Mon. Not. R. Astron. Soc. **000**, 000–000 (0000) Printed 29 July 2013 (MN LATEX style file v2.2)

Constraints on the mass and occupation fraction of dark matter halos hosting Lyman-alpha emitters at $z \sim 3$

Jaime E. Forero-Romero[1] and Julian E. Mejía-Restrepo[2]

[1] *Departamento de Física, Universidad de los Andes, Cra. 1 No. 18A-10, Edificio Ip, Bogotá, Colombia*
[2] *Departamento de Astronomía, Universidad de Chile, Camino el Observatorio 1515, Santiago, Chile*

29 July 2013

ABSTRACT

We derive constraints on the mass and occupation fraction of dark matter halos hosting Lyα Emitting galaxies (LAEs) at a redshift of $z = 3.1$, by comparing the number density and the angular correlation function between mock and observed fields. The mock fields are constructed in a large cosmological N-body simulation using a model where a dark matter halo with mass in the range $M_{\rm min} < M_{\rm h} < M_{\rm max}$ can only host one LAE with a probability $f_{\rm occ}$. We define three different families of parameters consistent with observations based on their mass range $\Delta M \equiv \log_{10} M_{\rm max} - \log_{10} M_{\rm min}$ dex and occupation fraction $f_{\rm occ}$. From this we conclude that the number density and spatial clustering information available from observations is not sufficient to derive a unique and narrow set of values for the parameters in the model. In particular, it is always possible to find a a model with any desired occupation fraction in the range $0.1 \leq f_{\rm occ} \leq 1.0$. However, there is a dominant family composed by models with a narrow mass range $\Delta M < 1.0$ dex, a low occupation fraction $f_{\rm occ} \leq 0.3$ and a maximum mass $M_{\rm max} < 10^{12} h^{-1} {\rm M}_{\odot}$. The existence of this dominant family gives support to the idea that the most massive dark matter halos at that epoch do not host the brightest LAEs and that only a small fraction of star forming galaxies can be actually detected as LAEs.

Key words: cosmology: theory cosmology: large-scale structure of universe galaxies: formation galaxies: high-redshift galaxies: statistics galaxy: haloes

1 INTRODUCTION

Lyman-α emitting galaxies (LAEs) have become in the last decade a central topic in studies of structure formation in the Universe. They are helpful in a diverse range of fields. 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 through cosmic history (Forero-Romero et al. 2012). and tracers of active star formation (Guaita et al. 2013)

At the same time, theoretical and observational developments have contributed to the emergence of a paradigm to describe structure formation in a cosmological context. In this context it is considered that dominant matter content of the Universe is to be found in dark matter, whereby each galaxy is hosted by larger dark matter structure known as a halo. (Peebles 1980)

Most models of galaxy formation find that the mass of the halo can be used to predict properties of the galaxy such as its stellar mass and star formation rate (Behroozi et al. 2012). Processes that regulate the star formation cycle are also though to be strongly dependent on its mass. Furthermore, the spatial clustering of galaxies on large scales is entirely dictated by the halo distribution. For the reasons mentioned above, finding the typical dark matter halo mass hosting LAEs represents a significant step forward to understand the nature of this population in the context of Lambda Cold Dark Matter (ΛCDM) paradigm.

Some theoretical approaches to this problem have been based on ab-initio approach modeling. Starting from the DM halo population, the corresponding intrinsic star formation properties are inferred together with other statistics such as the luminosity function, the correlation function and the equivalent width distributions. Such modelling has been implemented from analytic considerations, semi-analytic models and full N-body hidrodynamical simulations (Laursen & Sommer-Larsen 2007; Dayal et al. 2009; Forero-Romero et al. 2011; Yajima et al. 2012; Orsi et al. 2012; Walker-Soler et al. 2012).

Added to the uncertainties in the astrophysical processes describing star formation in galactic populations, a highly debated step in this approach is the calculation of

the fraction of Lyα photons that escape the galaxy to the observer. Given the resonant nature of the Lymanα line, the radiative transfer of Lyα is sensitive to the density, temperature, topology and kinematics of the neutral Hydrogen in the interstellar medium (ISM) (Neufeld 1991; Forero-Romero et al. 2011; Dijkstra & Kramer 2012; Laursen et al. 2013; Orsi et al. 2012).

This complexity can be tackeld with monte-carlo simulations for the radiative transfer a required tool to obtain physically sound results, although the degeneracy in the physical parameters involved in the problem makes it difficult to achieve a robust consensus on what is the theoretical expected value for the Lyα escape fraction at high redshift.

Our objective in the present paper is to use a method to bypass all the complexity involved in the estimation of the Lyα escape fraction to infer information relevant to the understanding of LAEs in a cosmological context. The analysis we present is thus based on one of the most robust prediction of ΛCDM, namely the clustering and abudance of dark matter halos. We construct a simple model that allow us to construct mock catalogs and explore a wide range of parameters to narrow the mass interval of halos hosting LAEs together with its occupation fraction.

Another central point in our study is the explicit inclusion of cosmic variance effects by using large volume cosmological simulations. This allows us to go closer to the space of observables and directly explore the compatibility of a model with observations.

This paper is structured as follows. In the next section we present the simulation and model used to produce the mock catalogs and the criteria used to compare them against observations. In SS3 we present the main results in terms of the parameters of our model. We continue with a discussion of these results under the light of other observational results for LAEs and abundance matching models for galaxy formation before presenting our conclusions in SS5.

Throughout this paper we assume a ΛCDM cosmology with the following values for the cosmological parameters, $\Omega_m = 0.27$, $\Omega_\Lambda = 0.73$ and $h = 0.70$, corresponding to the matter density, vacuum density and the Hubble constant in units of 100 km s^{-1} Mpc^{-1}.

2 METHODOLOGY

Our method is based on the comparison of observatios and mock catalogs on the surface number density and the angular correlation function of LAEs. In the next subsections we describe in detail the four key elements of this workflow. First, we present the observations we take as a benchmark. Second, the N-body simulation and the halo catalogs we use. Third, the simplified model that allows us to translated halo catalogs into mock LAE observations. Fourth, the statistics we use to compare observational results against our theoretical predictions.

