 6$ (solid line). The dotted lines represent the predicted angular clustering of dark matter computed using the linear dark matter power spectrum. The name of the field is labelled on each panel. Filled circles with error bars show the observed ACF measured from the individual field. The crosses with error bars show the observed combined ACF and are reproduced in each panel for reference. All errors are 1σ and estimated using bootstrap resampling (Ling, Barrow & Frenk 1986). Note that the data points of the combined ACF and GS-Wide field at the smallest angular separation bins ($\theta \lesssim 3$ arcsec and $\theta \lesssim 5$ arcsec, respectively) are not seen in this plot because the clustering amplitude is negative. The data point of HUDF09-1 at the largest angular separation bin ($\theta \gtrsim 80$ arcsec) is not seen for the same reason. We interpret these negative clustering amplitudes as being caused by insufficient numbers of galaxy pairs at small angular separations, and note that the ACFs binned on a linear scale do not show negative clustering at small scale (see Barone-Nugent et al. 2014). In each case the bottom sub-panels show the predicted galaxy bias, defined as $b^2(\theta) = w_{\rm LBGs}(\theta)/w_{\rm DM}(\theta)$.

differences between the measured ACFs from each survey field. These field-to-field variations can be explained by cosmic variance. In addition, the observed sample variance in each field contributes to the field-to-field variation, because there are order-of-magnitude differences in the number of LBGs used in the ACF analysis (as listed in Table 1). Because of this field-to-field variation, the measured ACFs do not clearly show the expected dependence on luminosity (as we will discuss in the next section). Fig. 1 of Barone-Nugent et al. (2014) also shows that the combined ACF from all fields shows a large variation in the amplitude on large scales. The presence of this variation reinforces our interpretation of it being driven by sample variance. On the other hand, the predicted ACFs do not show significant field-to-field variation. Overall, when considering the estimated uncertainties from observational measurements, the predicted ACFs for each individual field are in good agreement with the measured ACFs from observations. We note that the predicted ACF in GS-Deep field is also consistent with the observation to within 2σ when comparing the ACF with the measured power-law ACF, as discussed in the next section.

Fig. 5 shows the predicted angular clustering of LBGs at $z \sim 7$ and 8 selected from the model together with the measured ACFs at $z \sim 7.2$ and the combined ACF from observations. For this case, we plot the ACF on a linear scale because of the sparse number of observed pairs at small angular separations. Predicted ACFs for each individual field are again in good agreement with the measured ACFs from observations.

4.2 Dependence of clustering on luminosity

In analytic and numerical models for dark matter haloes, more massive haloes cluster more strongly than less massive haloes (e.g. Mo & White 1996). Over the past decade observational results show that the clustering strength of galaxies depends on luminosity both at high redshifts (Ouchi et al. 2004, 2005; Cooray & Ouchi 2006; Kashikawa et al. 2006; Lee et al. 2006; Hildebrandt et al. 2009) and in the local Universe (e.g. Norberg et al. 2001; Zehavi et al. 2002). This clustering dependence on luminosity supports the expectation that more massive haloes host brighter galaxies (e.g. Giavalisco & Dickinson 2001). Recent studies (Barone-Nugent et al. 2014; Harikane et al. 2016) also showed this clustering dependence on luminosity for LBGs up to $z \sim 7$.

The ACFs for each individual survey field allow us to investigate dependence of clustering on luminosity, since the flux limits of these fields listed in Table 1 are all different. The predicted clustering amplitude for each individual field at $z \sim 6$ (Fig. 4) increases from the deepest field (XDF) to the shallowest field (GN-Wide2). We further find that the model predictions are comparable with these measurements. The predicted clustering amplitudes at $z \sim 7$ and 8 (Fig. 5) also show a similar trend.

On large scales, the ACFs can be approximated by a power law using the angular correlation amplitude, A_w, and the correlation slope, β, i.e. $w(\theta) = A_w\,\theta^{-\beta}$. To more clearly show the dependence of clustering on luminosity, we calculate the best-fitting angular correlation amplitude, A_w, using the least square method for each

