

# UNCERTAINTY ON THE GALAXY-HALO CONNECTION FOR LYMAN- $\alpha$ EMITTERS AT $Z = 3.1$

JULIÁN E. MEJÍA-RESTREPO<sup>1,3</sup>, JAIME E. FORERO-ROMERO<sup>2</sup>

<sup>1</sup>Departamento de Astronomía, Universidad de Chile, Camino el Observatorio 1515, Santiago, Chile

<sup>2</sup>Departamento de Física, Universidad de los Andes, Cra. 1 No. 18A-10, Edificio Ip, Bogotá, Colombia

<sup>3</sup>FACom-Instituto de Física-FCEN, Universidad de Antioquia, Calle 70 No. 52-21, Medellín, Colombia

Submitted for publication in *ApJ*

## ABSTRACT

We study the impact of cosmic variance and observational uncertainties in constraining the mass and occupation fraction ( $f_{\text{occ}}$ ) of dark matter halos hosting Ly $\alpha$  Emitting Galaxies at high redshift (LAEs). We use a N-body simulation to construct mock catalogs with the same typical size of observed fields at  $z = 3.1$  ( $\sim 1\text{deg}^2$ ) to match the observed angular correlation function (ACF) and number density of LAEs. In our model a dark matter halo with mass in the range  $M_{\text{min}} < M_h < M_{\text{max}}$  can only host one detectable LAE at most. By following a thorough Markov Chain Monte-Carlo exploration of the parameter space determined by  $M_{\text{min}}$  and  $M_{\text{max}}$ , our analysis shows that  $f_{\text{occ}}$  is uniquely determined by  $M_{\text{min}}$  regardless of  $M_{\text{max}}$  using the relation  $f_{\text{occ}} \sim 0.1 (M_{\text{min}}/10^{10.5})^{0.93}$ . However, the current observational data only allows us to put weak constraints on  $M_{\text{min}}$ ,  $M_{\text{max}}$  and consequently on  $f_{\text{occ}}$ . Particularly, we find that  $10^{9.6}h^{-1}M_{\odot} \leq M_{\text{min}} \leq 10^{11.0}h^{-1}M_{\odot}$ ,  $10^{10.9}h^{-1}M_{\odot} \leq M_{\text{max}} \leq 10^{13.0}h^{-1}M_{\odot}$  and  $0.02 \leq f_{\text{occ}} \leq 0.30$ . Nevertheless, we also find that the upcoming large surveys can help to put tighter constraints on  $M_{\text{min}}$  and  $f_{\text{occ}}$  through the width measurement of the LAE number distribution function obtained over several fields of the same size of current observations. Analogously, the improvement in the precision of the measured ACF will assist in constraining  $M_{\text{max}}$ .

*Subject headings:* Cosmology: theory - large-scale structure of Universe - Methods: data analysis - numerical - N-body simulations

## 1. INTRODUCTION

(●REFS)ADD NEWER REFERENCES Lyman- $\alpha$  emitting galaxies (LAEs) are central to a wide range of subjects in extragalactic astronomy. LAEs can be used as probes of reionization (Dijkstra et al. 2011), tracers of large scale structure (Koehler et al. 2007), signposts for low metallicity stellar populations, markers of the galaxy formation process at high redshift (Dayal et al. 2009; Forero-Romero et al. 2012) and tracers of active star formation (Guaita et al. 2013).

In most of those cases, capitalizing the observations requires understanding how LAEs are formed within an explicit cosmological context. Under the current structure formation paradigm the dominant matter content of the Universe is dark matter (DM). Each galaxy is thought to be hosted by a larger dark matter structure known as a halo. (Peebles 1980; Springel et al. 2005). Understanding the cosmological context of LAEs thus implies studying the galaxy-halo connection. Galaxy formation models suggest that the physical processes that regulate the star formation cycle are dependent on halo mass (Behroozi et al. 2013). The mass becomes the most important element in the halo-galaxy connection.

The goal becomes finding the typical DM halo mass of halos hosting LAEs. In the case of LAEs there are different ways to find this mass range. One approach is theoretical, using general astrophysical principles to find the relationship between halo mass, intrinsic Ly $\alpha$  luminosities and observed Ly $\alpha$  luminosities. This approach is usually implemented through semi-analytic models (Garel et al. 2012; Orsi et al. 2012) and full N-body hydrodynamical simulations (Laursen & Sommer-Larsen 2007; Dayal et al. 2009; Forero-Romero et al. 2011; Ya-

jima et al. 2012).

The downside of these calculations is the uncertainty in the estimation of the escape fraction of Ly $\alpha$  photons. Given the resonant nature of the Ly $\alpha$  line, the escape fraction is sensitive to the dust contents, density, temperature, topology and kinematics of the neutral Hydrogen in the interstellar medium (ISM). The process of finding a consensus on the expected value for the Ly $\alpha$  escape fraction in high redshift galaxies is still matter of open debate (Neufeld 1991; Verhamme et al. 2006; Forero-Romero et al. 2011; Dijkstra & Kramer 2012; Laursen et al. 2013; Orsi et al. 2012).

