Mon. Not. R. Astron. Soc. **000**, 000–000 (0000) Printed 21 June 2013 (MN LATEX style file v2.2)

Using cosmic variance to constrain the dark matter halo mass of Lyman-alpha emitting galaxies at z=3.1

Jaime E. Forero-Romero[1] and Julian 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*

21 June 2013

ABSTRACT

We use cosmological N-body dark matter only simulations to constrain the characteristic mass of dark matter halos hosting Lyman-Alpha Emitting (LAE) galaxies at a redshift of $z = 3.1$. The method is based on matching the statistics for the number density between mock and observed fields. The mock fields are constructed using a simple model where a dark matter halo can only host one LAE with a probability $f_{\rm occ}$ if its mass is found withing a certain range mass range delimited by two threshold values, $M_{\rm min}$ and $M_{\rm max}$. We also maximize the number of mocks surveys consistent with observations and impose consistency with the angular correlation function. Under these conditions we find that LAEs are preferentially hosted by halos in very narrow mass ranges less than 0.5 dex in width, with the minimum mass $10 \leq \log_{10} M_{\rm min} \leq 10.9$ and the occupation fraction $f_{\rm occ} \leq 0.2$. Our finding suggest that the most massive dark matter halos at that epoch do not host the brightest LAEs. This also gives support to observational evidence that says that only a small fraction of star forming galaxies can be actually detected as LAEs.

Key words: galaxies: kinematics and dynamics, Local Group, methods:numerical

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 and markers of the the galaxy formation process through cosmic history (Forero-Romero et al. 2012).

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.

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 a forward modeling. Starting from the DM halo population, the corresponding intrinsic star formation properties are infered and 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 (Dayal et al. 2009; Forero-Romero et al. 2011; Yajima et al. 2012; Forero-Romero et al. 2012) .

Added to the uncertainties in the astrophysical processeses describing star formation in galactic populations, a highly debated steps in this approach is the calculation of the fraction of Lyman-α photons that escape the galaxy to the observer. Given the resonance nature of the line, the radiative transfer of Lyman-α 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; Laursen et al. 2013).

This complexity makes the use of 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 ex-

pected value for the Lyman-*alpha* escape fraction in high redshift.

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 to constrain the typical mass of a dark matter halos hosting LAEs at $z = 3.1$ is based on the comparison of observational results on the surface number density and the predictions of a simple model that uses the outputs from cosmological N-body simulations.

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

The observational benchamark 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 report a total 2161 LAEs with an observed equivalent width larger than 190 over a total survey area of 2.42 deg^2, this corresponds to averaga 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, 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. The third and fourth fields are the Subaru Deep Field (SDF) with 0.22 deg^2 and 196 LAEs, and the fild arotund the Great Observatory Optical Deep Survey North (GOODS-N) with 0.24 deg^2 and 185 LAEs.

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 intrumentational 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 due only to astrophysical reasons and not different observational conditions or criteria to construct the catalogs.

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. It includes 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 density $\Omega_\mathrm{m} = 0.27$, cosmological $\Omega_\Lambda = 0.73$, 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 lenght of 0.17 times the interparticle distance. The minimum halo mass in the models we construct in this paper correspond to groups of ~ 75 particles. 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.

2.3 A Model to Populate Halos with LAEs

In our model a dark matter halo can only host one or zero LAE. 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 $M_\mathrm{min} < M_\mathrm{h} < M_\mathrm{max}$ where LAEs reside and the fraction f_occ of such halos that effectively host a LAE. We do not assign a luminosity to each LAE. We are primaraly interested in constraining the halo mass range hosting detectable LAEs under the conditions defined by Yamada et al. (2012). In what follows will describe by the letter $\mathcal{M}$ a model defined by an specific choice of the three scalar parameters M_min, M_max y f_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 dinemsion 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 $M_\mathrm{min} < M_\mathrm{h} < M_\mathrm{max}$ and a random variable taken from an homogeneous distribution $0 \leq \xi < 1$ is smaller than the occupation fraction $\xi < f_\mathrm{occ}$.