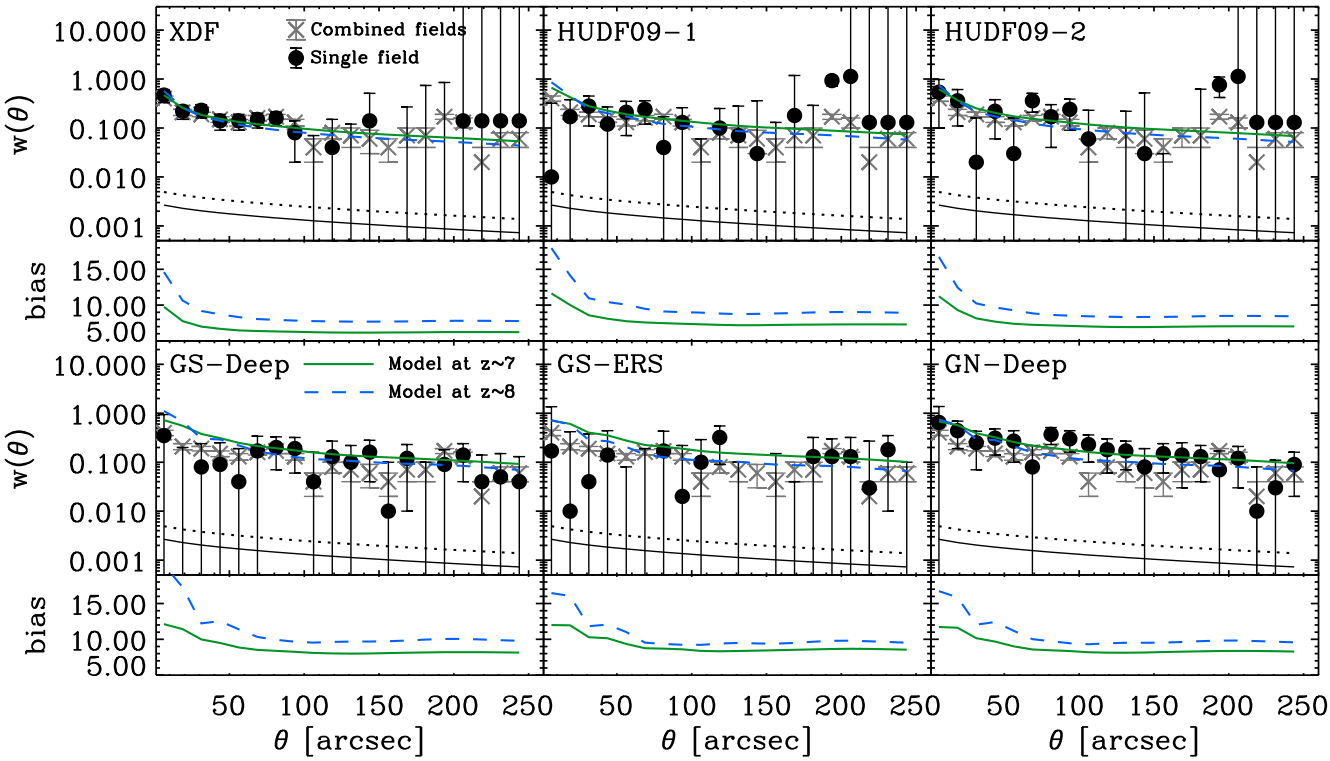

Figure 5. The same as Fig. 4, but at $z \sim 7$ and 8. Dotted and dashed–dot lines represent the predicted ACF of dark matter at $z \sim 7$ and 8, respectively. Note that we plot the ACF on a linear scale because of the sparse number of observed pairs at small angular separations.

survey field and compare this predicted value with observational measurements (Barone-Nugent et al. 2014; Harikane et al. 2016). Note that Barone-Nugent et al. (2014) and Harikane et al. (2016) fixed $\beta = 0.6$ and 0.8, respectively, to correct the clustering amplitude measured from their sample (e.g. Roche & Eales 1999; Lee et al. 2006). We will present the best-fitting value of A_w predicted from the model with fixed β ($\beta = 0.6$ and 0.8) for comparison, and analyse the values of the best-fitting value of A_w when allowing the value of β to vary. We also note that we find the best-fitting values by considering only angular separations larger than 10 arcsec We do not show uncertainties for these best-fitting values because statistical errors, e.g. bootstrap resampling, are too small to be presented.

The left-hand panels of Fig. 6 show the measured and predicted A_w values for each individual survey field at $z \sim 6$ (top), and $z \sim 7$ and 8 (bottom) as a function of the magnitude limit of Y_{105} and J_{125} band, respectively, corresponding to the rest-frame UV magnitude. In addition to computing clustering in individual field galaxy samples, Barone-Nugent et al. (2014) also measured ACFs of bright and faint subsamples split using a UV luminosity of $M_{\rm AB(1600)} = -19.4$, which is the median magnitude for the overall sample. The corresponding values for A_w are shown in the middle-left and -right panels of Fig. 6. We split model LBGs into bright and faint subsets using the same magnitude cut and show the corresponding model clustering predictions for bright and faint subsets. Note that since the median UV luminosity is for all samples, A_w values for bright LBGs are not measured in GS-ERS, GN-Deep and GN-Wide2 fields from observations, and also are not predicted in GS-Wide, GS-ERS, GN-Deep and GN-Wide2 fields from the model. Fig. 6 shows that whilst there is significant field-to-field variation due to cosmic variance (Trenti & Stiavelli 2008), the data shows the trend that brighter LBGs are more highly clustered, based both on the dependence with limiting magnitude, and comparison between bright and faint samples.

Before comparing model predictions with observations, we first consider the model predictions in isolation. At $z \sim 6$ (top-left-hand panel in Fig. 6) the predicted angular correlation amplitude increases from the deepest survey field (XDF) towards the shallowest survey field (GN-Wide2). In addition, the comparison between the bright and faint subsets (middle-left and -right in top panels) show a similar trend. The predicted amplitudes for bright subsets are higher than those predicted for faint subsets. At $z \sim 7$ and 8 (bottom panels) the model predictions show similar trends to the predictions at $z \sim 6$. We note that, in wide and shallow fields (GS and GN fields), the correlation amplitudes of bright LBGs are almost identical to those of all LBGs. This is because the wide and shallow fields contain mainly bright LBGs, so that the two samples are very similar.

When comparing the model predictions with observations, we find that the trends of clustering with luminosity are similar between models and observation. Thus, clustering measurements imply that brighter LBGs reside in more massive haloes at $z > 6$, as expected from simulations (Liu et al. 2016). We note that the model predictions for the HUDF deviates from the measured values, although it is still consistent within 3σ. This is likely due to large sample variance because the HUDF09-1 and HUDF09-2 have only a few tens of LBGs. Overall, we find that the trends are much clearer than in the observed samples because our single simulation realization does not include cosmic variance.