A different approach to infer the typical mass of halos hosting LAEs is based on the spatial clustering information. This approach uses the fact that in CDM cosmologies the spatial clustering of galaxies on large scales is entirely dictated by the halo distribution (Colberg et al. 2000), which in turn has a strong dependence on halo mass. Using measurements of the angular correlation function of LAEs, observers have put constraints on the typical mass and occupation fraction of the putative halos hosting these galaxies (Hayashino et al. 2004; Gawiser et al. 2007b; Nilsson et al. 2007; Ouchi et al. 2010; Bielby et al. 2016). In these studies the observations are done on fields of  $\sim 1\text{deg}^2$  and the conclusions derived on the halo host mass do not elaborate on the uncertainty resulting from the cosmic variance on these fields.

In this letter we investigate the impact of cosmic variance in constraining the mass and occupation fraction of halos hosting LAEs at  $z = 3$ . We build mock surveys from a cosmological N-body to compare them against the observations in Bielby et al. (2016) using the angular correlation function. We use a simple model to populate a

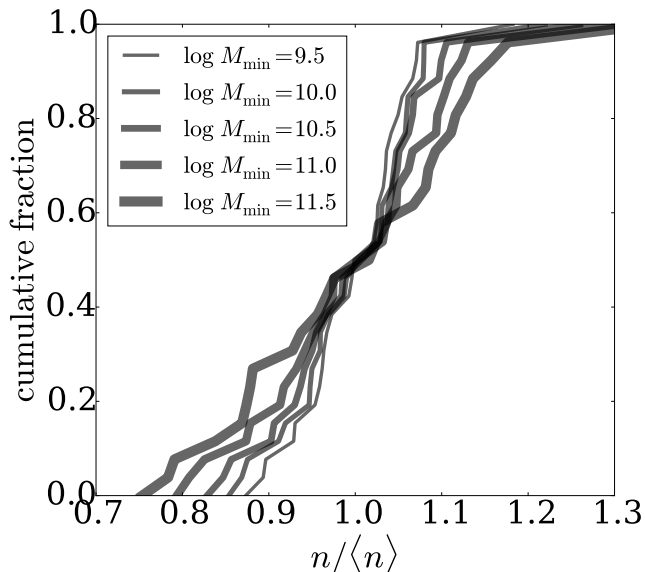


FIG. 1.— Cumulative halo number density distribution function over 27 mock fields. Each line corresponds to a different model with increasing values of  $M_{\min}$ . Different models produce different number density distributions. The width increases with  $M_{\min}$ . halo in the simulation with a LAE assuming a minimum ( $M_{\min}$ ) and maximum mass ( $M_{\max}$ ) for the dark matter halos hosting LAEs without predicting a Ly $\alpha$  luminosity. This approach bypasses all the physical uncertainties associated to star formation and radiative transfer. We use the Markov Chain Monte Carlo technique to obtain the likelihood of the parameters given the observational constraints.

Throughout this letter we assume a  $\Lambda$ CDM cosmology with the following values for the cosmological parameters,  $\Omega_m = 0.30711$ ,  $\Omega_\Lambda = 0.69289$  and  $h = 0.70$ , corresponding to the matter density, vacuum density and the Hubble constant in units of  $100 \text{ km s}^{-1} \text{ Mpc}^{-1}$ .

## 2. METHODOLOGY

The base of our method is the comparison between observations and mock catalogs. This approach allows us to take explicitly into account cosmic variance. The comparison has four key elements. First, the observations we take as a benchmark. Second, the N-body simulation and the halo catalogs we use to build the mocks. Third, the parameters describing our model to assign a LAE to a halo. Fourth, the statistical method we adopt to compare observations and simulations. We describe in detail these four elements in the following subsections.

### 2.1. Observational constraints

Bielby et al. (2016) used narrow band imaging to detect 643 LAE candidates at  $z \sim 3$  with equivalent widths of  $\gtrsim 65\text{\AA}$  and a flux limit of  $2 \times 10^{17} \text{ erg/cm}^2/\text{s}$  ( $L \sim 7 \times 10^{42} \text{ erg/s}$ ). Using spectroscopy they found a 22% contamination fraction. Their observations cover 5 (out of 9) independent and co-spatial fields of the VLT LBG Redshift Survey (VLRS). The total observed area corresponds to  $1.07 \text{ deg}^2$  that translates to  $\sim 80^2 h^{-1} \text{ Mpc}^2$  in a comoving scale. Bielby et al. (2016) used the NB497 narrow-band filter whose  $77\text{\AA}$  FWHM and  $154\text{\AA}$  FWTM correspond to a total observational depth of  $44 h^{-1} \text{ Mpc}$  and  $82 h^{-1} \text{ Mpc}$  comoving, respectively.

### 2.2. Simulation and halo catalog

We use results from the Bolshoi simulation (Klypin et al. 2011, 2014) which was performed in a cubic volume of  $250 h^{-1} \text{ Mpc}$  comoving on a side. The dark matter distribution was sampled using 2048<sup>3</sup> particles. The cosmological parameters are consistent with Planck results (Planck Collaboration et al. 2014) with a matter density  $\Omega_m = 0.307$ , cosmological constant  $\Omega_\Lambda = 0.693$ , dimensionless Hubble constant  $h = 0.678$ , slope of the power spectrum  $n = 0.96$  and normalization of the power spectrum  $\sigma_8 = 0.823$ . This translates into a particle mass of  $m_p = 1.5 \times 10^8 h^{-1} \text{ M}_\odot$ .