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 contigous 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-contigous 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

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

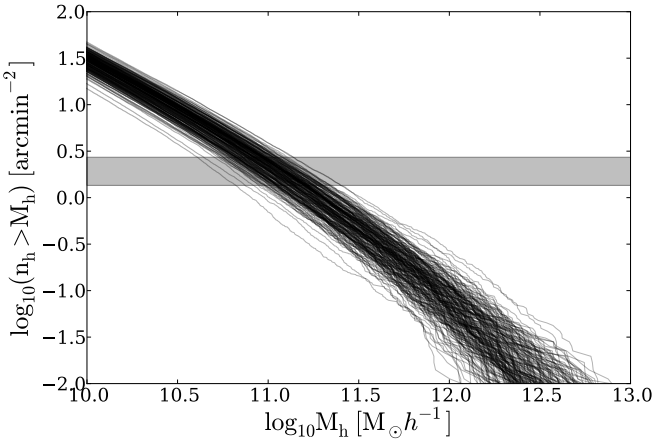

Figure 1. Surface density of dark matter halos as a function of a minimum halo mass to count the total number of elements in a volume. Each line represents on of the 210 volumes of dimensions $46 \times 35 \times 41$ h−3Mpc^3 in the Bolshoi simulation. The horizontal grey band represents the range of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

mock fields is called `random`, whereby all the 11 fields are selected in such a way as to avoid that any two volumes are contigous.

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 regulary spaced by 0.1. In total the number of different models $\mathcal{M}$ that are explored is $30 \times 30 \times 10 = 9000$.

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 hypotesis, 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.

In this paper we consider that a model $\mathcal{M}$ that has at least one (1) mock survey (out of 15) consistent with the observed distribution of LAE number densities has viable parameters thatdeserve to be considered for further analysis.

3 BASIC RESULTS

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 a range of models $\mathcal{M}$ can be expected to be consistent with observations. In Figure 1 we can read that models with a minimum mass $\log_{10} M_{\rm min} > 11.5 h^{-1}{\rm M}_\odot$will always have a surface number density lower than the observational constrain. The oppossite 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 numeber density to match the gray band values.

Conversely, there are regions in the plot where the halo surface density is always higher than the observational constraints correspond to models $\mathcal{M}$ with a minimum mass below $M_{\rm min} < 3 \times 10^{10} h^{-1}{\rm M}_\odot$. Models with this minimum mass hava a chance for successfuly reproducing observations if the occupation fraction $f_{\rm occ} < 1$ is tuned as to lower the halo number density down to the observed value.

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

3.2 Three Regions in Parameter Space

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 (1) mock survey consistent with observations.

In Figure **??** there are three regions of parameter space that can be clearly distinguised. 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 section. For these models the number density of LAEs is too low.

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 very extended, larger than 1.0 dex, going up to the maximum halo mass present in the simulation at that redshift.

The third region in parameter space corresponds to $\log_{10} M_{\rm min} < 10.5$. In this case only models with a 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 succesful 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 using more stringent constraints to define the