We can also study the dependence of clustering on luminosity using the measured ACFs by Harikane et al. (2016). They split the observed sample at $z \sim 6$ into two subsamples with thresholds of $M_{\rm AB(1600)} < -20.0$ and $M_{\rm AB(1600)} < -19.1$, respectively. They also

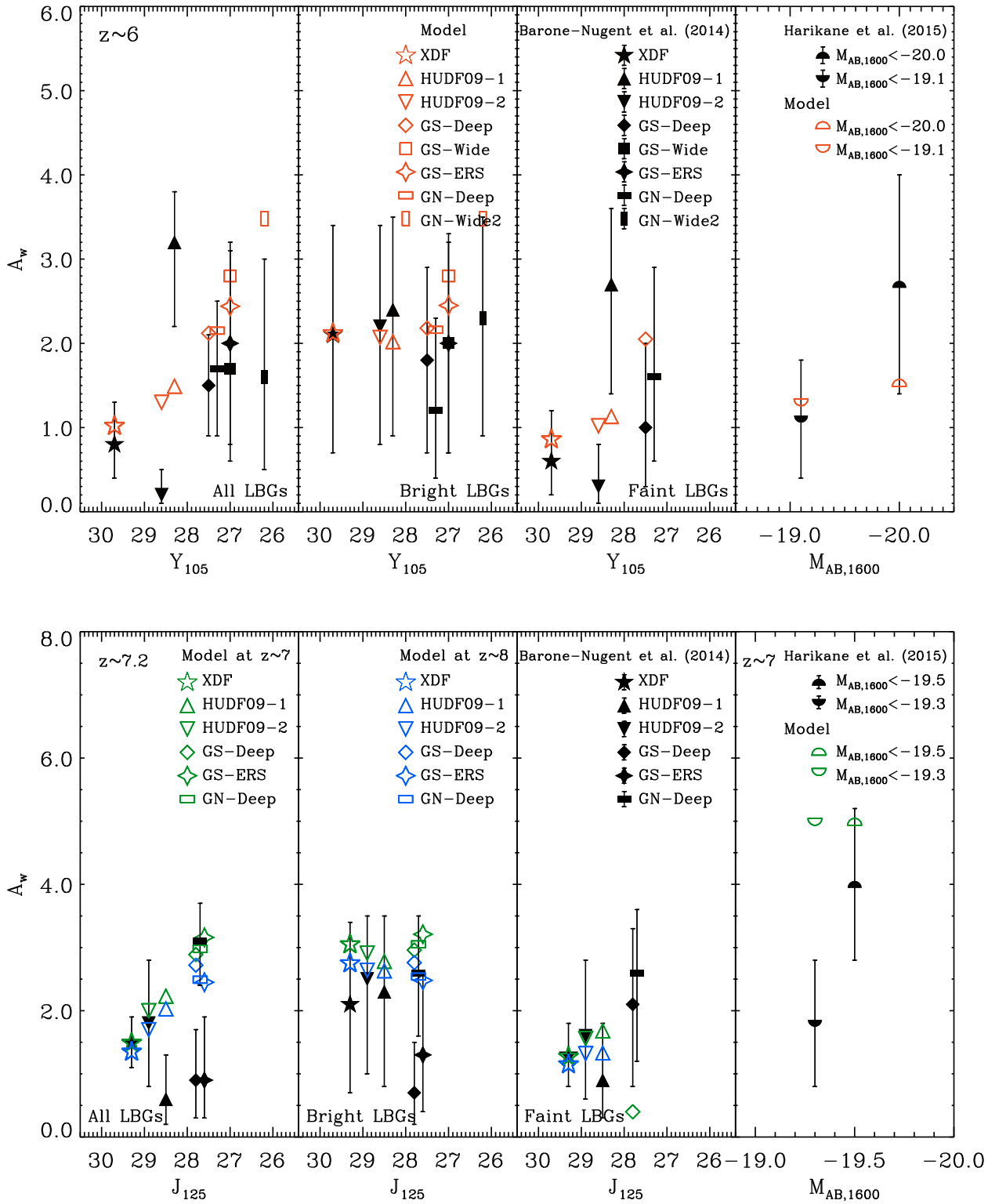

Figure 6. Top: comparison of the angular correlation amplitude, A_w, with observations at $z \sim 6$. The filled symbols with error bars represent the measured A_w by Barone-Nugent et al. (2014) and the positions of x-axis indicate the detection limit for Y_{105} band magnitude corresponding to $M_{\rm AB,1600}$. The empty symbols represent the model predictions. Left-hand, middle-left and middle-right panels show the galaxy bias of all, bright and faint LBGs, respectively. The right-hand panel shows the A_w measured by Harikane et al. (2016) and the model prediction. Bottom: the same as top panels, but we compare the predicted A_w values at $z \sim 7$ and 8 with the measured A_w at $z \sim 7.2$ by Barone-Nugent et al. (2014). Note that for $z \sim 8$ LBGs H_{160} band corresponds to the rest-frame UV magnitude, but we plot the predicted A_w at $z \sim 8$ as a function of J_{125} band for simplicity, since the detection limit of H_{160} band decreases the same rate as the J_{125} band (see Table 1).

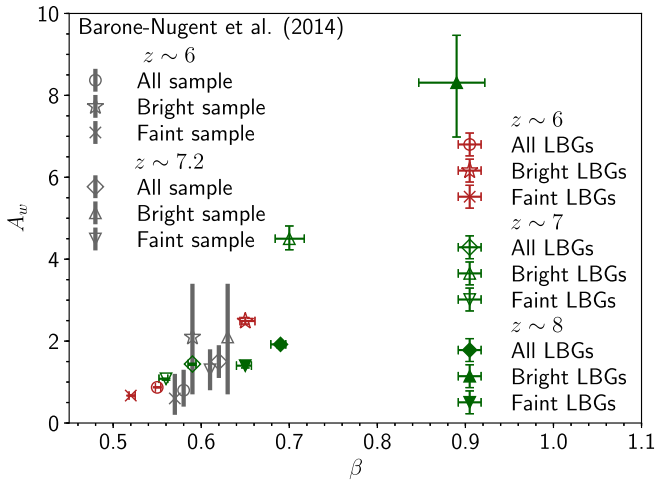

Figure 7. Predicted best-fitting values of A_w and β at $z \sim 6$–8 for XDF field. Different symbols correspond to all, bright and faint subset of model LBGs at different redshifts as indicated by the legend. Estimated A_w values from Barone-Nugent et al. (2014) are shown for comparison. We note that the observational A_w values are estimated with an assumption of $\beta = 0.6$. These values are slightly shifted on β for clarity. All errors are 1σ and estimated using bootstrap resampling.

split the sample at $z \sim 7$ into two subsamples with thresholds of $M_{\rm AB(1600)} < -19.5$ and $M_{\rm AB(1600)} < -19.3$, respectively.