We use halo catalogs constructed with a Bound-Density-Maxima (BDM) algorithm. The catalogs were obtained from the publicly available Multidark database<sup>1</sup> (Riebe et al. 2013). For each halo in the box we extract its comoving position and mass. We focus our work on halos more massive than  $1.5 \times 10^9 h^{-1} \text{ M}_\odot$  resolved with at least 10 particles.

We split the simulation volume at  $z \sim 3$  into 27 smaller mock volumes mimicking the area and depth reported in Bielby et al. (2016) and described in §2.1.

### 2.3. A model to populate halos with LAEs

We build a model to assign LAEs to a DM halo. We are not concerned on the exact LAE luminosity, we only care about a yes/no answer to the following question. Does this halo host a LAE?

We first assume that a dark matter halo can host one detectable LAE at most. This assumption is consistent with theoretical analysis of the correlation function (Jose et al. 2013) and observations that confirm a lack of class pairs in LAEs Bond et al. (2009).

Then we say that a halo will host a LAE with probability  $f_{\text{occ}}$  if and only if the halo mass is in the range  $M_{\min} < M_h < M_{\max}$ . The probability  $f_{\text{occ}}$  can be thought as the occupation fraction of halos which can be automatically set as the ratio of the observed number of LAEs to the number of halos within the considered mass range, that is  $f_{\text{occ}} \equiv N_{\text{LAE}}/N_{\text{halos}}$ .

Fig. 1 shows the cumulative halo number distribution (HND) in the mock fields of the simulation for different models  $\mathcal{M}$ . From this figure we can see that the number of dark matter halos along the mock fields for different models  $\mathcal{M}$  varies within a factor of  $\sim 2$  ( $\sim 0.3 \text{ dex}$ ) tracing the cosmic variance. As a consequence, the occupation fraction varies along the mock fields by the same factor. The interpretation of the occupation fraction  $f_{\text{occ}}$  involves two phenomena: the actual presence of a star forming galaxy in a halo and its detectability as a LAE.

We do not explore any physical model to disentangle these two effects. This means that our model does not assign a luminosity or escape fraction to each LAE. We avoid this in order to maintain theoretical uncertainties to a minimum. This flexibility allows us to explore a wide range of possible masses for the host halos without any strong theoretical prejudice regarding the details of star formation and Ly $\alpha$  escape fraction.

We also artificially contaminate our mock catalogs with 22% randomly distributed data points to mimic the frac-