2.1 Observational Constraints

One of the main observational benchmarks we use in this paper is the LAE number density information at $z = 3.1$ obtained by the panoramic narrow-band survey presented by Yamada et al. (2012) from a survey conducted with the Subaru 8.2m telescope and the Subaru Prime Focus Camera, which has a field of view covering 34×27 arcmin, corresponding to a comoving scale of 46×35 Mpc h^{-1} at $z = 3.09$. The narrow band filter is centered at 4977 Åwith 77Åwidth, corresponding to the redshift range $z = 3.062 - 3.125$ and 41 Mpc h^{-1} comoving scale for the detection of the Lyman-α line centered at $z = 3.09$. The authors reported a total 2161 LAEs with an observed equivalent width larger than 190Åover a total survey area of 2.42 deg^2 that includes 12 subfields, this corresponds to average surface number density of 0.20 ± 0.01 arcmin^{-2}.

The survey covered four independent fields. The first is the SSA22 field of 1.38 deg^2 with 1394 detected LAEs (7 subfields), this field has been known to harbor a region with a large density excess of galaxies. The second observed region is composed by the fields Subaru/*XMM-Newton* Deep Survey (SXDS)-North, -Center and -South, with a total of 0.58 deg^2 and 386 LAEs (3 subfields). The third and fourth fields are the Subaru Deep Field (SDF) with 0.22 deg^2 and 196 LAEs (1 subfield), and the field around the Great Observatory Optical Deep Survey North (GOODS-N) with 0.24 deg^2 and 185 LAEs (1 subfield).

There is abundant observational work done on LAEs at redshift $z = 3.1$ (Kudritzki et al. 2000; Matsuda et al. 2005; Gawiser et al. 2007; Nilsson et al. 2007; Ouchi et al. 2008). However, we decide to focus on the data from Yamada et al. (2012) because it has the largest covered area with homogeneous instrumentation conditions (telescope, narrow band filter), data reduction pipeline and conditions to construct the LAE catalog. This ensures that the number density variations among fields are *not* due to different observational conditions or criteria to construct the catalogs.

The second constraint is the angular correlation function (ACF). Yamada et al. (2012) does not report an ACF measurement. Instead we use the results by Hayashino et al. (2004) and Ouchi et al. (2008, 2010).

Hayashino et al. (2004) observed in the densest field of SSA22 while Ouchi et al. (2008) observed on a 1 deg^2 sky of the Subaru/XMM-Newton Deep Survey (SXDS) Field. There are some differences between these observations and those iby Yamada et al. (2012). The details in the color selection, corresponding limiting luminosities and EW thresholds are different in these references. Nevertheless we use these observations as additional constraints in spite of the fact that the first selected models are based only on the surface density statistics by (Yamada et al. 2012).

2.2 Simulation and Halo Catalogs

The Bolshoi simulation (Klypin et al. 2011) we use in this paper was performed in a cubic volume of 250 h^{-1} Mpc on a side. The dark matter distribution is sampled using 2048^3 particles, which translates into a particle mass of $m_\mathrm{p} = 1.35 \times 10^8\ h^{-1}\ \mathrm{M}_\odot$. The cosmological parameters are consistent with a WMAP5 and WMAP7 data with a matter density $\Omega_\mathrm{m} = 0.27$, cosmological constant $\Omega_\Lambda = 0.73$, dimensionless Hubble constant $h = 0.70$, slope of the power spectrum $n = 0.95$ and normalization of the power spectrum$\sigma_8 = 0.82$ (Komatsu et al. 2009; Jarosik et al. 2011).

We use halo catalogs constructed with a Friend-of-Friends (FOF) algorithm with a linking length of 0.17 times

the inter-particle distance. The catalogs were obtained from the publicly available Multidark database [1] (Riebe et al. 2011). We focus our work on halos more massive than $1 \times 10^{10} h^{-1} \mathrm{M}_{\odot}$that are resolved with at least 70 particles, the reasons for this choice are explained in the next section.

2.3 A Model to Populate Halos with LAEs

In our model a dark matter halo can only host one LAE at most. There are three parameters in the model that decide whether a halo can host a galaxy or not: the lower and upper bounds for the mass range $M_{\rm min} < M_{\rm h} < M_{\rm max}$ where LAEs reside and the fraction $f_{\rm occ}$ of such halos that host a *detectable LAE*. The reader must keep in mind that the physical interpreation of the occupation fraction $f_{\rm occ}$ convolves two things: the presence of a star forming galaxy in a halo and its detectability as a LAE. We do not assign a luminosity to each LAE. We are primarily interested in constraining the halo mass range hosting detectable LAEs under the conditions defined by Yamada et al. (2012). In what follows we will describe by the letter $\mathcal{M}$ a model defined by an specific choice of the three scalar parameters $M_{\rm min}$, $M_{\rm max}$ and $f_{\rm occ}$.

For each model $\mathcal{M}$ we create a set of mock fields from disjoint volumes in the simulation. Each volume has the same geometry probed by Suprime-CAM and the narrow band filter, namely rectangular cuboids of dimensions $46 \times 35 \times 41\ h^{-3}\mathrm{Mpc}^3$ where the last dimension goes in the redshift direction. This corresponds to a total area of 880 arcmin^2 in each mock field. We construct a total $5 \times 7 \times 6 = 210$ of such volumes from a snapshot in the Bolshoi simulation. In each mock field a LAE is assigned to the position of a dark matter halo if the halo mass is in the range allowed by the model $M_{\rm min} < M_{\rm h} < M_{\rm max}$ and a random variable taken from an homogeneous distribution $0 \leq \xi < 1$ is smaller than the occupation fraction $\xi < f_{\rm occ}$.

Next we construct mock surveys by making groups of 11 mock fields out of the 210 available volumes. In total 15 mock surveys are constructed for each model $\mathcal{M}$. The grouping is done in two different ways. In the first way, called `match`, we follow the clustering of the observed fields. From the 11 mock fields, 7 are constructed from contiguous fields in the simulation to mimic the SSA22 region, 3 are also contiguous between them but not to the first 7 fields to mimic the SXDS fields and finally 2 non-contiguous fields to imitate the SDF and GOODS-North field. Our main goal with this selection is to test the impact on the final statistics of having 7 clustered fields. The second way to group the mock fields is called `random`, whereby all the 11 fields are selected in such a way as to avoid that any two volumes are contiguous. In this paper we only report the results obtained by the `match` method that are also obtained in the `random` selection.

2.4 Sampling and Selecting Models

We make a thorough exploration of the parameter space for the models $\mathcal{M}$. $\log_{10} M_{\rm min}$ takes 30 values from 10.0 up to 12.9 with an even spacing of 0.1 dex. $\log_{10} M_{\rm max}$ takes values in the same range as $\log_{10} M_{\rm min}$ only with a displacement of 0.1 dex in the whole range. The occupation fraction $f_{\rm occ}$ takes 10 different values from 0.1 to 1 regularly spaced by 0.1. In total the number of different models $\mathcal{M}$ that are explored is $30 \times 30 \times 10 = 9000$.