We find that the predicted best-fitting A_w values using the same rest-frame UV magnitude thresholds show a dependence of increased clustering with luminosity (right-hand panels of Fig. 6), supporting conclusions based on HUDF and COSMOS fields. We note that the predicted values of A_w at $z \sim 7$ are identical. This is caused by the fixed value of β as discussed below. We checked that when we allow the value of β to vary, the predicted values of A_w show the same trend with the observation and are consistent within 2σ errors.

Fig. 7 shows the predicted best-fitting values of A_w and β for model LBGs at $z \sim 6$–8 when allowing the value of β to vary. As shown in Fig. 6, the best-fitting values of A_w increase with luminosity. We also find that the best-fitting values of β increase with luminosity. This is consistent with observational measurements at $z \sim 4$–5 from Kashikawa et al. (2006) and a semi-analytical model prediction at $z \sim 4$ from Park et al. (2016). This trend shows that A_w and β increase with increasing redshift. When comparing predicted values with estimated A_w at $z \sim 6$ and $z \sim 7.2$ from Barone-Nugent et al. (2014), the predicted best-fitting values are still consistent with observations within 1σ errors. The predicted values of bright LBGs at $z \sim 7$–8 are consistent with observations within $\sim 2.5\sigma$ errors. We note that Fig. 6 shows A_w at $z \sim 8$ to be higher than that at $z \sim 7$, which is in contrast to this result. We interpret that this results from the fixed β. We checked that the best-fitting A_w is underestimated when the best-fitting β is larger than a fixed β and vice versa. Although this difference arising from the fixed β is not significant when taking into account the current observational uncertainty, this factor should be considered in future surveys.

Overall, the model prediction of the correlation amplitude reproduces the measured dependence of clustering on luminosity from observations within 2σ errors.

4.3 Galaxy bias and halo mass

We show the predicted galaxy bias for each survey field at $z \sim 6$–8 (bottom sub-panels in Fig. 4 and Fig. 5), defined as the ratio of angular correlation function of model LBGs to the angular correlation function of dark matter, $b^2(\theta) = w(\theta)/w_{\rm DM}(\theta)$. We computed the ACF of dark matter using the initial linear dark matter power spectrum linearly extrapolated using the growth factor. We compute the ACF of dark matter using the same redshift distribution as the galaxies (Fig. 4 and Fig. 5) for each survey field.

In general, biases for deep fields (XDF and HUDFs) are lower than biases for shallow fields, i.e. brighter galaxies show higher bias. This result follows directly from and reinforces the clustering dependence on luminosity discussed in Section 4.2. Scale dependent bias is a general prediction of the hierarchical structure formation framework (e.g. Colín et al. 1999). We find that model biases increase with decreasing angular separation, but are almost constant on large angular separations ($\theta \gtrsim 70$ arcsec). In addition, this small scale increase is more significant for shallow fields consisting only of bright galaxies. Although it is not possible to compare this result directly with observations at $z \sim 6$, for observed LBGs at lower redshifts ($z \sim 4$ and $z \sim 5$) scale dependent bias has been reported (Hamana et al. 2004; Kashikawa et al. 2006; Lee et al. 2006).

To compare model predictions with observations, we therefore use large scale bias. Barone-Nugent et al. (2014) estimated the galaxy bias using the ratio of the galaxy variance at $8\,h^{-1}$ Mpc, $\sigma_{8,\rm g}$, to the linear matter fluctuation at $8\,h^{-1}$ Mpc, $\sigma_8(z)$ (e.g. Lee et al. 2006). The galaxy variance, $\sigma_{8,\rm g}$, is calculated from the ACF using (Peebles 1980)

$$\sigma^2_{8,\rm g} = \frac{72(r_0/8\ h^{-1}\ {\rm Mpc})^\gamma}{(3-\gamma)(4-\gamma)(6-\gamma)\ 2^\gamma}, \quad (20)$$

where r_0 is the correlation length and $\gamma = \beta + 1$ are parameters to approximate the real-space correlation function using a power law, $\xi(r) = (r/r_0)^\gamma$.

In the model, we calculate the best-fitting parameters of r_0 and γ from the predicted two-point correlation function in the range $1\,{\rm Mpc} < r < 10\,{\rm Mpc}$ for each snapshot and determine a weighted average value of r_0 and γ using the redshift distribution. We note that we do not fix γ as was done with the observations. We checked that any resulting biases from fixed γ are not significant. Fig. 8 shows the predicted large scale bias at $z \sim 6$ (left three panels in top row), $z \sim 7$ and $z \sim 8$ (left three panels in bottom row), respectively. In agreement with the angular correlation amplitude (Fig. 6), the predicted bias increases from the deepest field to the shallowest field. Bright subsamples for each survey field also show a higher bias than faint subsamples. When comparing the model prediction with observations, all predicted biases are consistent with observations (except HUDF fields). We interpret this to be because of large sample variance in these fields as discussed in Section 4.2.

Waters et al. (2016) studied galaxy clustering at $z = 8$, 9 and 10 using their hydrodynamic simulation. They predicted the linear galaxy bias of 13.4 ± 1.8 at $z = 8$ from the real-pace correlation function. This value is higher than the prediction from our model of $b = 9.67$ for bright LBGs ($M_{\rm AB(1600)} < -19.4$) in the XDF field. However, considering their assumed detection limit ($M_{\rm AB(1600)} \lesssim -20.5$), our result is consistent with their prediction.