<sup>1</sup> <http://www.multidark.org/MultiDark/>

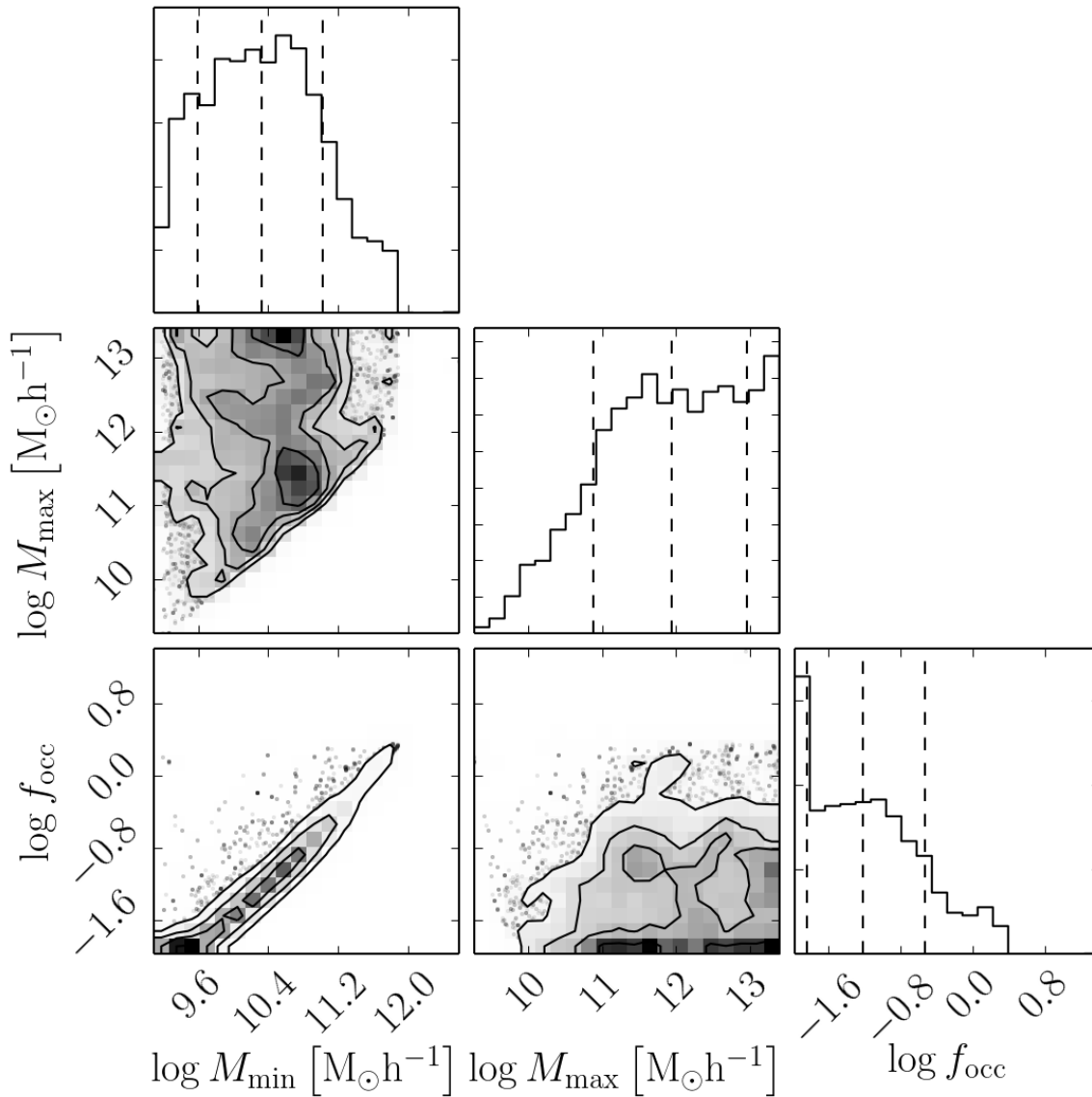


FIG. 2.— One and two dimensional projections of the posterior probability distributions of  $M_{\min}$ ,  $M_{\max}$  and  $f_{\text{occ}}$ . The models with  $\log f_{\text{occ}} > 0.00$  ( $f_{\text{occ}} > 1$ ) correspond to models where the number density of halos is smaller than the number of observed LAEs but we consider them as consistent because of the uncertainty in the median number density due to cosmic variance. See Fig. 1 and §2.4 for details.

tion of interlopers in observations Bielby et al. (2016). On top of that we apply rejection sampling to our LAE selection taking the transmission function of the NB479 filter used in their observations as a radial selection factor.

In what follows we note by the letter  $\mathcal{M}$  a model defined by a particular choice of the two parameters  $M_{\min}$ ,  $M_{\max}$ . For each model  $\mathcal{M}$  we define  $\tilde{f}_{\text{occ}}$  as the median occupation fraction within the mock fields and  $\Delta M \equiv M_{\max} - M_{\min}$ .

#### 2.4. Exploring and selecting consistent models

We use the angular correlation function to compare the observations with our mock catalogs. This is done through a thorough exploration of the parameter space of the models  $\mathcal{M}$  by means of a Monte Carlo Markov Chain minimization using the EMCEE python package

(Foreman-Mackey et al. 2013). We put a flat prior on  $\log M_{\min}$  and  $\log M_{\max}$  to vary between 9.2 up to 13.4, corresponding to the halo mass range of the simulation at  $z = 3$ . Given that the typical scatter in  $N_{\text{halos}}$  is about 0.3 dex (i.e. about a factor of 2), our parameter space is restricted to models where the median number of dark matter halos is larger than  $N_{\text{LAE}}/3$ . Because of this possibility of having  $N_{\text{halos}}$  smaller than  $N_{\text{LAE}}$  there will be some models where  $f_{\text{occ}} \equiv N_{\text{LAE}}/N_{\text{halo}} > 1$ . Those models only reflect the uncertainty in the number density of LAEs due to cosmic variance.

The MCMC exploration is done using a total of 24 seeds and 400 iterations (9600 models) to sample the posterior PDF,  $P(M_{\min}, M_{\max}, f_{\text{occ}}|\mathcal{M}) \propto \exp(-\chi_{\mathcal{M}}^2/2)$ , with

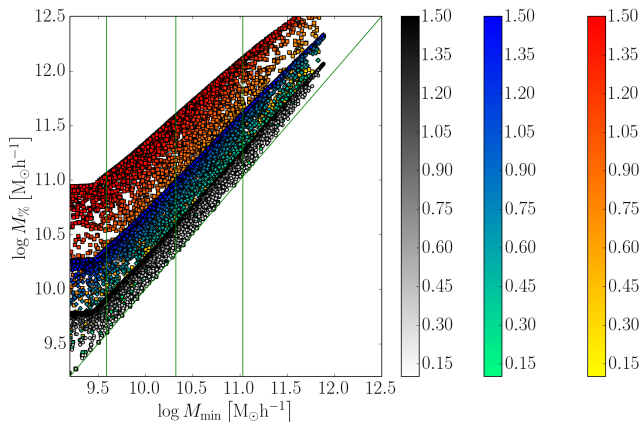


FIG. 3.—  $\log M_{\min}$  vs  $\log M_{50\%}$  (black),  $\log M_{16\%}$ - $\log M_{84\%}$  (blue) and  $\log M_{2.5\%}$ - $\log M_{97.5\%}$  (red) for different values of  $\log M_{\max}$ . The points are color coded by  $\Delta \log M$  [ $M_{\odot} h^{-1}$ ]. The green line diagonal line accounts for the 1:1 relation and the green vertical lines represent the 16, 50 and 84 percentiles of the posterior probability distribution of  $M_{\min}$ .

