CLUSTERING OF $i_{775}$ DROPOUT GALAXIES AT $z\sim 6$ IN GOODS AND THE UDF

RODERIK A. OVERZIER$^2$, RYCHARD J. BOUWENS$^3$, GARTH D. ILLINGWORTH$^3$ and MARLIN FRANX$^2$

Draft version October 31, 2006

ABSTRACT

We measured the angular clustering at $z \sim 6$ from a large sample of $i_{775}$ dropout galaxies (293 with $5800 \leq 27.5$ from GOODS and 95 with $5800 \leq 29.0$ from the UDF). Our largest and most complete subsample (having $L \gtrsim 0.5 L^*_{i_{775}}$) shows the presence of clustering at 94% significance. For this sample we derive a (co-moving) correlation length of $r_0 = 4.5 \pm 0.2 h_{72}^{-1}$ Mpc and bias $b = 4.1 \pm 1.5$, using an accurate model for the redshift distribution. No clustering could be detected in the much deeper but significantly smaller UDF, yielding $b < 4.4$ (1$\sigma$). We compare our findings to Lyman break galaxies at $z \sim 3-5$ at a fixed luminosity. Our best estimate of the bias parameter implies that $i_{775}$ dropouts are hosted by dark matter halos having masses of $\sim 10^{11} M_{\odot}$, similar to that of $V_{606}$ dropouts at $z \sim 5$. We evaluate a recent claim that $z_{\sim 6}$ star formation might have occurred more efficiently compared to that at $z = 3-4$. This may provide an explanation for the very mild evolution observed in the UV luminosity density between $z = 6$ and 3. Although our results are consistent with such a scenario, the errors are too large to find conclusive evidence for this.

Subject headings: cosmology: observations – early universe – large-scale structure of universe

1. INTRODUCTION

The Advanced Camera for Surveys (ACS; Ford et al. 1995) aboard the Hubble Space Telescope has made the detection of star-forming galaxies at $z \sim 6$ ($i_{775}$ dropouts) relatively easy. The largest sample of $i_{775}$ dropouts currently available (Bouwens et al. 2003) comes from the Great Observatories Origins Deep Survey (GOODS; Giavalisco et al. 2004), allowing the first quantitative analysis of galaxies only 0.9 Gyr after recombination (Stanway et al. 2003; Bouwens et al. 2003; Yan & Windhorst 2004; Dickinson et al. 2004; Mahotra et al. 2004; see also Shimasaku et al. 2004; Ouchi et al. 2005). Bouwens et al. (2006) found evidence for strong evolution of the luminosity function between $z \sim 6$ and 3, while the (unextincted) luminosity density at $z \sim 6$ is only $\sim 0.8$ times lower than that at $z \sim 3$. Some $i_{775}$ dropouts have significant Balmer breaks, indicative of stellar populations older than 100 Myr and masses comparable to those of $L^*$ galaxies at $z \sim 0$ (Evles et al. 2003; Yan et al. 2003).

Through the study of the clustering we can address fundamental cosmological issues that cannot be answered from the study of galaxy light alone. The strength of clustering and its evolution with redshift allows us to relate galaxies with the underlying dark matter and study the bias. The two-point angular correlation function (ACF) has been used to measure the clustering of Lyman break galaxies (LBGs) at $z = 3-5$ (e.g., Adelberger et al. 1998; 2003; Arnouts et al. 1999; 2002; Maiolino et al. 1999; Giavalisco & Dickinson 2001; Ouchi et al. 2001). LBGs are highly biased ($b \sim 2-8$), and this bias depends strongly on rest frame UV luminosity and, to a lesser extent, on dust and redshift. The clustering statistics of LBGs have reached the level of sophistication that one can measure two physically different contributions. At small angular scales the ACF is dominated by the non-linear clustering of galaxies within single dark matter halos, whereas at large scales its amplitude tends to the “classical” clustering of galaxies residing in different halos (Ouchi et al. 2005a; Lee et al. 2006), as explained within the framework of the halo occupation distribution (e.g., Zehavi et al. 2004; Hamana et al. 2004). Understanding the clustering properties of galaxies at $z \sim 6$ is important for the interpretation of “overdensities” observed towards luminous quasars and in the field (Ouchi et al. 2005b; Stiavelli et al. 2005; Wang et al. 2005; Zheng et al. 2006) that could demarcate structures that preceded present-day massive galaxies and clusters (Springel et al. 2003). Our aim here is to “complete” the census of clustering by extending it to the highest redshift regime with sizeable samples. In §§ 2 and 3 we describe the sample, and present our measurements of the ACF. In § 4 we discuss our findings. Throughout we use the cosmology ($\Omega_m$, $\Omega_{\Lambda}$, $h_{72}$, $n$, $\sigma_8$)=$(0.27,0.73,1.0,1.0,0.9)$ with $H_0 = 72 h_{72} \text{ km s}^{-1} \text{ Mpc}^{-1}$.