Harikane et al. (2016) estimated the effective galaxy bias from their HOD modelling,

$$b_{\rm g}^{\rm eff} = \frac{1}{n_{\rm g}} \int {\rm d}M_{\rm h}\, \frac{{\rm d}n}{{\rm d}M_{\rm h}}(M_{\rm h}, z)\ N(M_{\rm h})\ b_{\rm h}(M_{\rm h}, z), \quad (21)$$

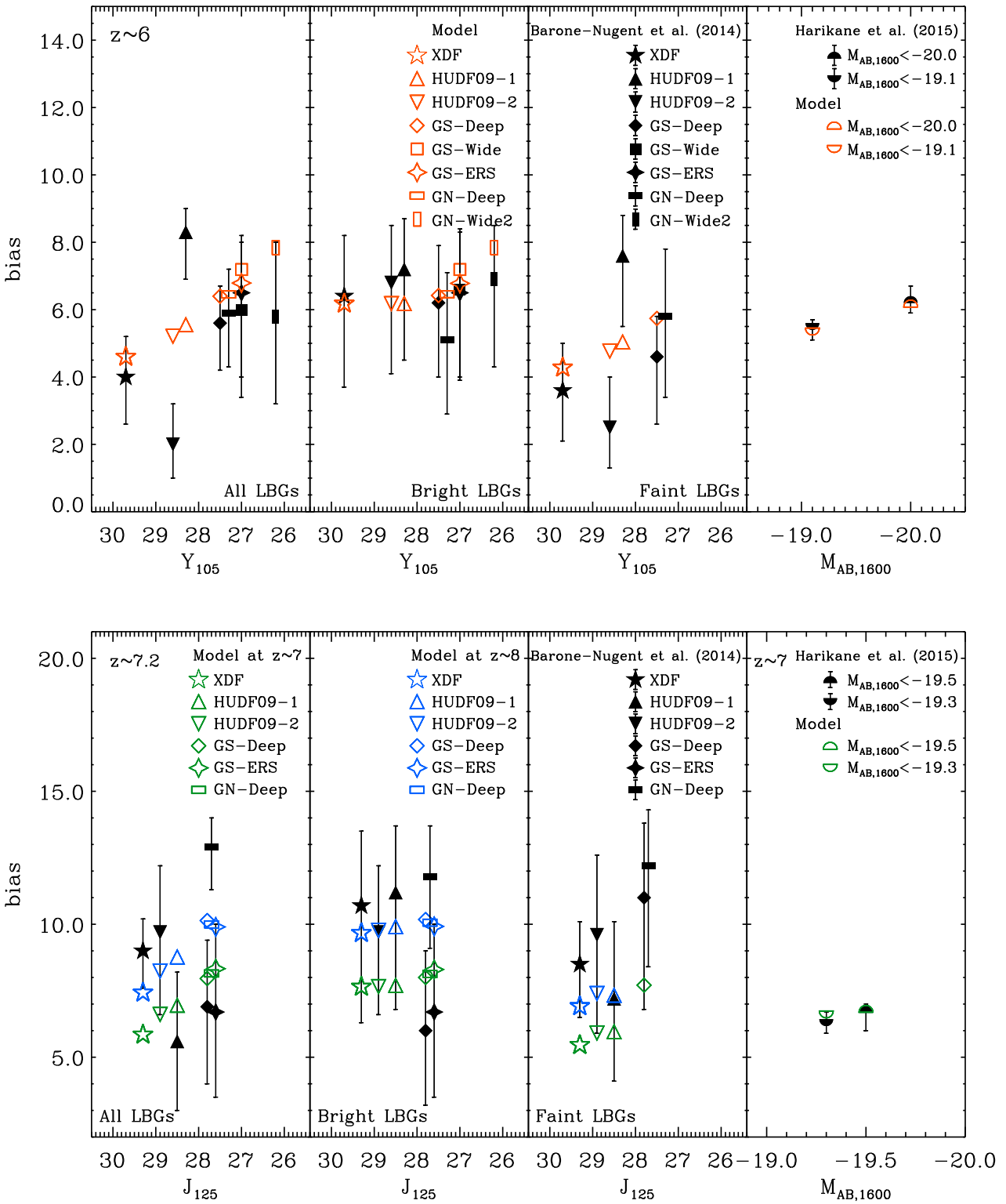

Figure 8. Top: comparison of the predicted galaxy bias with observations at $z \sim 6$. The galaxy bias is defined as $\sigma_{8,\mathrm{g}}/\sigma_8$, where $\sigma_{8,\mathrm{g}}$ is the galaxy variance at $8h^{-1}$ Mpc. The filled symbols with error bars represent the measured bias by Barone-Nugent et al. (2014) and the positions on the x-axis indicate the detection limit for Y_{105} band magnitude corresponding to $M_{\mathrm{AB},1600}$. The empty symbols represent the model predictions. Left-hand, middle-left and middle-right panels show the galaxy bias of all, bright and faint LBGs, respectively. The right-hand panel shows the galaxy bias measured by Harikane et al. (2016) and the model prediction. Bottom: the same as the top panels, but we compare the predicted biases at $z \sim 7$ and 8 with the estimated bias at $z \sim 7.2$ from Barone-Nugent et al. (2014). Note that for $z \sim 8$ LBGs the H_{160} band corresponds to the rest-frame UV magnitude, but we plot the predicted bias at $z \sim 8$ as a function of J_{125} band for simplicity since the detection limit of H_{160} band decreases at the same rate as J_{125} band (see Table 1).