$$\chi^2_{\mathcal{M}} = \sum_{\theta} \left[ \frac{(\text{ACF}_{\mathcal{M}}(\theta) - \text{ACF}_{\text{obs}}(\theta))^2}{\sigma_{\mathcal{M}}^2(\theta) + \sigma_{\text{obs}}^2(\theta)} \right] \quad (1)$$

where  $\text{ACF}_{\mathcal{M}}$  and  $\text{ACF}_{\text{obs}}$  are the ACF of the explored model  $\mathcal{M}$  and the observational ACF reported by Bielby et al. (2016) respectively.  $\sigma_{\mathcal{M}}$  is the associated  $1\text{-}\sigma$  scatter of the  $\text{ACF}_{\mathcal{M}}$  as a product of cosmic variance and  $\sigma_{\text{obs}}$  is the observational error associated to  $\text{ACF}_{\text{obs}}$ . The  $\text{ACF}_{\mathcal{M}}$ s are computed using the Landy & Szalay estimator (Landy & Szalay 1993).

### 3. RESULTS AND DISCUSSION

#### 3.1. Constraining $M_{\min}$ , $M_{\max}$ and $f_{\text{occ}}$ in Dark Matter Halos hosting LAEs

In Fig. 2 we present the one and two dimensional projections of the posterior probability distributions of  $M_{\min}$ ,  $M_{\max}$  and  $f_{\text{occ}}$ . We find that the  $M_{\min}$  and  $M_{\max}$  parameter space cannot be tightly constrained mainly because of cosmic variance and the large observational Poissonian uncertainty in the ACF. We however found that  $9.6 < \log M_{\min} < 11.0$  and  $10.9 < \log M_{\max} < 13.0$ . We interestingly find that  $f_{\text{occ}}$  is basically completely determined by  $M_{\min}$  going from  $f_{\text{occ}}=0.015$  when  $\log M_{\min}=9.6$  to  $f_{\text{occ}}=0.30$  when  $\log M_{\min}=11.0$  ( $f_{\text{occ}} \sim 0.1 (M_{\min}/10^{10.5})^{0.93}$ ). We remark that those models where  $\log f_{\text{occ}} > 0.00$  ( $f_{\text{occ}} > 1$ ) correspond to models where the number density of halos is smaller than the number of observed LAEs but are still consistent because of the uncertainty in the median number density of LAEs in the universe due to cosmic variance (see Fig. 1 and §2.4).

#### 3.2. Halo mass distribution within models

Previous attempts to constrain the DMH mass of LAEs by means of the ACF present their result in terms of the median mass of LAEs that is derived from the derived correlation length by means of semi-analytical prescriptions (e.g. Hayashino et al. 2004; Gawiser et al. 2007a; Ouchi et al. 2010; Bielby et al. 2016). Instead, in this work we directly compare the ACF of our simulated mock

catalogs with the observational ACF and present our results in terms of a minimum and maximum mass DMH hosting LAEs. It is therefore necessary to connect our approach by computing the median mass of the DMH in each of the models  $\mathcal{M}$ . In Fig. 3 we show the 50 ( $\log M_{50}$ , black dots), the 84 ( $\log M_{84}$ , blue diamonds) and 95 ( $\log M_{95}$ , red squares) percentiles of the LAE halo mass as a function of  $\log M_{\min}$  for each of the models that we run in our MCMC simulation. The points are color coded according to their  $\Delta \log M$  associated value. We can see that the median mass ( $M_{50}$ ) and  $M_{84}$  are not very sensitive to  $\log M_{\max} = \log M_{\min} + \Delta \log M$  specially when  $\Delta \log M \gtrsim 1.0\text{dex}$ . We particularly found that  $\log M_{\min} \lesssim \log M_{50} \lesssim \log M_{\min} + 0.2\text{dex}$  and that  $\log M_{\min} + 0.1\text{dex} \lesssim \log M_{50} \lesssim \log M_{\min} + 0.5\text{dex}$  regardless of the value of  $\Delta \log M$ . The latter is a consequence of the very steepen distribution of the dark matter mass function and is at the same time the reason for the almost perfect one to one relation between  $M_{\min}$  and  $f_{\text{occ}}$ . However, it can also be seen in Fig. 3 that  $\log M_{84}$  is very sensitive to  $\log M_{\max}$  ( $\log M_{\min} + 0.2\text{dex} \lesssim \log M_{84} \lesssim \log M_{\min} + 1.5\text{dex}$ ). Thereby, any differences in the clustering strength of models sharing the same  $M_{\min}$  but different  $M_{\max}$  should be mainly driven by the  $\sim 16\%$  most massive halos of each model  $\mathcal{M}$ .