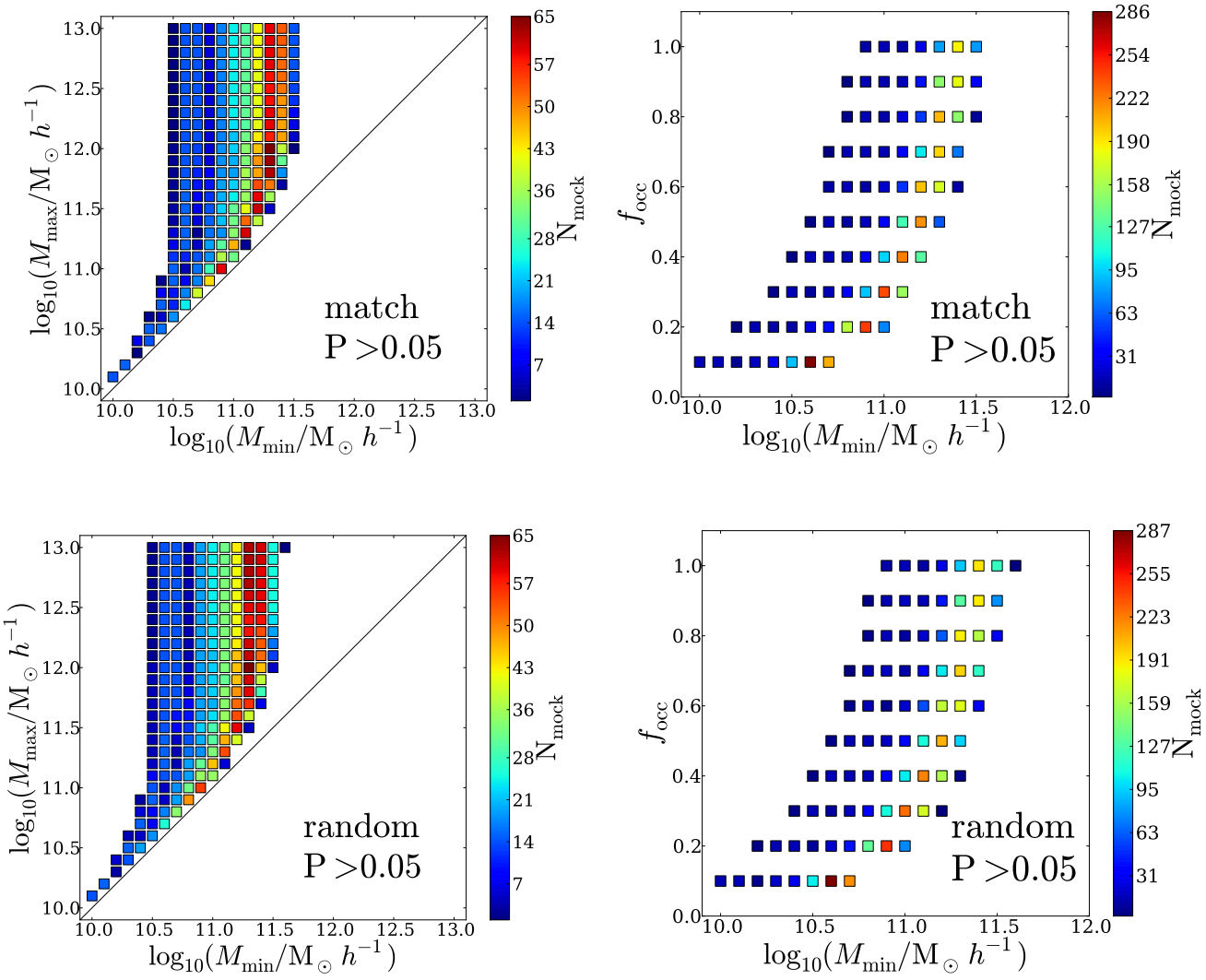

Figure 2. $M_{\rm min}$-$M_{\rm max}$ (left) and $M_{\rm min} - f_{\rm occ}$ (right) planes for all models with $P > 0.05$ in two different ways used to construct the mock surveys. The color code corresponds to the number of mock surveys that are found to be compatible with observations in terms of the KS test with $P > 0.05$. Only regions of parameter space with at least one (1) consistent mock survey are included.

agreement with observations and including additional observational information onpossible values for the occupation fraction at $z \sim 3$.

4 ADDITIONAL CONSTRAINTS

In this section we consider two additional constraints on the models we accept as succesul. First we take models with the highest possible number of mock surveys consistent with observations. Second by comparing the angular correlation function against observational constraints by Hu et al. (2004).

4.1 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 (1) 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 $P > 0.05$. Selction of these models as successful represents a reduction of a factor of ~ 6 with repect 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 $\log_{10} M_{\rm min} < 11.7$ 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 selection effect on the occupation fraction $f_{\rm occ}$.