The lower limit for the parameter $M_{\rm min}$ is set by the minium occupation fraction we are able to consider. At $M_{\rm min} = 10^{10} h^{-1} \mathrm{M}_{\odot}$the halo number density is already ~ 10 times higher than the observational constraints for LAEs. This means that models in that mass range and an occupation fraction $f_{\rm occ}$ have the possibility to be compatible with observations. Lower values for $M_{\rm min}$ require $f_{\rm occ} < 0.1$, which are not considered in this paper .

For each mock survey generated in a given model $\mathcal{M}$ we compute the surface density in the 12 mock fields. We perform a Kolmogorov-Smirnov (KS) to compare this mock date against the 12 observational values. From this test we obtain a value $0 < P < 1$ to reject the null hypothesis, namely that two data sets come from the same distribution. In this paper we consider that for values $P > 0.05$ the two distributions can be thought as coming from the same distribution.

We begin by considering that a model $\mathcal{M}$ that has at least one mock survey (out of 15) consistent with the observed distribution of LAE number densities has viable parameters that deserve to be considered for further analysis. Later on we impose harder constraints to reduce the number of models by asking that all the 15 mocks to be consistent with observations.

We then calculate the angular correlation function (ACF) for all the models having the 15 mock surveys consistent with observations . The ACF is computed (using the Davis & Peebles estimator (Davis & Peebles 1983)) only on the densest subfield in all the 15 mock surveys corresponding to the SSA22 region. We then take the mean ACF over the 15 mock surveys for each model as well as its the standard deviation to include the effect of cosmic variance.

We finally compare the observational ACF and the mean mock ACF of the models in therms of the angular correlation lenght (θ_0) by fitting them to a power-law function:

$$\omega(\theta) = \left(\frac{\theta}{\theta_0}\right)^{-\beta} \tag{1}$$

Both mock and observational ACF are derived by a least square minimization procedure. The observational errors of the ACF as well as the standard deviation of the mean modelled ACFs are considered to compute the error in the fitted parameters to ensure ourselves to include poissionian and cosmic-variance errors. We consider that a model is consistent with observations if the two parameters β and θ_0 are equal within a 1-σ range.

3 RESULTS

The main purpose of this section is to show how different observational constraints narrow down the parameters space of allowed models. In each one of the four sub-sections we add a new constraint. First we consider the average halo

[1] `http://www.multidark.org/MultiDark/`

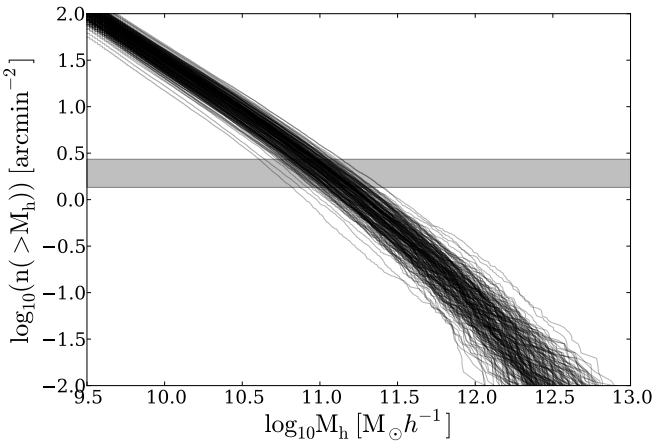

Figure 1. Surface density of dark matter halos as a function of a minimum halo mass to count the total number of elements in a volume. Each line represents one of the 210 volumes of dimensions $46 \times 35 \times 41$ h−3Mpc^3 in the Bolshoi simulation. The horizontal gray band represents the range of surface densities observed for LAEs at $z = 3.1$ as reported by (Yamada et al. 2012).

number density, then we consider the surface density distributions for all the mock fields in a given mock survey. From this result we move to narrower constraints by requiring a high number of mock surveys with consistent surface density distributions to finally consider the impact of the angular correlation function.

3.1 Dark Matter Halo Number Density

In Figure 1 we present the results for the integrated dark matter halo surface density as a function of halo mass. Each line corresponds to one of the 210 sub-volumes in the Bolshoi simulation. The gray band indicates the surface density values for LAEs allowed reported in observations (Yamada et al. 2012).

This result provides the basis to understand why only a specific range of models $\mathcal{M}$ can be expected to be consistent with observations. From Figure 1 we can read that models with a minimum mass $\log_{10} M_{\rm min} > 11.5 h^{-1}\mathrm{M}_\odot$ will always have a surface number density lower than the observational constrain. This makes impossible that models with that minimum mass can be compatible with observations.

The opposite is true in models with $\log_{10} M_{\rm min} < 10.5$ that will show surface number density larger than observations, this implies that in such models the occupation fraction has to be tuned $f_{\rm occ} < 1.0$ as to lower the halo number density to match the gray band values.

In the next subsection we quantify this intuition by means of the Kolmogorov-Smirnov tests between mock surveys and observations.

3.2 Models consistent with the surface density distributions

Figure 2 presents regions in parameter space $M_{\rm min} - M_{\rm max}$, $M_{\rm min} - f_{\rm occ}$ where the KS test yields values of $P > 0.05$ at least for one mock survey. For those models it is not possible reject the hypothesis that the simulated and observed data for the surface number density come from the same parent distribution.

The upper (lower) panels correspond to the `match` (`random`) method to build the mock surveys from individual fields. The plot shows number of mock surveys consistent with observations. There are between 550 to 600 models out of the original 9000 models that have at least one mock survey consistent with observations.

In Figure 2 there are three regions of parameter space that can be clearly distinguished. The first region corresponds to models where the minimum mass is high $\log_{10} M_{\rm min} > 11.5$. None of this models is compatible with observations as expected from the results in the previous subsection. For these models the number density of LAEs is too low compared to observations.

The second region corresponds to an intermediate range for the minimum mass $10.5 < \log_{10} M_{\rm min} < 11.5$ where regardless of the value of the maximum mass $M_{\rm max}$ it is possible to tune the occupation fraction $f_{\rm occ}$ to bring some of the mock observations into good agreement with observations. In this region in parameter space one can thus find two extreme kinds of models. One kind where the mass interval is very narrow with sizes smaller than < 0.3 dex (a factor of two in mass) and others where the mass interval is extended (larger than 1.0 dex) going up to the maximum halo mass present in the simulation at that redshift (up to 2.5 dex).