In order to estimate the effect of most massive halos in  $\text{ACF}_{\mathcal{M}}$  in Fig. 4 we show the computed  $\text{ACF}_{\mathcal{M}}$  of models with  $\log M_{\min} = 0.5$  and different values of  $\Delta \log M$ . We can see that the clustering gets stronger for larger values of  $\Delta \log M$ . Nevertheless, due to the large impact of cosmic variance at the volume of the current observations all the models are basically consistent within errors. The last result together with the large Poissonian observational error in the ACF explain the current difficulty to put tighter constraints in  $\log M_{\max}$  in our model.

In Fig. 3 the green vertical lines represent the 16, 50 and 84 percentiles of the posterior probability distribution of  $M_{\min}$  that we previously obtained. From this figure we find that the median dark matter halo mass is  $\tilde{M}_h = 10^{10.6 \pm 0.7}$ . This result is consistent with previous estimations of the median dark matter halo mass reported by Bielby et al. (2016) ( $\tilde{M}_h = 10^{11.0 \pm 0.6}$ ), Gawiser et al. (2007b) ( $\tilde{M}_h = 10^{10.9 \pm 0.9}$ ) and Ouchi et al. (2010) ( $\tilde{M}_h = 6.7^{+42.0}_{-6.7} \times 10^{10}$ ) using semi-analytical approaches.

#### 3.3. Constraining Dark matter halos mass with cosmic variance

Fig. 1 shows the halo number distribution (HND) in the mock fields of the simulation for different models  $\mathcal{M}$ . By simple inspection one can infer an increase in the distribution with as  $\log M_{\min}$  increase. This trend is confirmed in Fig. 5 where we plot the central  $1\text{-}\sigma$  (blue diamonds) and  $2\text{-}\sigma$  (red dots) widths of the NDH as a function of  $\log M_{\min}$ . We particularly found that when we consider survey fields of  $\sim 1\text{deg}^2$ , the central  $1\text{-}\sigma$  ( $2\text{-}\sigma$ ) width of the HND increases monotonically from  $0.05\text{dex}$  ( $0.10\text{dex}$ ) when  $\log M_{\min} = 9.5$  to  $0.20\text{dex}$  ( $0.35\text{dex}$ ) when  $\log M_{\min} = 12.0$ . The latter result opens the possibility to constrain the  $\log M_{\min}$  (as well as the median mass) of LAEs by simply measure the width of the distribution of observed LAE along several observational fields mapping  $\sim 1\text{deg}^2$  in area and the observational depth determined by the NB479 filter.

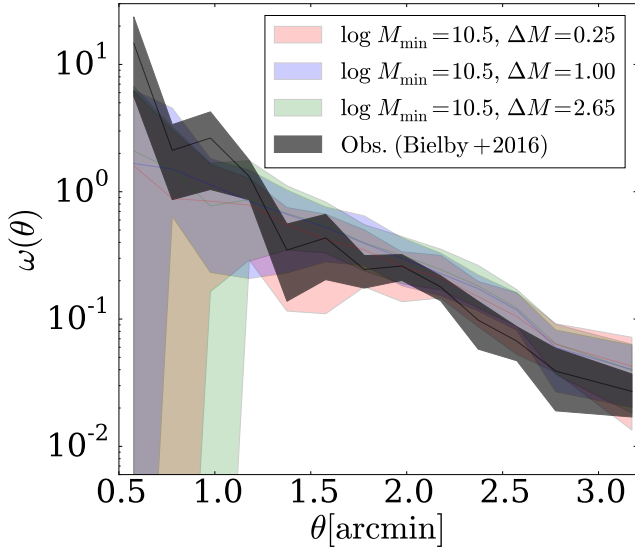


FIG. 4.— Angular Correlation functions for  $\log M_{\min}[\text{M}_{\odot} h^{-1}] = 10.5$  and different values of  $\Delta M$ . The shaded region represents the 1-sigma deviation due to cosmic variance. Radical different models in  $\Delta M$  are consistent with observations once cosmic variance is modelled in detail.

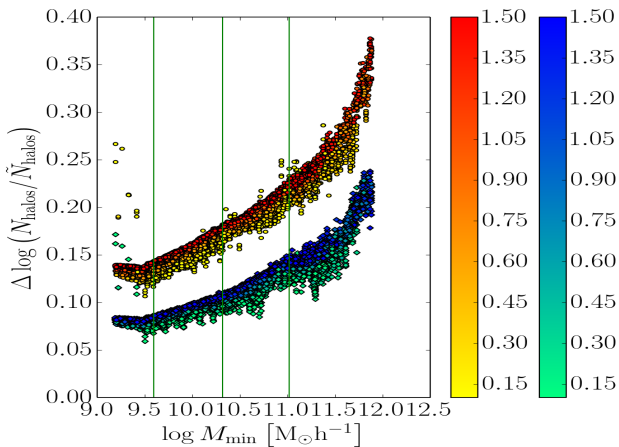


FIG. 5.— Width of the halo number distribution over the 54 mock field of the simulation ( $\Delta \log(N_{\text{halos}}/\tilde{N}_{\text{halos}})$ ) of the central 68 (blue diamonds) and 95 (red circles) percentiles vs  $M_{\min}$ . Points are color coded according to their associated value of  $\Delta M \equiv M_{\min} - M_{\text{min}}$ . The green line diagonal line accounts for the 1:1 relation and the green vertical lines represent the 16, 50 and 84 percentiles of the posterior probability distribution of  $M_{\min}$ .