2. DATA

The present analysis is based on the sample of $i_{775}$ dropouts described in detail by Bouwens et al. (2006). We used the ACS data from the GOODS v1.0 release, consisting of two spatially disjoint, $\sim 160$ arcmin$^2$ fields. These data were processed with Apsis (Blakeslee et al. 2003), along with a substantial amount of overlapping data available from the Galaxy Evolution from Morphology and Spectral energy distributions (GEMS; Rix et al. 2001), supernova searches (A. G. Riess et al. 2006 and S. Perlmutter et al. 2006, both in preparation), and Ultra...
Deep Field (UDF) NICMOS programs \cite{Thompson2003}. The processed images were brought to a uniform signal-to-noise level by degrading the deeper parts of the area. The 10σ detection limit of the degraded data was 27.5 in z_{850} in a 0′.2 diameter aperture. We also used a deep sample of i_{775} dropout sources. The ACFs for various limiting magnitudes. The results of GOODS and the UDF are shown in Figure 1. The effective rest-frame UV luminosity of the sample is \( L \approx 0.5 L^* \) for z_{850}>27.5 \cite{Bouwens2006}. Note that the luminosity is quite sensitive to redshift due to the Gunn-Peterson trough entering z_{850} at z > 6, with \( L^* \) corresponding to z_{850}=26.5 (~28) at z = 5.5 (z = 6.5).

3. THE ANGULAR CORRELATION FUNCTION

We measured the ACF, \( w(\theta) \), defined as the excess probability of finding two sources in the solid angles \( \delta \Omega_1 \) and \( \delta \Omega_2 \) separated by the angle \( \theta \), over that expected for a random Poissonian distribution \cite{Peebles1980}. We used the estimator \( w(\theta) = [D D(\theta) - 2 D R(\theta) + R R(\theta)] / R R(\theta) \) of \cite{Landy1993}, where \( D D(\theta) \), \( D R(\theta) \) and \( R R(\theta) \) are the number of pairs of sources with angular separations between \( \theta \) and \( \theta + \Delta \theta \) measured in the data, random, and data-random cross catalogs, respectively. We used 16 random catalogs containing ~100 times more sources than in the data, but with a similar angular geometry. Errors on \( w(\theta) \) were bootstrapped \cite{Ling1986}. We assumed a power-law ACF of the form \( w(\theta) = A_w \theta^{\beta} \) and determined its amplitude, \( A_w \), by fitting the function \( w(\theta) = A_w \theta^{\beta} IC \). The integral constraint \( IC = \int w(\theta) d\Omega_1 d\Omega_2 / \Omega^2 \), where \( \Omega \) is the survey area, was 0.033A_w for GOODS and 0.074A_w for the UDF. We did not attempt to fit the slope of the ACF and assumed \( \beta = 0.6 \) based on the results of \cite{Lee2006}. The ACF was fitted over the range 10″–300″ (10″–200″ for the UDF), corresponding to roughly 0.4–10 h\(^{-1}\) Mpc comoving at z ≈ 6. The lower value of 10″ is larger than the virial radius of a 10^{12} M_\odot halo to ensure that we are measuring the large-scale clustering (and not receiving a contribution at small scales from the subhalo component). Because the results of the fits are sensitive to the size of the bins used, we determined \( A_w \) using Monte Carlo simulations of the data. Finally, we note that if the contaminants to our samples (~7% of the total) have a uniform distribution, the measured amplitude should be multiplied by ~1.16 to yield the corrected clustering amplitude.