The third region in parameter space corresponds to $\log_{10} M_{\rm min} < 10.5$. In this case only models with a very narrow mass interval of at most 0.5 dex ($\log_{10} M_{\rm max} < 11.0$) and low occupation fractions $f_{\rm occ} < 0.3$ are allowed.

Without any additional information our method allows us to infer that most of the successful models are found in the second and third region of parameter space where this result was already expected from halo abundance calculations shown in Figure 1. In the next section we reduce the size of this region by imposing tighter constraints on the KS tests and consistency with the angular correlation function.

3.3 Models with the highest success rates

For each model $\mathcal{M}$ there are 15 different mock surveys. In the previous section we presented the models that had at least one mock survey with $P > 0.05$.

Figure 3 shows the number of models that have at least $N_{\rm high-P}$ mocks with $P > 0.05$ for both the `match` and `random` methods. This shows that there are between 80 to 100 models with all the 15 realizations with $P > 0.05$. Selection of these models as successful represents a reduction of a factor of ~ 6 with respect to the total number of mocks with at least one consistent mock.

Figure 4 presents the locii of these models in the parameter space $M_{\rm min} - M_{\rm max}$ and $M_{\rm min} - f_{\rm occ}$. The results are very similar between the `match` and `random` methods. With this constraints the number of consistent models with $\log_{10} M_{\rm min} < 11.5$ are greatly reduced. This corresponds to the regions in the parameter space in Figure 2 that already had a low number of consistent mock surveys. On the other hand, from the right panel in Figure 4 one can see that there is not a strong constraint on the occupation fraction $f_{\rm occ}$.

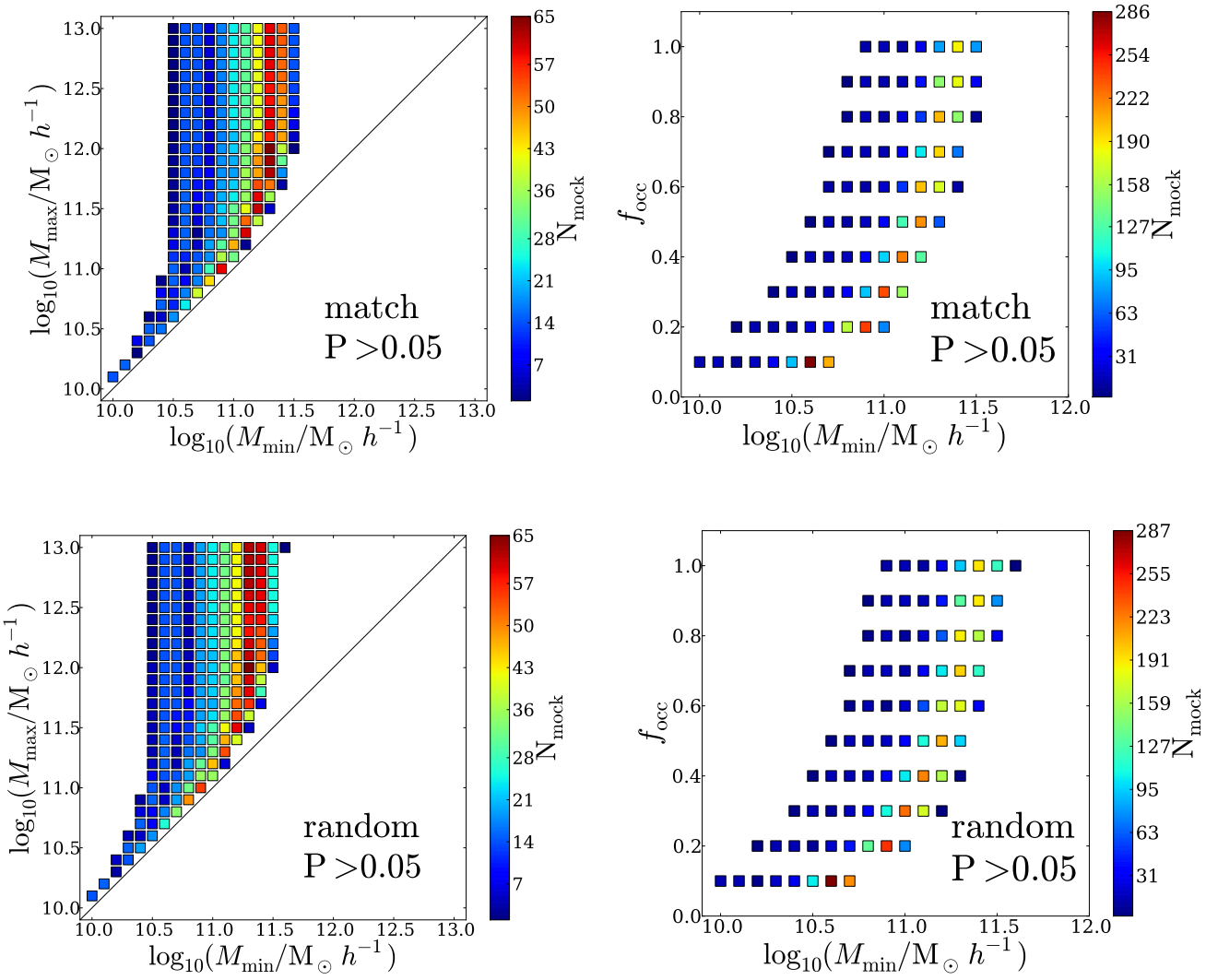

Figure 2. $M_{\rm min}$-$M_{\rm max}$ (left) and $M_{\rm min} - f_{\rm occ}$ (right) planes for all models with $P > 0.05$ in two different ways used to construct the mock surveys. The color code corresponds to the number of mock surveys that are found to be compatible with observations in terms of the KS test with $P > 0.05$. Only regions of parameter space with at least one consistent mock survey are included.

3.4 Consistency with the Angular Correlation Function

Figure 5 shows the main results in a $\theta_0 - \beta$ plane where the average and standard deviation for each mock is shown in comparison with the result derived from observations. The error bars in this Figure represents the standard deviation of the ACF over all the sub-fields in the 15 mock observations. These error bars are larger than the statistical uncertainty from the fitting procedure on a single field.

In the left panel Figure 5 we see how the observational ACF measured by Hayashino et al. (2004) (green dot) is successful in reducing the total number of possible models. Only those with angular-correlation length within $15'' < \theta_0 < 23''$ are considered to reproduce observations.