#### 4. CONCLUSIONS

In this letter we studied the impact of cosmic variance and observational uncertainties in constraining the mass range and occupation fraction of dark matter halos hosting LAEs. To this end, we used the BolshoiP N-body simulation to construct 27 mock fields with the

same typical size of observed fields at  $z = 3.1$  ( $\sim 1 \text{ deg}^2$ ). In our model a dark matter halo with mass in the range  $M_{\min} < M_h < M_{\max}$  can only host one detectable LAE at most. We explored the parameter space determined by  $M_{\min}$  and  $M_{\max}$  using Monte Carlo Markov-Chain minimization to match the observed ACF and mean number density of LAEs.

Our analysis allowed us to put weak constraints on  $M_{\min}$ ,  $M_{\max}$  and  $f_{\text{occ}}$  where  $10^{9.7} h^{-1} \text{M}_{\odot} \leq \log M_{\min} \leq 10^{11.2} h^{-1} \text{M}_{\odot}$ ,  $10^{10.9} h^{-1} \text{M}_{\odot} \leq \log M_{\max} \leq 10^{13.0} h^{-1} \text{M}_{\odot}$  and  $0.02 h^{-1} \text{M}_{\odot} \leq f_{\text{occ}} \leq 0.5$ , spanning three orders of magnitude in halo mass. Previous works (Hayashino et al. 2004; Gawiser et al. 2007b; Ouchi et al. 2008; Bielby et al. 2016) have found typical masses within somehow narrower mass ranges ( $\sim 10^{10.5}-10^{12}$ ). The main reason for our weaker constraints and large discrepancies with previous works resides in the cosmic variance on the typical field size in observations. A thorough exploration of cosmic variance impact as we present in this letter had not been presented so far in the literature.

Our analysis also allowed us to draw three results that can be used to put tighter constraints on  $M_{\min}$ ,  $M_{\max}$  and  $f_{\text{occ}}$  once upcoming large LAE surveys, as the HETDEX project (Adams et al. 2011) and the HSC ultra deep survey, are available :

1.  $f_{\text{occ}}$  and the median mass of LAEs are almost uniquely determined by  $M_{\min}$  regardless of  $M_{\max}$ .
2. measuring the width of the LAE number distribution function obtained over several fields of  $\approx 1 \text{ deg}^2$  one will be able to tightly constrain  $M_{\min}$ ,  $M_{\text{med}}$  and  $f_{\text{occ}}$  up to a factor of 2.
3.  $M_{\max}$  drives the ACF strength at short angular distances ( $< 1 \text{ arcmin}$ ), therefore precise measurements of the ACF at such scales are crucial to accurately determine  $M_{\max}$ .

#### ACKNOWLEDGMENTS

JM acknowledges “CONICYT-PCHA/doctorado Nacional para extranjeros/2013-63130316” for their PhD scholarship support.

The authors gratefully acknowledge the Gauss Centre for Supercomputing e.V. ([www.gauss-centre.eu](http://www.gauss-centre.eu)) and the Partnership for Advanced Supercomputing in Europe (PRACE, [www.prace-ri.eu](http://www.prace-ri.eu)) for funding the MultiDark simulation project by providing computing time on the GCS Supercomputer SuperMUC at Leibniz Supercomputing Centre (LRZ, [www.lrz.de](http://www.lrz.de)). The Bolshoi simulations have been performed within the Bolshoi project of the University of California High-Performance Astro-Computing Center (UC-HiPACC) and were run at the NASA Ames Research Center.

#### REFERENCES

- Adams, J. J., Blanc, G. A., Hill, G. J., Gebhardt, K., Drory, N., Hao, L., Bender, R., Byun, J., Ciardullo, R., Cornell, M. E., Finkelstein, S. L., Fry, A., Gawiser, E., Gronwall, C., Hopp, U., Jeong, D., Kelz, A., Kelzenberg, R., Komatsu, E., MacQueen, P. J., Murphy, J., Odms, P. S., Roth, M., Schneider, D. P., Tufts, J. R., & Wilkinson, C. P. 2011, *ApJS*, 192, 5
- Bielby, R. M., Tummuangpak, P., Shanks, T., Francke, H., Crighton, N. H. M., Bañados, E., González-López, J., Infante, L., & Orsi, A. 2016, *MNRAS*, 456, 4061
- Bond, N. A., Gawiser, E., Gronwall, C., Ciardullo, R., Altmann, M., & Schawinski, K. 2009, *ApJ*, 705, 639
- Brozgi, P. S., Wechsler, R. H., & Conroy, C. 2013, *ApJ*, 770, 57