3.1. Results from GOODS and the UDF

Figure 2 (top panels) shows the GOODS and UDF ACFs for various limiting magnitudes. The results of the fits are given in Table 1. For the three brightest subsamples (z_{850}<28.5) we measured a positive signal out to \( \theta \approx 1′ \). In GOODS, we found \( A_w = 2.71 \pm 0.05 \) for z_{850}<27.0, and \( A_w \approx 0.80 \pm 0.09 \) for z_{850}<27.5. The analysis of the UDF is hampered by the relatively small number of sources available, owing to its ~30 times smaller area, although its greater depth (1.5 mag) partially makes up for this lack of area. We found \( A_w = 1.40 \pm 0.64 \) and \( A_w = 0.00 \pm 0.93 \) for the z_{850}<28.5 and z_{850}<29.0 samples, respectively.

Because the objects were selected from data of uniform depth, signal in the ACF is unlikely to be caused by variations in the object surface density. Given the large errors on \( A_w \), it is useful to ask whether the \( w(\theta) \) observed at \( \theta < 1′ \) could be the result of shot noise in a random object distribution. We created 1000 random distributions with the same geometry and the same number of points as our GOODS and UDF data, and calculated the ACF in each of the random samples. The mean and standard deviation at each \( \theta \) is plotted in Figure 2 (top panels, offset by –0.4 for clarity). We calculate the chance of reproducing the observed clustering in the random realizations, using the average \( w(\theta) \) measured over the first four bins (\( \theta < 100″ \)) as a gauge of this clustering. This chance is 0.1% for our z_{850}<27.0 sample, and 6%, 10% and 35% for the fainter samples.

Another test of the clustering signal detected in this sample was as follows. We used the formalism of \cite{Sonier1978} to create mock samples with a choice ACF in two dimensions. A 250′×250′ mock field with surface density similar to that of the i_{775} dropout data was used to mimic the measured \( A_w \) to an accuracy of 98%, determined from a fit. Next, we randomly extracted 100 mock “GOODS” surveys and measured the mean \( w(\theta) \) and its standard deviation using identical binning and fitting to that for the real samples. The result is indicated in Figure 2 (hatched region) for our z_{850}<27.5 sample, which being our largest and most complete sample provides the most reliable constraint on this clustering. The simulation demonstrates that the amplitude of the observed \( w(\theta) \) at \( \theta < 1′ \) lies within \( \pm 1\sigma \) of the amplitudes predicted based on our model ACF, although the scatter in the expected amplitudes is large.

In the above analysis we restricted ourselves to cluster-
ing at $\theta \geq 10''$. Our measurements also showed an excess of pair counts at $\theta < 10''$. Upon closer inspection it was found that the excess was strictly limited to $\theta < 5''$, with $w(2''5) \sim 2.0 \pm 0.9$. The excess is significant with an enhancement of $w(\theta)$ due to subhalo clustering at $\lesssim 30$ kpc to 1.7$\sigma$ confidence, but the exact amplitude cannot be determined accurately due to the small number of pairs (11 pairs at $z_{850} < 28.0$). The excess is similar to that found for Ly$\alpha$ emitters at $z = 5.7$ \cite{Shimasaku2000}. While it is possible that the positive signal out to $\sim 1'$ is the result of strong subhalo clustering (see \cite{Lee2005, Ouchi2005}), the occurrence of such halos becomes increasingly rare with redshift, and by limiting the fits to $\theta \geq 10''$ we minimized any contribution.

4. DERIVATION OF COSMOLOGICAL QUANTITIES

Although the uncertainties are large, we estimate the spatial correlation length ($r_0$) from $A_w$, using the Limber equation adopted for our cosmology and the redshift distributions of Figure 1 (see Table 1). The clustering was assumed to be fixed in comoving coordinates across the redshift range. We found $r_0 = (4.5^{+2.1}_{-3.2}) \, h_7^{-1}$ Mpc for the $z_{850} < 27.5$ sample. At $z_{850} < 27$, the best-fit value was found to be twice as high, $r_0 = (9.6^{+4.6}_{-3.8}) \, h_7^{-1}$ Mpc, but consistent with the fainter subsample within the errors. For the UDF samples, the best-fit values correspond to upper limits for the clustering amplitude of $\sim 10$ and $\sim 5 \, h_7^{-1}$ Mpc for $z_{850} < 28.5$ and $z_{850} < 29$, respectively. If we apply the contamination correction, $r_0$ increases by $\sim 10\%$.