In order to better illustrate the results from this test we show in the right panel of Figure 5 the plane θ_0-$M_{\rm min}$ for all models. There we divide the models into two disjoint sets: those with $\log_{10} M_{\rm max} < 12$ (blue dots) and $\log_{10} M_{\rm max} > 12$ (green dots). The colored rectangle includes the parameter region which is consistent with the observational constraint in θ_0. With this restriction we find that most models with $\log_{10} M_{\rm max} > 12$ and $\log_{10} M_{\rm min} > 11.1$ can be safely ruled out.

In Figure 6 we present the preferred models in the planes $M_{\rm min} - M_{\rm max}$ and $M_{\rm min} - f_{\rm occ}$ for the `match` method after applying the observational constraints in the occupation fraction and the ACF. Here we consider a model consistent if there is an overlap of $1-\sigma$ in its values for the correlation length θ_0 and the power β in the correlation function. As before, we found similar results using the `random` method

4 DISCUSSION

Matching the galaxy surface number density statistics to our mock data sets the median mass of all successful models in the range $(10^{10} - 10^{12})\ h^{-1}{\rm M}_\odot$. Including more strict crite-

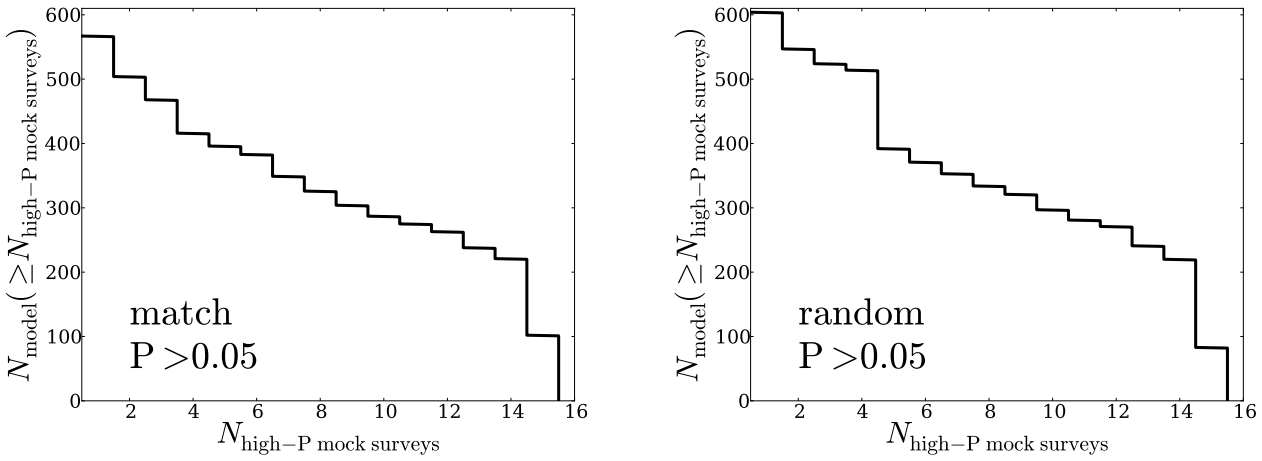

Figure 3. Number of models with a minimum number of mock survey realizations that are consistent with observations. .

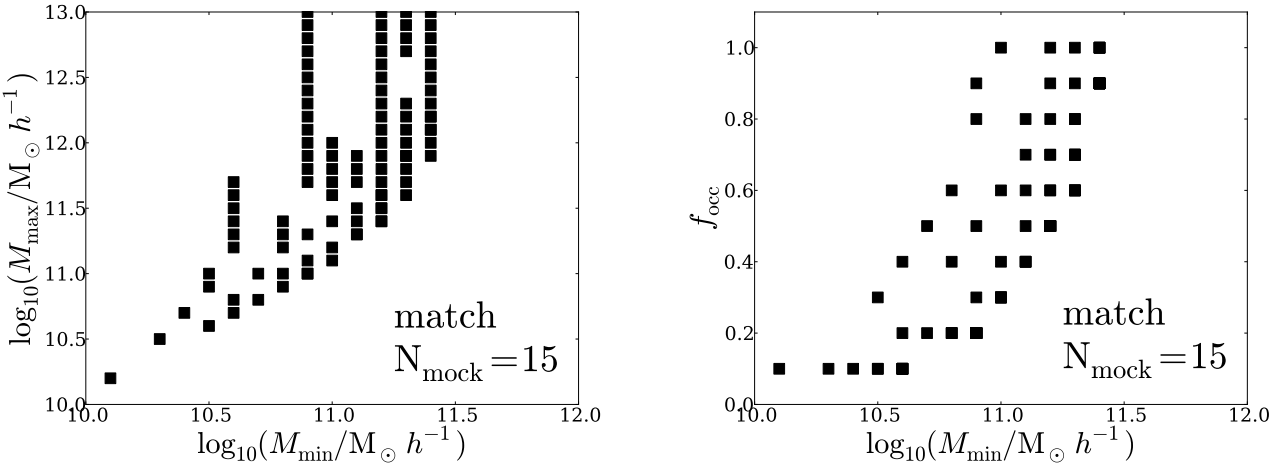

Figure 4. Favored regions in parameter space when the constraints on the maximal number of consistent mocks is imposed. The results for the `random` methodology (not shown here) are very similar to the ones presented here for the `match` method.

ria on the number of mock surveys that must be consistent with those observations and additional information from the angular correlation function greatly reduces the number of possible models. We end up with $\sim$ 50 models out of the initial 9000 possible combination of parameters.

We find that to make a physical interpretation of these models it is useful to do it in terms of the size of halo mass range. We now use the variable $\Delta M = \log_{10} M_{\rm max} - \log_{10} M_{\rm min}$, together with the occupation fraction $f_{\rm occ}$ to help us build a classification of all the successful models into three families:

(1) Low $f_{\rm occ} \leq 0.3$ and low $\Delta M \leq 1.0$ dex: 24 models.
(2) Low $f_{\rm occ} \leq 0.3$ and high $\Delta M > 1.0$dex: 11 models
(3) High $f_{\rm occ} > 0.3$ and low $\Delta M \leq 1.0$: 14 models

There is a clear majority of models with a narrow $\Delta M \leq 1$, compared to the 2.5dex of halo mass available for occupation at that redshift. Such models imply that there is a cut at lower and higher halo masses that render inefficient the presence and/or detection of LAEs.

At the low mass end, such cut can be readily interpreted in terms of the minimal halo star formation rate needed to produce the necessary Lymanα luminosity to be above a given detection threshold. However, under the reasonable assumption of star formation rate increasing with halo mass, the cut at higher halo masses requires a different explanation. There are two different physical scenarios to explain this.