- Colberg, J. M., White, S. D. M., Yoshida, N., MacFarland, T. J., Jenkins, A., Frenk, C. S., Pearce, F. R., Evrard, A. E., Couchman, H. M. P., Efstathiou, G., Peacock, J. A., Thomas, P. A., & Virgo Consortium. 2000, *MNRAS*, 319, 209
- Dayal, P., Ferrara, A., Saro, A., Salvaterra, R., Borgani, S., & Tornatore, L. 2009, *MNRAS*, 400, 2000
- Dijkstra, M., & Kramer, R. 2012, *MNRAS*, 424, 1672
- Dijkstra, M., Mesinger, A., & Wyithe, J. S. B. 2011, *MNRAS*, 414, 2139
- Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, *PASP*, 125, 306
- Forero-Romero, J. E., Yepes, G., Gottlöber, S., Knollmann, S. R., Cuesta, A. J., & Prada, F. 2011, *MNRAS*, 415, 3666
- Forero-Romero, J. E., Yepes, G., Gottlöber, S., & Prada, F. 2012, *MNRAS*, 419, 952
- Garel, T., Blaizot, J., Guiderdoni, B., Schaerer, D., Verhamme, A., & Hayes, M. 2012, *MNRAS*, 422, 310
- Gawiser, E., Francke, H., Lai, K., Schawinski, K., Gronwall, C., Ciardullo, R., Quadri, R., Orsi, A., Barrientos, L. F., Blanc, G. A., Fazio, G., & Feldmeier, J. J. 2007a, *ApJ*, 671, 278
- Gawiser, E., Francke, H., Lai, K., Schawinski, K., Gronwall, C., Ciardullo, R., Quadri, R., Orsi, A., Barrientos, L. F., Blanc, G. A., Fazio, G., Feldmeier, J. J., Huang, J.-s., Infante, L., Lira, P., & Padilla, N. 2007b, *ApJ*, 671, 278
- Guaita, L., Francke, H., Gawiser, E., Bauer, F. E., Hayes, M., Östlin, G., & Padilla, N. 2013, *A&A*, 551, A93
- Hayashino, T., Matsuda, Y., Tamura, H., Yamauchi, R., Yamada, T., Ajiki, M., Fujita, S. S., Murayama, T., Nagao, T., Ohta, K., Okamura, S., Ouchi, M., Shimasaku, K., Shioya, Y., & Taniguchi, Y. 2004, *AJ*, 128, 2073
- Jose, C., Srianand, R., & Subramanian, K. 2013, *ArXiv e-prints*
- Klypin, A., Yepes, G., Gottlober, S., Prada, F., & Hess, S. 2014, *ArXiv e-prints*
- Klypin, A. A., Trujillo-Gomez, S., & Primack, J. 2011, *ApJ*, 740, 102
- Koehler, R. S., Schuecker, P., & Gebhardt, K. 2007, *A&A*, 462, 7
- Landy, S. D., & Szalay, A. S. 1993, *ApJ*, 412, 64
- Laursen, P., Duval, F., & Östlin, G. 2013, *ApJ*, 766, 124
- Laursen, P., & Sommer-Larsen, J. 2007, *ApJ*, 657, L69
- Neufeld, D. A. 1991, *ApJ*, 370, L85
- Nilsson, K. K., Möller, P., Möller, O., Fynbo, J. P. U., Michałowski, M. J., Watson, D., Ledoux, C., Rosati, P., Pedersen, K., & Grove, L. F. 2007, *A&A*, 471, 71
- Orsi, A., Lacey, C. G., & Baugh, C. M. 2012, *MNRAS*, 425, 87
- Ouchi, M., Shimasaku, K., Akiyama, M., Simpson, C., Saito, T., Ueda, Y., Furusawa, H., Sekiguchi, K., Yamada, T., Kodama, T., Kashikawa, N., Okamura, S., Iye, M., Takata, T., Yoshida, M., & Yoshida, M. 2008, *ApJS*, 176, 301
- Ouchi, M., Shimasaku, K., Furusawa, H., Saito, T., Yoshida, M., Akiyama, M., Ono, Y., Yamada, T., Ota, K., Kashikawa, N., Iye, M., Kodama, T., Okamura, S., Simpson, C., & Yoshida, M. 2010, *ApJ*, 723, 869
- Peebles, P. J. E. 1980, *The large-scale structure of the universe*
- Planck Collaboration, Ade, P. A. R., Aghanim, N., Armitage-Caplan, C., Arnaud, M., Ashdown, M., Atrio-Barandela, F., Aumont, J., Baccigalupi, C., Banday, A. J., & et al. 2014, *A&A*, 571, A16
- Riebe, K., Partl, A. M., Enke, H., Forero-Romero, J., Gottlber, S., Klypin, A., Lemson, G., Prada, F., Primack, J. R., Steinmetz, M., & Turchaninov, V. 2013, *Astronomische Nachrichten*, 334, 691
- Springel, V., White, S. D. M., Jenkins, A., Frenk, C. S., Yoshida, N., Gao, L., Navarro, J., Thacker, R., Croton, D., Helly, J., Peacock, J. A., Cole, S., Thomas, P., Couchman, H., Evrard, A., Colberg, J., & Pearce, F. 2005, *Nature*, 435, 629
- Verhamme, A., Schaerer, D., & Maselli, A. 2006, *A&A*, 460, 397
- Yajima, H., Choi, J.-H., & Nagamine, K. 2012, *MNRAS*, 427, 2889