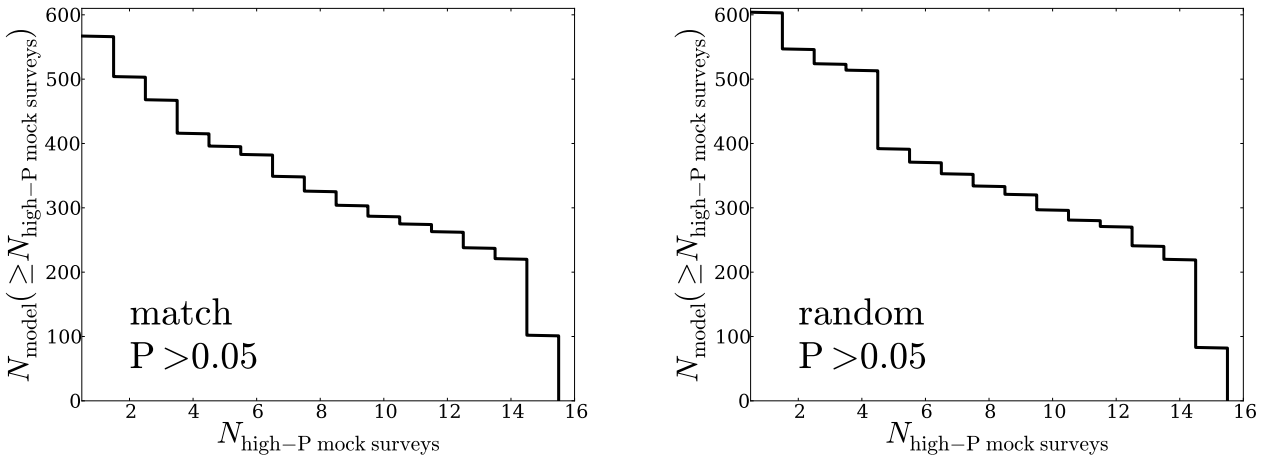

Figure 3. Number of models with a mininium number of mock survey realizations that are consistent with observations. .

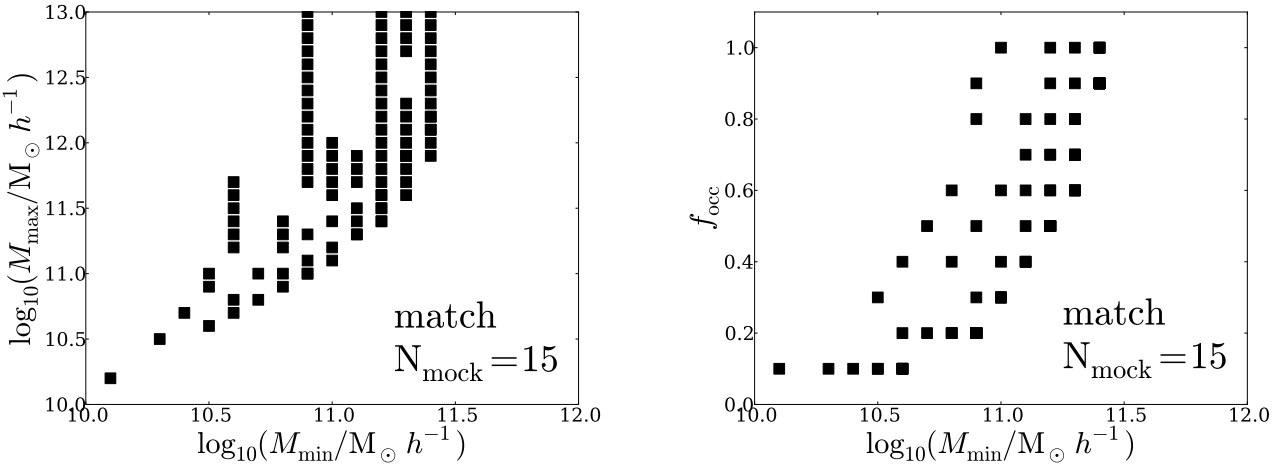

Figure 4. Favored regions in parameter space when the constraints on the maximal number of consistent mocks is imposed. The results for the `random` methodology (not shown here) are very similar to the ones presented here for the `match` method.

4.2 Consitency with the Angular Correlation Function

As an additional observational constraint we requiring that the models selected in the previous section must be compatible with the observational angular correlation function (ACF). The measurementens presented in (Yamada et al. 2012) do not report an ACF measurement. Instead we use the results reported by (Hu et al. 2004) on the LAEs observed in the densest field of SSA22.