One possible interpretation can be made in terms of a decreasing escape fraction of Lyα radiation in massive systems. There are detailed models for radiative transfer that support the idea that massive galaxies with higher metallicities have larger dust contents and less concentrated ISM than lower mass systems which, due to the resonant nature of the Lyα line that increase the probability of absortion ofLyα photons, produce high absorption of Lyα photons but

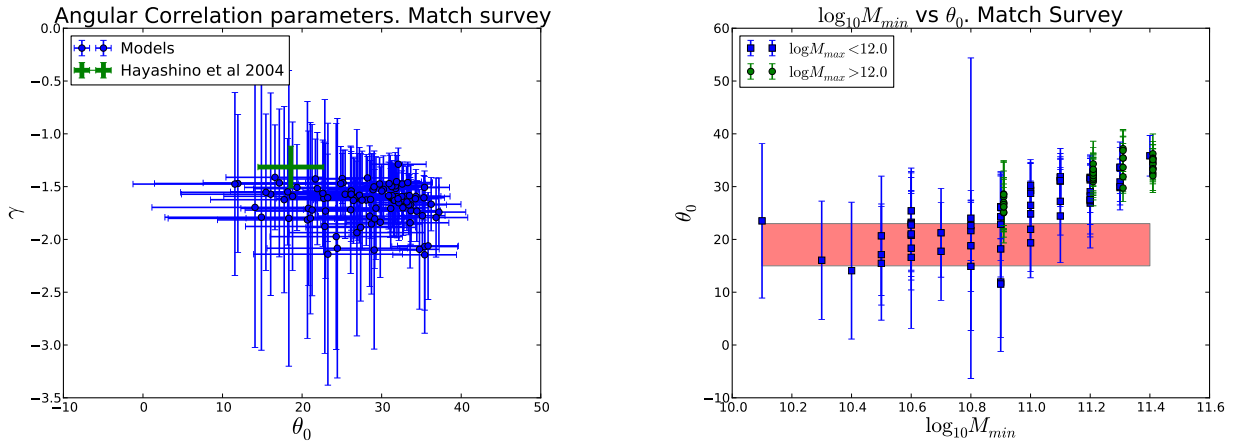

Figure 5. Left: Values for the free parameters (θ_0 vs β) in the fitting formula (Eq. 1) for the mean angular correlation function. Blue dots correspond to simulations and the green cross to observations by Hayashino et al. (2004). The error bars in the theoretical data correspond to the standard deviation from the different mocks surveys. Right: Same results as before in the shown in the θ_0-$M_{\rm min}$ plane. This time the observational constraints are represented by the red rectangle. Blue squares represent models with $\log_{10} M_{\rm max} < 12.0$, green circles correspond to models with $\log_{10} M_{\rm max} > 12.0$.

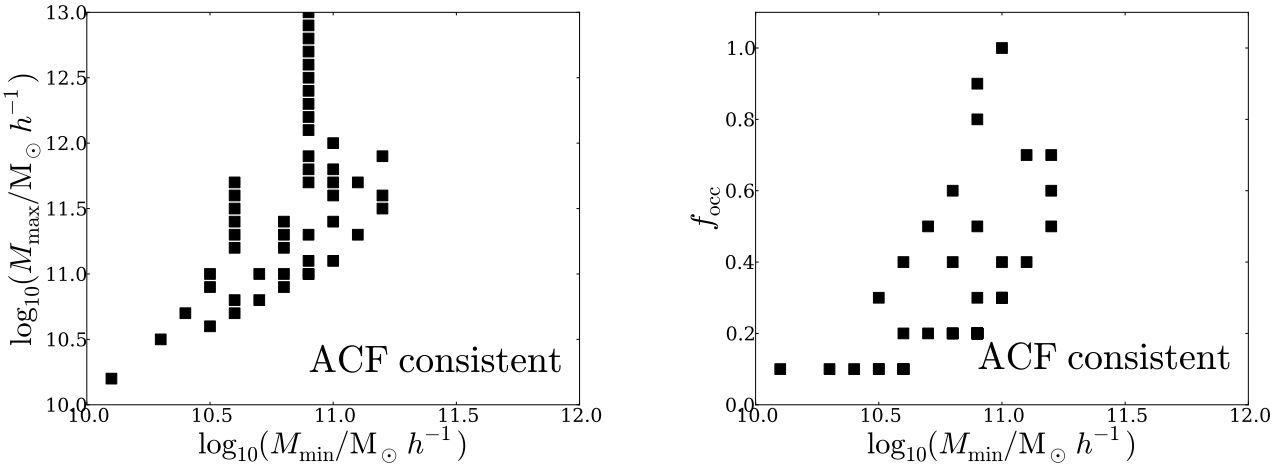

Figure 6. Best models parameter space when both the constraints on the maximal number of consistent mocks and the occupation fraction $f_{\rm occ} \leq 0.2$ are included. $M_{\rm min}$ vs $M_{\rm max}$ (left), $M_{\rm min}$ vs $f_{\rm occ}$ (rigth).

not of continuum or other n-resonant lines (Forero-Romero et al. 2011; Laursen et al. 2009).

A second scenario is that larger systems have more extended gaseous envelopes which, due to resonance effects of the Lyα line, leading to a low surface brightness and a broader line, making these systems less observable in narrow band filter surveys (Zheng et al. 2010).

The preference for narrow ΔM ranges to hosting LAEs together with the minor presence of reasonable models with $M_{\rm max} > 10^{12} h^{-1}{\rm M}_\odot$ shows that our models support theoretical and observational insights where the most massive systems are not bright Lyα sources (Forero-Romero et al. 2012), (Shapley et al. 2003)

4.1 Comparison against results from blind surveys

We also find that 70% of the best models are found in families (1) and (2), with a low occupation fraction $f_{\rm occ} \leq 0.3$. This preference goes in the same direction as the observational constraint on the escape fraction $f_{\rm esc} \sim 0.1 - 0.2$ derived at $z = 2.2$ by Hayes et al. (2010) and recent theoretical study of observational data in a wide redshift range $0 < z < 6$ (Dijkstra & Jeeson-Daniel 2013).

The observational estimation by Hayes et al. (2010) was based on blind surveys of the Hα and Lyman α line. Using corrections by extinction to obtain an estimate for the intrinsic Hα luminosity, and using values for the theoretical expectation of the ratio Lymanα/Hα they derive an bulk escape fraction for the Lymanα radiation of $f_{\rm esc} = (5.3 \pm 3.8)\%$ or $f_{\rm esc} = (10.7 \pm 2.8)\%$ if a different dust correction is used.