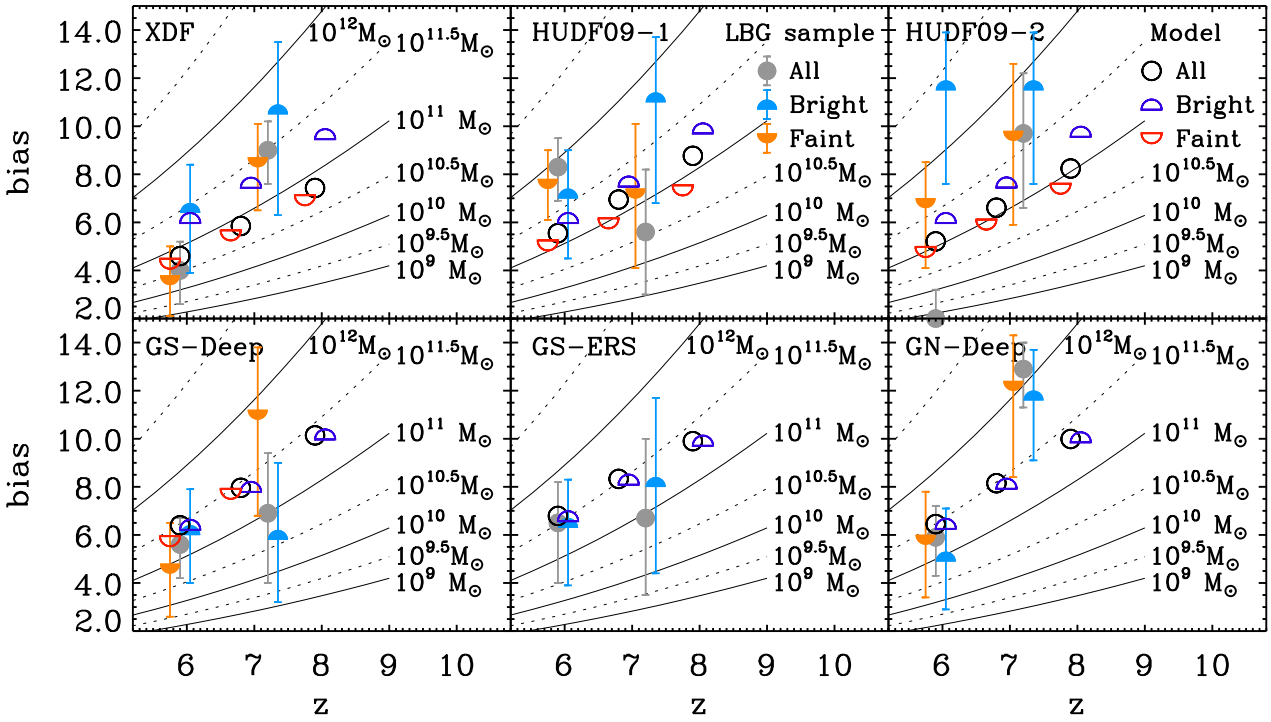

Figure 9. Galaxy bias as a function of redshift for each survey field. The galaxy bias is defined as $\sigma_{8,\,g}/\sigma_8$. Filled circles, upper half circles and lower half circles with error bars are measured bias of all, bright and faint sample, respectively. Empty circles, upper half circles and lower half circles are the predicted bias for all, bright and faint subset from the model. The solid and dotted lines represent the average bias for haloes with $M \geq M_{\rm halo}$ from Tinker et al. (2010). Note that the model does not predict the bias of the faint subset for the GS-ERS and GN-Deep fields due to an insufficient number of galaxies. Similarly, we do not show the measured bias of the faint subsample from the GS-ERS field for the same reason.

where $n_{\rm g}$ is the mean galaxy number density over the redshift distribution, $N(M_{\rm h})$ is the mean number of galaxies in a dark matter halo of mass $M_{\rm h}$, $\frac{{\rm d}n}{{\rm d}M_{\rm h}}(M_{\rm h}, z)$ is the halo mass function, and $b_{\rm h}(M_{\rm h}, z)$ is the halo bias. To compare the model prediction with observations we assume the halo bias of Tinker et al. (2010) following the assumption of Harikane et al. (2016), but do not use HOD modelling since the model directly provides the number of galaxies as a function of halo mass.

To investigate dependence with luminosity for this sample we split model LBGs into two subsets using the rest-frame UV magnitude thresholds corresponding to those in observations from Harikane et al. (2016); $M_{\rm AB(1600)} < -20.0$ and $M_{\rm AB(1600)} < -19.1$ for LBGs at $z \sim 6$, and $M_{\rm AB(1600)} < -19.5$ and $M_{\rm AB(1600)} < -19.3$ for LBGs at $z \sim 7$, respectively. As shown in the case of the large scale bias, the predicted effective bias increases with luminosity, and the resulting biases are in good agreement with the estimated bias from observations at both $z \sim 6$ and $z \sim 7$ (right-hand panels in Fig. 8).

Fig. 9 shows the evolution of observed galaxy bias as a function of redshift for each survey field. The biases increase with increasing redshift at fixed mass. As already noted, the model prediction also shows that brighter LBGs reside in more massive dark matter haloes. The halo mass corresponds to the mass range $10^{10.5}{\rm M}_\odot \lesssim M_{\rm halo} \lesssim 10^{11.5}{\rm M}_\odot$, which is consistent with the predicted halo mass for LBGs (see Liu et al. 2016). We note that this mass range is slightly higher than the estimated halo mass from Barone-Nugent et al. (2014), since we use a different bias calculation. The predicted biases are consistent with observations within approximately 2σ errors.