It is important to keep in mind that there are some differences between this work and (Yamada et al. 2012). First, the color selection by Yamada et al. (2012) is less stringent compared to the one by Hu et al. (2004). Also the EW threshold is different, Hu et al. (2004) uses a cut around 154Åinstead of 190Åused by (Yamada et al. 2012)

The comparison between the theoretical and observational ACF is done in therms of the angular correlation lenght obtained by fitting to a power-law function:

$$\xi(\theta) = \left(\frac{r}{\theta_0}\right)^{-\gamma} \quad (1)$$

Both the theoretical and observational ACF are derived by a least square minimization procedure. The results are shown in Figure 5 in a $\theta_0 - \gamma$ plane where the average and standard deviation for each mock is shown in comparison with the resut derived from observations. The Error bars in these figures 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..

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

In order to interpret the results from this test we show

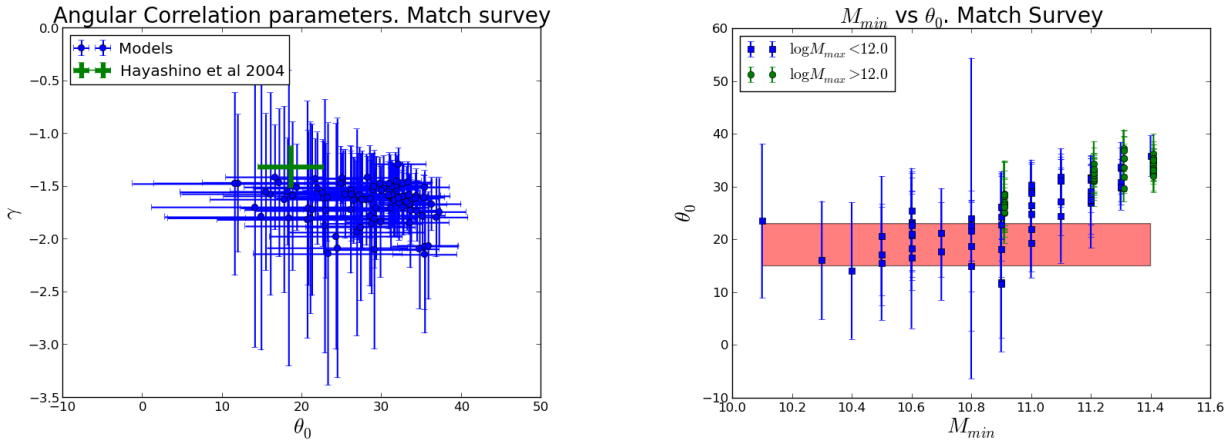

Figure 5. Left: Values for the free parameters (θ_0 vs γ) of the mean angular correlation function of the 15 mock surveys for each model. Blue dots corresponds to simulations and the green cross to observations by Hu et al. (2004). The error bars in the theoretical data correspond to the starndard deviation from the different mocks. Right: Same results as before 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 are models with $\log_{10} M_{\rm max} > 12.0$.