They also showed that that the luminosity function for

LAEs at $z = 2.2$ is consistent with the escape fraction being constant for every galaxy regardless of its luminosity. From this results they derive that almost 90% of the star forming galaxies emit insufficient Lyman α to be detected, effectively setting the occupation fraction to be $f_{\rm occ} = 0.10$.

Dijkstra & Jeeson-Daniel (2013) used a similar principle to derive their results. They compared observationally derived star formation functions to LAE luminosity functions. At $z \sim 3.0$ they derive an effective escape fraction of $f_{\rm esc} = (17 \pm 5)\%$ could be interpreted as an occupation fraction $f_{\rm occ} \sim 0.2$. We consider a success of our method the fact that we find that most of the consistent models show a low occupation fraction.

4.2 Comparison to other clustering estimates

Observational based on the ACF inferred from photometric measurements in the Extended Chandra Deep Field South have shown that the median dark matter masses of h los hosting LAEs is $\log_{10} M_{\rm med} = 10.9^{+0.5}_{-0.9} {\rm M}_\odot$, with a corresponding occupation fraction of $1 - 10\%$ (Gawiser et al. 2007). Our results are in a general good agreement with those estimates for the host mass. This is not completely unexpected given that we have also required consistency with ACF measurements.

The novelty in our approach is that we have a detailed estimate for host halo mass range together with the escape fraction. This allows us to demonstrate that the halo mass range could be very narrow $\Delta M < 0.2$dex, something that cannot be inferred from ACF analysis alone.

We also find interesting that our ACF analysis is also not enough to rule out models with a high occupation fraction $f_{\rm occ} > 0.3$, which represent almost one fourth of our best models, coinciding with a wide range in halo masses $\Delta M > 1.0$ dex. These models can only be considered unfavorable based on the density and abundance observations as we have described in the previous subsection.

Ouchi et al. (2010) presents an analysis of LAE observations in the redshift interval $3.1 < z < 7.0$. At $z = 3.1$ They get to an average mass for the host dark matter halos of $M_h = 2.9^{+24.0}_{-2.9} \times 10^{10}\ h^{-1}{\rm M}_\odot$with a corresponding duty cycle of 0.008 ± 0.03. This broadly matches the expectations from the first family of models, also summarized in Table 1. In particular it seems to favor only the very low occupations fractions.

4.3 In the context of abundance matching models

Considering the additional evidence for a low escape fraction and assuming a direct relation between f_{esc} and f_{occ} we can say that the preferred models are in families (1) and (2) with a clear majority composed by those with narrow mass range.

Abundance matching methods throughout cosmic time from redshifts $0 < z < 8$ Behroozi et al. (2013b,a) work based on observational results for Lyman Break Galaxies. In those studies the minimum halo mass considered to be relevant in their analysis is $10^{11.4} h^{-1}{\rm M}_\odot$. For that mass, Behroozi et al. (2013b) report a stellar mass mass around $(1.0 \pm 0.3) \times 10^{9.0}\ h^{-1}{\rm M}_\odot$, while their star formation rate is in the range 0.6 ± 0.2 Msun yr^{-1}, which nevertheless are close to the lower bound of values inferred for LAEs at high redshift (Gawiser et al. 2007; Nilsson et al. 2009; Pentericci et al. 2009).

In our results, all the preferred models have a halo mass range lower than the minimum of $M_{\rm min} < 10^{11.4} h^{-1}{\rm M}_\odot$considerend in abundance matching at $z = 3$. This mass scale is also strictly superior to the great majority of $M_{\rm max}$ values allowed in the models with low escape fraction. This suggests that a detailed and careful study of the spectral and photometric properties of LAEs coupled to the kind of analysis performed in this paper can be a guide in the study of the properties of low mass dark matter halos at $z = 3.1$, extending the capabilities of abundance matching methods.

4.4 On the reproducibility of our results

... All the software to produce the results in this paper is publicly available.

... The raw catalogs can be obtained from the MultiDark database but can also be obtained in the repository of this paper on github.

5 CONCLUSIONS

In this paper we constrain the preferred mass and occupation fraction of dark matter halos hosting Lyman Alpha Emitters at redshift $z = 3.1$ in a ΛCDM cosmology. We use a method that matches the cosmic variance of the surface density number of LAEs between mock and real observations and is also consistent with estimations of the angular correlation function. The mock catalogs are constructed using a model with three basic parameters: the halo mass range where LAEs can be found, $M_{\rm min} < M_{\rm h} < M_{\rm max}$, and the fraction of the halos in this range that host a detectable LAE, $f_{\rm occ}$. After a thorough exploration of the parameter space we find that a wide variety of models are consistent with the observational constraints, where only a minority has been considered so far in the literature. Out of 9000 initial combinations for the model parameters we end up with 49 successful arrangement of parameters.

The most important conclusion of this work is that available observational information for the clustering properties of LAEs at $z = 3.1$ does not provide enough constraints to uniquely narrow the range of masses and occupation fraction of dark matter halos hosting those galaxies. In particular we are not able to put a stringent constraint on the occupation fraction $f_{\rm occ}$.

In our results the models can be split into three families depending the mass range $\Delta M = \log_{10} M_{\rm max} - \log_{10} M_{\rm min}$ and the occupation fraction $f_{\rm occ}$. The first family is narrow both in ΔM and $f_{\rm occ}$, a second family is only narrow in ΔM and the third is narrow in $f_{\rm occ}$ and broad in ΔM.

All the halo mass ranges in the best models fall into the broad range obtained in previous analysis (Gawiser et al. 2007). The improvement of our method is being able to describe in high detail all the possible models consistent with observations, including the discovery of models previously overlooked like those with a high occupation fraction $f_{\rm occ} > 0.3$.

A central result in this exploration of parameter espace is the existence of a large number of models with a very

$\log_{10} M_{\rm min}$	$\log_{10} M_{\rm max}$	$f_{\rm occ}$
10.1	10.2	0.1
10.3	10.5	0.1
10.4	10.7	0.1
10.5	10.6	0.3
10.5	10.9	0.1
10.5	11.0	0.1
10.6	10.8	0.2
10.6	11.2	0.1
10.6	11.3	0.1
10.6	11.4	0.1
10.6	11.5	0.1
10.6	11.6	0.1
10.7	11.0	0.2
10.8	11.2	0.2
10.8	11.3	0.2
10.8	11.4	0.2
10.9	11.3	0.3
10.9	11.7	0.2
10.9	11.8	0.2
10.9	11.9	0.2
11.0	11.6	0.3
11.0	11.7	0.3
11.0	11.8	0.3
11.0	12.0	0.3

Table 1. List of parameters for the first family of models. Narrow mass range $\Delta M \leq 1.0$dex and low occupation fraction $f_{\rm occ} \leq 0.3$.

narrow mass range $\Delta M < 0.1$ dex that are consistent in every aspect with the spatial distribution of observed LAEs. In order to have these characteristics it is required that halos around the mass $M_{\rm max}$ for each model become inefficient in hosting detectable galaxies. One way to achieve this is having LAEs with a decreasing Lyα escape fraction with increasing mass.