We calculated the galaxy–dark matter bias, defined as $b(\theta) = \sqrt{w(\theta)/w_{dm}(\theta)}$, where $w_{dm}(\theta)$ is the ACF of the dark matter as “seen” through our redshift window; $w_{dm}(\theta)$ was calculated using the nonlinear fitting function of \cite{Peeock1996} (Fig. 2 middle panels). In the bottom panels of Figure 2 we have indicated the bias as a function of $\theta$ (points). Our best-fit ACF at $z_{850} < 27.5$ implies $b(\theta\sim 30'') \sim 4.1^{+1.5}_{-2.6}$ (solid line), bracketed by $b \sim 8$ for the brightest GOODS sample and $b < 4.4$ for the faintest UDF sample. The contamination correction yields values that are $\sim 5\%$ higher.

It is important to evaluate how our results might be influenced by cosmic variance. Using \cite{Somerville2004}, we estimate $\sigma_c \sim 0.2$ for GOODS and $\sigma_c \sim 0.5$ for the UDF ($\sigma_c$ being the square root of the cosmic variance). Our best constraint on clustering at $z = 6$ is therefore currently provided by the $z_{850} < 27.5$ sample, given the relatively small variance, large sample size, and large completeness. Our best value for the bias of $i_{775}$ dropouts is very similar to the bias of $b = 3.4 \pm 1.8$ found for faint Ly$\alpha$ emitters in the Subaru/XMM-Newton Deep Field \cite{Ouchi2005}. An alternative method of estimating the bias is to directly compare the $i_{775}$ dropout
The clustering signal in the UDF is likely further diminished due to the small number statistics, as well as large cosmic variance. Also, the decrease in the effective halo mass from $z = 4$ to 5 at $M_V \lesssim -19.5$ is not as dramatic as observed at luminosities of $M_V \lesssim -20$, making it difficult to verify this result based on the present i775 dropout sample. However, we note that if a decrease in the star-forming efficiency with decreasing redshift is true (and can be confirmed for galaxies at $z \gtrsim 6$), it would largely offset changes that are occurring in the mass function over this range. As such, this may provide at least a partial explanation for the mild evolution in the luminosity density from $z = 6$ to 3.

In conclusion, we used the largest available sample of i775 dropouts to study clustering at $z \sim 6$. We found a small signal, although its amplitude is not well constrained due to the large errors on the individual data points. The present analysis is therefore reminiscent of that performed at $z \sim 3-5$ based on the original Hubble Deep Fields. The clustering of galaxies at $z \sim 6$ will continue to be studied from deep, wide surveys (e.g., see Ouchi et al. 2005, Shimakawa et al. 2003, 2004). Although it might become possible in the near future to increase the size of our faint ACS samples by relaxing our current i775 dropout detection threshold, to perform an analysis at the same level of detail as currently performed at $z \sim 5$ would require another six GOODS fields, for $\sim$1200 arcmin$^2$ in total.

We thank the referee, Masami Ouchi, for a thorough report and helpful comments. ACS was developed under NASA contract NAS 5-32865, and this research has been supported by NASA grant NAG5-7097.

**REFERENCES**


![Fig. 3.— Bias of i775 dropouts compared to bias at $z = 3 - 5$ from Lee et al. (2006). Dashed lines indicate the bias of dark halos from Sheth & Tormen (1999) for $M_{halo} > 10^{10}$, $5 \times 10^{10}$, $10^{11}$, $5 \times 10^{11}$, and $10^{12} M_\odot$ (bottom to top). The relatively small halo mass inferred at $z = 5$ compared to that at $z = 4$ for objects with $M_V \lesssim -20$ (squares) cannot be confirmed at $z = 6$ based on the present data. The best-fit halo mass inferred for objects at $z = 6$ is consistent with the average halo mass of V606 dropouts at $z = 5$ at a fixed luminosity of $M_V < -19.5$ (circles). The star indicates the bias of Ly$\alpha$ emitters at $z = 5.7$ from Ouchi et al. (2005b).]{fig3.pdf}