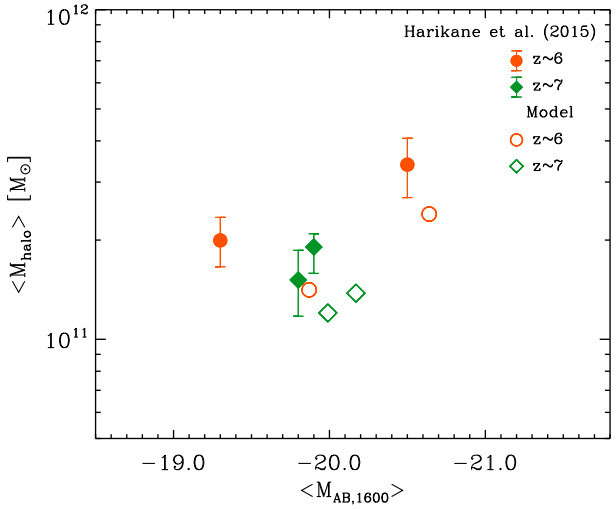

Figure 10. Mean dark matter halo mass as a function of mean UV luminosity. Filled circles and diamonds with error bars are estimated halo masses using HOD modelling from samples of Harikane et al. (2016) at $z \sim 6$ and $z \sim 7$, respectively. Empty circles and diamonds are model predictions at $z \sim 6$ and $z \sim 7$, respectively.

Fig. 10 compares the predicted mean dark matter halo mass as a function of mean UV luminosity with the estimated halo mass from Harikane et al. (2016). We use the same magnitude threshold for direct comparison with the observations for model LBGs (see Section 4.2). This analysis shows that brighter LBGs reside in more

massive dark matter haloes. The model predicts slightly brighter UV luminosity and less massive halo masses than the observations. However, the difference is not significant. Both observations and the model show that the host dark matter halo mass slightly decreases with increasing redshift.

While Harikane et al. (2016) found no significant evolution of the dark matter halo mass hosting LBGs with redshift, Barone-Nugent et al. (2014) reported the estimated halo mass using the combined sample from all survey fields (see Fig. 2 in Barone-Nugent et al. 2014). They show that the dark matter halo mass hosting LBGs may be increasing from $z \sim 6$ to $z \sim 7$ ($\lesssim 2\sigma$ confidence). On the other hand, the estimated dark matter halo masses from individual survey fields do not show this trend. Harikane et al. (2016) explained that this difference may be due to either the different sample selections or to large sample variance. Another possible difference is the method used to estimate the dark matter halo mass. Barone-Nugent et al. (2014) measured the large scale bias using a variance at $8\,h^{-1}$ Mpc, and estimated the mass by calculating a mass function weighted bias. Harikane et al. (2016) estimated a mean halo mass using HOD modelling. Since the mean halo mass is affected by contribution from a range of haloes including central and satellite galaxies, the mean halo mass may be different from the halo mass estimated from large scale bias. We checked the estimated halo mass from the large scale bias of the Harikane et al. (2016) sample, and found that this shows a similar trend to that of Barone-Nugent et al. (2014), although the evolution is weaker. Confirming the evolution of dark matter halo mass hosting LBGs of fixed luminosity between $z \sim 6$ and 7 will require larger surveys of LBGs.

Overall, the model prediction of galaxy bias is consistent with estimates of the observed bias and we find that dark matter halo masses at fixed luminosity do not show significant evolution with redshift.

5 CONCLUSION

We have investigated the clustering properties of LBGs at $z \sim 6$–8 using the semi-analytical model MERAXES developed as part of the DRAGONS project.

This is the first study to compare model predictions using a semi-analytical model with the new clustering measurements of LBGs up to $z \sim 8$. We predict the ACF of LBGs, and compare these model predictions with clustering measurements from survey fields consisting of the Hubble XDF, the HUDF and CANDELS (Barone-Nugent et al. 2014), and from the HUDF and CANDELS (Harikane et al. 2016).

We find that the predicted ACFs at $z \sim 6$–8 are in good agreement with ACFs measured from observations. The model predictions show a dependence of clustering amplitude on luminosity, with brighter LBGs being more strongly clustered than fainter LBGs. The predicted dependence on luminosity is consistent with observational results. This result implies that the trend of more massive haloes hosting brighter galaxies (e.g. Giavalisco & Dickinson 2001) holds during the epoch of reionisation. Consequently, the predicted galaxy bias increases with increasing luminosity as expected in a hierarchical galaxy formation model (e.g. Mo & White 1996). We also find that the predicted galaxy bias at fixed apparent magnitude increases with increasing redshift, as seen in observations.

We find that the model LBGs of magnitude $M_{\rm AB(1600)} < -19.4$ at $6 \lesssim z \lesssim 8$ reside in dark matter haloes of mean mass $\sim 10^{11.0} - 10^{11.5}\,{\rm M}_{\odot}$, and this dark matter halo mass does not evolve significantly during reionisation. The predicted dark matter halo mass is consistent both with the estimate of the dark matter halo mass by calculating a mass function weighted bias (Barone-Nugent et al. 2014), and with the estimate of the mean dark matter halo mass using HOD modelling (Harikane et al. 2016).

ACKNOWLEDGEMENTS

This research was supported by the Victorian Life Sciences Computation Initiative (VLSCI), grant ref. UOM0005, on its Peak Computing Facility hosted at the University of Melbourne, an initiative of the Victorian Government, Australia. Part of this work was performed on the gSTAR national facility at Swinburne University of Technology. gSTAR is funded by Swinburne and the Australian Governments Education Investment Fund. This research program is funded by the Australian Research Council through the ARC Laureate Fellowship FL110100072 awarded to JSBW. HSK is supported by a Discovery Early Career Researcher Awards (DE140100940) from the Australian Research Council. AM acknowledges support from the European Research Council (ERC) under the European Unions Horizon 2020 research and innovation program (grant agreement No 638809 AIDA).