All the models also present a range of masses where the halos are below the mass scale of $10^{11.4} h^{-1} {\rm M}_{\odot}$ inferred as a lower threshold for the LBGs used in halo abundance matching studies. Detailed study of high-z LAES will have a positive impact on our understanding of the connection of low mass halos to star forming galaxies.

We foresee that the new observations with new instruments (such as MUSE, Hyper SuprimeCam and HETDEX) covering larger fields and a wider range of luminosities will be key in imposing tighter constraints on the properties of dark matter halos hosting LAEs.

ACKNOWLEDGMENTS

J.E.F-R thanks the hospitality of Changbom Park and the Korea Institute for Advanced Study where the first full draft of this paper was completed. J.E.F-R also thanks Peter Laursen for helpful comments on the physical interpretation and the presentation of our results.

REFERENCES

Behroozi P. S., Wechsler R. H., Conroy C., 2012, ArXiv e-prints

$\log_{10} M_{\rm min}$	$\log_{10} M_{\rm max}$	$f_{\rm occ}$
10.6	10.7	0.4
10.7	10.8	0.5
10.8	10.9	0.6
10.8	11.0	0.4
10.9	11.0	0.8
10.9	11.0	0.9
10.9	11.1	0.5
11.0	11.1	1.0
11.0	11.4	0.4
11.1	11.3	0.7
11.1	11.7	0.4
11.2	11.5	0.7
11.2	11.6	0.6
11.2	11.9	0.5

Table 2. List of parameters for the second family of models. Narrow mass range $\Delta M \leq 1.0$dex and high occupation fraction $f_{\rm occ} > 0.3$.

$\log_{10} M_{\rm min}$	$\log_{10} M_{\rm max}$	$f_{\rm occ}$
10.6	11.7	0.1
10.9	12.1	0.2
10.9	12.2	0.2
10.9	12.3	0.2
10.9	12.4	0.2
10.9	12.5	0.2
10.9	12.6	0.2
10.9	12.7	0.2
10.9	12.8	0.2
10.9	12.9	0.2
10.9	13.	0 0.2

Table 3. List of parameters for the third family of models. Broad mass range $\Delta M > 1.0$dex and low occupation fraction $f_{\rm occ} \leq 0.3$.

Behroozi P. S., Wechsler R. H., Conroy C., 2013a, ApJL, 762, L31
Behroozi P. S., Wechsler R. H., Conroy C., 2013b, ApJ, 770, 57
Davis M., Peebles P. J. E., 1983, ApJ, 267, 465
Dayal P., Ferrara A., Saro A., Salvaterra R., Borgani S., Tornatore L., 2009, MNRAS, 400, 2000
Dijkstra M., Jeeson-Daniel A., 2013, ArXiv e-prints
Dijkstra M., Kramer R., 2012, MNRAS, 424, 1672
Dijkstra M., Mesinger A., Wyithe J. S. B., 2011, MNRAS, 414, 2139
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
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., 2007, 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., 2007, 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
Hayes M., Östlin G., Schaerer D., Mas-Hesse J. M., Leitherer C., Atek H., Kunth D., Verhamme A., de Barros S., Melinder J., 2010, Nature, 464, 562
Jarosik N., Bennett C. L., Dunkley J., Gold B., Greason M. R., Halpern M., Hill R. S., Hinshaw G., Kogut A., Komatsu E., Larson D., Limon M., 2011, ApJS, 192, 14
Klypin A. A., Trujillo-Gomez S., Primack J., 2011, ApJ, 740, 102
Koehler R. S., Schuecker P., Gebhardt K., 2007, A&A, 462, 7
Komatsu E., Dunkley J., Nolta M. R., Bennett C. L., Gold B., Hinshaw G., Jarosik N., Larson D., Limon M., Page L., Spergel D. N., Halpern M., 2009, ApJS, 180, 330
Kudritzki R.-P., Méndez R. H., Feldmeier J. J., Ciardullo R., Jacoby G. H., Freeman K. C., Arnaboldi M., Capaccioli M., Gerhard O., Ford H. C., 2000, ApJ, 536, 19
Laursen P., Duval F., Östlin G., 2013, ApJ, 766, 124
Laursen P., Razoumov A. O., Sommer-Larsen J., 2009, ApJ, 696, 853
Laursen P., Sommer-Larsen J., 2007, ApJL, 657, L69
Matsuda Y., Yamada T., Hayashino T., Tamura H., Yamauchi R., Murayama T., Nagao T., Ohta K., Okamura S., Ouchi M., Shimasaku K., Shioya Y., Taniguchi Y., 2005, ApJL, 634, L125
Neufeld D. A., 1991, ApJL, 370, L85
Nilsson K. K., Møller 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
Nilsson K. K., Tapken C., Møller P., Freudling W., Fynbo J. P. U., Meisenheimer K., Laursen P., Östlin G., 2009, A&A, 498, 13
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
Pentericci L., Grazian A., Fontana A., Castellano M., Giallongo E., Salimbeni S., Santini P., 2009, A&A, 494, 553
Riebe K., Partl A. M., Enke H., Forero-Romero J., Gottloeber S., Klypin A., Lemson G., Prada F., Primack J. R., Steinmetz M., Turchaninov V., 2011, ArXiv e-prints
Shapley A. E., Steidel C. C., Pettini M., Adelberger K. L., 2003, ApJ, 588, 65
Walker-Soler J. P., Gawiser E., Bond N. A., Padilla N., Francke H., 2012, ApJ, 752, 160
Yajima H., Choi J.-H., Nagamine K., 2012, MNRAS, 427, 2889
Yamada T., Nakamura Y., Matsuda Y., Hayashino T., Yamauchi R., Morimoto N., Kousai K., Umemura M., 2012, AJ, 143, 79
Zheng Z., Cen R., Trac H., Miralda-Escudé J., 2010, ApJ, 716, 574