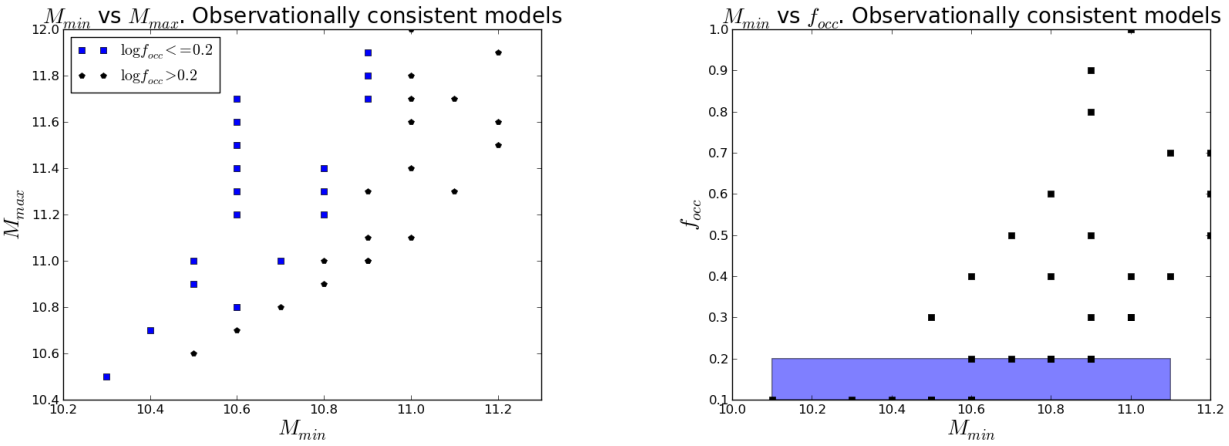

Figure 6. Best models paramter space when both the constraints on the maximal number of consistent mocks and the occupation fraction $f_{\rm occ} \leq 0.2$ are included. $M_{\rm min}$ vs $M_{\rm max}$ (left), $M_{\rm min}$ vs $f_{\rm occ}$ (rigth).

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 sents: those with $\log \frac{M_{\rm max}}{M_\odot} < 12$ (blue dots) and $\log \frac{M_{\rm max}}{M_\odot} > 12$ (green dots). The colored rectangle includes the parameter region wich is consistent with the observational constraint in θ_0. With this restriction we find that most models with $\log \frac{M_{\rm max}}{M_\odot} > 12$ and $\log \frac{M_{\rm min}}{M_\odot} > 11.1$can be safely ruled out.

In Figure 6 we present the prefered 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 γ in the correlation function.

5 DISCUSSION

Imposing additional constraints with the maximal number of consistent mock surveys and the angular correlation function we end up with 50 models out of the initial 90000 models. In what follows we discuss the implications of the selection of these models on the escape fraction of Lyman α radiation, the stellar abundance and the star formation efficiency in LAEs at $z = 3.1$.

5.1 Escape fraction and allowed halo mass range

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

The observational estimation by **?** 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 $f_{\rm occ} = 0.10$.

The results by Dijkstra & Jeeson-Daniel (2013) use a similar principle. They compared observationally dereived star formation functions to LAE luminosity functions. At $z \sim 3.0$ they derive an affective escape fraction of $f_{\rm esc} = (17 \pm 5)\%$, which in turn using the same arguments in Hayes et al. (2010) could be interpreted as an occupation fraction $f_{\rm occ} \sim 0.2$. We consider a sucess of our method the fact that we find that most of the consistent models show a low occupation fraction.

The study of the escape farction is also complementary to the size of the halo mass range $\Delta M = \log_{10} M_{\rm max} - \log M_{\rm min}$for LAEs. All the successful model we have found can be split into three families.

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

There is a clear majority of models with a narrow $\Delta M \leq 1$, such models imply that there is a cut for lower and higher halo masses that render inefficient the presence and/or detection of LAEs. A cut a lower masses can be readily interepreted in terms of the minimal halo star formation rate needed to produce the necessary Lymanα luminosity to be above a given detection threshold.

The cut at higher halo masses requires a different explanation. A possible interpretation can be made in terms of a lower escape fraction of Lyαradiation in massive systems. There are detailed models for radiative transfer that support the idea that massive galaxies with higher metallicities have larger dust contents than lower mass systems, which due to the resonant nature of the Lyαline are enough to produce high absorption of Lyαphotons but not of continuum or other non-resonant lines.

This preference for narrow ΔM ranges for hosting LAEs together with the total absence of reasonable models with $M_{\rm max} > 10^{12} h^{-1}{\rm M}_{\odot}$shows that our models supports theoretical insights where the most massive systems are not bright Lyαsources.

5.2 In the context of abundance matching models

Using abundance matching methods throughout cosmic time from redshifts $0 < z < 8$ Behroozi et al. (2013b,a) report that the instantaneous star formation efficiency (star formation rate divided by the stellar mass) presents a clear maximum around $10^{11.7}{\rm M}_{\odot}$at all redshifts $z < 4$.