REFERENCES

Barone-Nugent R. L. et al., 2014, ApJ, 793, 17
Baugh C. M., 2006, Rep. Prog. Phys., 69, 3101
Benson A. J., 2010, Phys. Rep., 495, 33
Bouwens R. J. et al., 2014, ApJ, 793, 115
Bouwens R. J. et al., 2015, ApJ, 803, 34
Bouwens R. J. et al., 2016, ApJ, 830, 67
Bradley L. D. et al., 2012, ApJ, 760, 108
Calzetti D., Armus L., Bohlin R. C., Kinney A. L., Koornneef J., Storchi-Bergmann T., 2000, ApJ, 533, 682
Cole S., 1991, ApJ, 367, 45
Colín P., Klypin A. A., Kravtsov A. V., Khokhlov A. M., 1999, ApJ, 523, 32
Cooray A., Ouchi M., 2006, MNRAS, 369, 1869
Cooray A., Sheth R., 2002, Phys. Rep., 372, 1
Croton D. J. et al., 2006, MNRAS, 365, 11
Dijkstra M., Haiman Z., Rees M. J., Weinberg D. H., 2004, ApJ, 601, 666
Duncan K. et al., 2014, MNRAS, 444, 2960
Fan X. et al., 2006, AJ, 132, 117
Giavalisco M., 2002, ARA&A, 40, 579
Giavalisco M., Dickinson M., 2001, ApJ, 550, 177
Guo Q. et al., 2011, MNRAS, 413, 101
Guo Q., White S., Angulo R. E., Henriques B., Lemson G., Boylan-Kolchin M., Thomas P., Short C., 2013, MNRAS, 428, 1351
Hamana T., Ouchi M., Shimasaku K., Kayo I., Suto Y., 2004, MNRAS, 347, 813
Harikane Y. et al., 2016, ApJ, 821, 123
Henriques B. M. B., White S. D. M., Thomas P. A., Angulo R., Guo Q., Lemson G., Springel V., Overzier R., 2015, MNRAS, 451, 2663
Hildebrandt H., Pielorz J., Erben T., van Waerbeke L., Simon P., Capak P., 2009, A&A, 498, 725
Jose C., Subramanian K., Srianand R., Samui S., 2013, MNRAS, 429, 2333
Jose C., Srianand R., Subramanian K., 2014, MNRAS, 443, 3341
Jose C., Baugh C. M., Lacey C. G., Subramanian K., 2017, MNRAS, 469, 4428
Kashikawa N. et al., 2006, ApJ, 637, 631
Lacey C. G. et al., 2016, MNRAS, 462, 3854
Lee K.-S., Giavalisco M., Gnedin O. Y., Somerville R. S., Ferguson H. C., Dickinson M., Ouchi M., 2006, ApJ, 642, 63
Leitherer C. et al., 1999, ApJS, 123, 3
Leitherer C., Ekström S., Meynet G., Schaerer D., Agienko K. B., Levesque E. M., 2014, ApJS, 212, 14

Limber D. N., 1954, ApJ, 119, 655
Ling E. N., Barrow J. D., Frenk C. S., 1986, MNRAS, 223, 21P
Liu C., Mutch S. J., Angel P. W., Duffy A. R., Geil P. M., Poole G. B., Mesinger A., Wyithe J. S. B., 2016, MNRAS, 462, 235
McLure R. J., Cirasuolo M., Dunlop J. S., Foucaud S., Almaini O., 2009, MNRAS, 395, 2196
McLure R. J. et al., 2013, MNRAS, 432, 2696
Mesinger A., Furlanetto S., Cen R., 2011, MNRAS, 411, 955
Meurer G. R., Heckman T. M., Calzetti D., 1999, ApJ, 521, 64
Mo H. J., White S. D. M., 1996, MNRAS, 282, 347
Mutch S. J., Geil P. M., Poole G. B., Angel P. W., Duffy A. R., Mesinger A., Wyithe J. S. B., 2016, MNRAS, 462, 250
Norberg P. et al., 2001, MNRAS, 328, 64
Norberg P. et al., 2002, MNRAS, 336, 907
Ocvirk P. et al., 2016, MNRAS, 463, 1462
Oesch P. A. et al., 2012, ApJ, 759, 135
Orsi A., Lacey C. G., Baugh C. M., Infante L., 2008, MNRAS, 391, 1589
Ouchi M. et al., 2004, ApJ, 611, 685
Ouchi M. et al., 2005, ApJ, 635, L117
Park J., Kim H.-S., Wyithe J. S. B., Lacey C. G., Baugh C. M., Barone-Nugent R. L., Trenti M., Bouwens R. J., 2016, MNRAS, 461, 176
Peebles P. J. E., 1980, The Large-scale Structure of the Universe. Princeton Univ. Press, Princeton, NJ
Planck Collaboration XIII, 2016, A&A, 594, A13
Poole G. B., Angel P. W., Mutch S. J., Power C., Duffy A. R., Geil P. M., Mesinger A., Wyithe S. B., 2016, MNRAS, 459, 3025
Roche N., Eales S. A., 1999, MNRAS, 307, 703
Samui S., 2014, New Astron., 30, 89
Samui S., Srianand R., Subramanian K., 2007, MNRAS, 377, 285
Schaye J. et al., 2015, MNRAS, 446, 521
Sobacchi E., Mesinger A., 2013, MNRAS, 432, 3340
Springel V. et al., 2005, Nature, 435, 629
Steidel C. C., Hamilton D., 1993, AJ, 105, 2017
Steidel C. C., Giavalisco M., Dickinson M., Adelberger K. L., 1996, AJ, 112, 352
Tinker J. L., Robertson B. E., Kravtsov A. V., Klypin A., Warren M. S., Yepes G., Gottlöber S., 2010, ApJ, 724, 878
Trenti M., Stiavelli M., 2008, ApJ, 676, 767
Waters D., Di Matteo T., Feng Y., Wilkins S. M., Croft R. A. C., 2016, MNRAS, 463, 3520
White S. D. M., Frenk C. S., 1991, ApJ, 379, 52
Yoshida M. et al., 2006, ApJ, 653, 988
Zehavi I. et al., 2002, ApJ, 571, 172
Zehavi I. et al., 2005, ApJ, 630, 1
Zehavi I. et al., 2011, ApJ, 736, 59
Zitrin A. et al., 2014, ApJ, 793, L12

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