This mass scale is strictly superior to the all the $M_{\rm max}$ values allowed in the consistent models. Using the results published in Behroozi et al. (2013b) we find that the typical stellar mass in halos of $10^{11.4} h^{-1}{\rm M}_{\odot}$is $(1.0 \pm 0.3) \times 10^{9.0}$ $h^{-1}{\rm M}_{\odot}$, while their star formion rate is in the range 0.6 ± 0.2 Msun yr^{-1}. These halo masses actually correspond to the lower bound of what is computed in the abundance matching model in order to fit the observational data for the stellar mass function.

We have the opposite situation in the second familiy of models. If we have a wide range in halo masses, where the upper end of the halo masses can be considered as observed LAEs, one can expect that bright LBGs will have a correspondence as observed LAEs. The most interesting aspect is that there is a clear cut in the minimal mass that can be attained by observed LAEs $M_{\rm min} > 10^{11} h^{-1}{\rm M}_{\odot}$. This puts a tight constraint on the relationship between the minimum star formation rate required to be observed as a LAE and this minimal halo mass.

... Intrinsic emission and escape fraction.
... Star formation rate efficiency at this redshift.
... Mass dependence of the escape fraction.
... Discuss all this in terms of the star formation efficiency in Behroozi et al.

5.3 Implications for large LAEs surveys

... The bias for the preferred halo mass.
... The scale at which cosmic variance drops.
... This can be observationally tested with HETDEX.

5.4 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.

6 CONCLUSIONS

In this paper we constrain the preferred mass for dark matter halos hosting Lyman Alpha Emitters at a redshift $z = 3.1$. We use a method that matches the cosmic variance in the surface density number of LAEs between mock and real observations. The mock catalogs are based on a simplified 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 thisrange that are actully occupied, $f_{\rm occ}$. After a thorugh exploration of the parameter space we are able to constrain the mass range of dark matter halos hosting LAEs to be in the range $< M_{\rm h} <$ and a corresponding occupation fraction that escales as $f_{\rm occ} = M_{\rm min}$.

We use three additional constraints to reduce the allowed range of models. The first imposes a tighter criterion to consider a model succesful, namely that all the mock surveys for a given model must be consistent with observations. This restriction narrows down the allowed range of models to be.

The second constraint is based on the observational results that high redshift LAEs have a bulk Lyman alpha escape fraction of XX which can be also interepreted as an average occupation fraction of XX.

Including additional observational constraints on the

$\log_{10} M_{\rm min}$	$\log_{10} M_{\rm max}$	$f_{\rm occ}$
10.1	10.2	0.1
10.3	10.5	0.1
10.4	10.7	0.1
10.5	10.9	0.1
10.5	11.0	0.1
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
10.6	11.7	0.1
10.6	10.8	0.2
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	11.7	0.2
10.9	11.8	0.2
10.9	11.9	0.2

Table 1. List of model parameters for the best models that have all mock surveys consistent with observations and an occupation fraction $f_{\rm occ} \leq 0.2$. This corresponds to the `match` method to construct the mock surveys. .

occupation fraction allows us to reduce the range of allowed halo masses to be in a narrower range of $< M_{\rm h} <$. Including the information from the angular correlation function (ACF) does not allows us to impose further constraints. This is due to the scatter in the ACF due to the cosmic variance on the field observed by XXX

We simulation allows us to extract 210 sub-boxes each of which has a comparable volume to the individual fields of view observed by Yamada et al. (2012). The comparison of the observed number density distribution against the results from our model is based on three different ways of constructing mock surveys. The first reproduces the spatial correlation between the 12 observational fields (`match`), the second breaks this spatial correlation while keeping the number of fields (`random`) and the third one simply includes all the 210 sub-boxes (`full`). We find that the methods `match` and `random` allow a larger set of models than the `random` method. We do not find a significant difference between the two first methods.

ACKNOWLEDGMENTS

REFERENCES

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

Behroozi P. S., Wechsler R. H., Conroy C., 2013a, ApJL, 762, L31

Behroozi P. S., Wechsler R. H., Conroy C., 2013b, ApJ, 770, 57

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., 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

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

Hu E. M., Cowie L. L., Capak P., McMahon R. G., Hayashino T., Komiyama Y., 2004, AJ, 127, 563

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

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

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

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

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