

# THE CONNECTION BETWEEN STAR FORMATION RATE AND DARK MATTER HALO MASS IN THE EPOCH OF REIONIZATION

FELIPE L. GOMEZ-CORTES, JAIME E. FORERO-ROMERO

Departamento de Física, Universidad de los Andes, Cra. 1 No. 18A-10, Edificio Ip, Bogotá, Colombia

Submitted for publication in *ApJ*

## ABSTRACT

We present updated constraints on the relationship between the star formation rate and dark matter halo mass at redshift  $z \sim 6$ . The observational basis for our work is the restframe UV luminosity function data obtained with CFHTLS, HST Legacy Survey and UKIDSS. The constraints are based on an abundance matching methodology to the observational data using cosmological N-body simulations. We also take into account the influence on the results of the dust extinction scaling derived from observations by Bouwens. We show dependence of our parameter uncertainties related to the comoving volume and magnitude range in observed surveys, taking advantage of the wide mass resolution range of latest available dark matter halo catalogs. We compare our results against the results of abundance matching methods (to the stellar mass), a semi-analytic model of galaxy formation (GALFORM) and a hydrodynamical simulation (Illustris).

*Subject headings:* galaxies: high-redshift — methods: numerical

## 1. INTRODUCTION

All magnitudes are in AB system.

## 2. OBSERVATIONAL CONSTRAINTS

We use information compiled in four different publications. All of them select galaxy candidates at  $z \sim 6$  using the drop-out technique (Steidel et al. 1996). In what follows we describe the relevant details of each reference.

### 2.1. Bouwens et al. 2015

Bouwens et al. (2015) presented results from a compilation of observations taken with the Advanced Camera for Surveys (ACS) and near-infrared Wide Field Camera 3 (WFC3/IR) since 2002 through 2012. The study includes the following fields of view. XDF, HUDF09-1, HUDF09-2, CANDELS-S/Deep, CANDELS-S/Wide, ERS, CANDELS-N/Deep, CANDELS-N/Wide, CANDELS-UDS, CANDELS-COSMOS and CANDELS-EGS.

The total survey area is  $740.8 \text{ arcmin}^2$  over five different lines of sight, with a total estimated volume of  $1.8 \times 10^6 \text{ Mpc}^3$  comoving. The limiting magnitude ranges between  $\sim 27.5 \text{ mag}$  in CANDELS-EGS and  $\sim 30 \text{ mag}$  in the deepest field (XDF). The total number of  $z = 6$  LBG candidates is 940, most of them in the faint end of the LF given the relatively small survey volume. The restframe UV magnitudes for these candidates are in the range.  $-22.52 \leq M_{1600} \leq -16.77$ . The estimated Schechter parameters:  $\phi^* = (0.33^{+0.15}_{-0.10}) \times 10^{-3} \text{ Mpc}^{-3}$ ,  $M_{1600}^* = -21.16 \pm 0.20$  and  $\alpha = -1.91 \pm 0.09$ .

### 2.2. Finkelstein et al. 2014

Finkelstein et al. (2014) worked also with HST data. They used results from the HUDF, CANDELS and GOODS fields, along with two of the Hubble Frontier Fields (HFF) of deep parallel observations (unlensed fields) near the Abell 2744 and MACS J0416.1-2403 clusters. The HFF uses the ACS and the WFC3/IR with the

same filters aforementioned except  $z_{850}$ . The total survey area is around  $\sim 300 \text{ arcmin}^2$  with a total estimated volume of  $8 \times 10^5 \text{ Mpc}^{-3}$ . There are 706 photometric candidates at redshift 6 defined as the interval  $5.5 < z < 6.5$ . The Schechter function parameters estimated for this data set are  $\phi^* = (1.86^{+0.94}_{-0.80}) \times 10^{-4} \text{ Mpc}^{-3}$ ,  $M_{1600}^* = -21.1^{+0.25}_{-0.31}$  and  $\alpha = -2.02^{+0.10}_{-0.10}$ .

### 2.3. Willott et al. 2013

Willott et al. (2013) presented results from the sixth release of the Canada-France-Hawaii Telescope Legacy Survey (CFHTLS). The observations were performed over four separated fields covering a total area  $\sim 4 \text{ deg}^2$  giving a survey volume of  $\sim 1 \times 10^7 \text{ Mpc}^3$ , which is over one order of magnitude larger than the compilations by Bouwens et al. (2015) and Finkelstein et al. (2014).

They performed optical observations with MegaCam. Their main selection criteria was that all objects must be brighter than magnitude  $z' = 25.3$ . The final number of LBGs founded was 40. Moreover, they get spectroscopic confirmation for 7 candidates using GMOS spectrograph on the Gemini Telescopes, which has a  $\ll 5.5$ -square arcmin field of view.

The survey was focused on the most luminous LBGs. The full LF at  $z = 6$  cannot be obtained as in other studies, however its large volume allows constraints on the bright end. The LF Schechter parameters are calculated in the magnitude range  $-20.5 > M_{1350} > -22.5$ .

### 2.4. McLure et al. 2009

McLure et al. (2009) used data obtained with the United Kingdom Infrared Telescope (UKIRT) in the near-IR imaging and Subaru Telescope for the optical imaging. They covered  $2268 \text{ arcmin}^2$  for a volume  $\sim 3 \times 10^6 \text{ Mpc}^3$ . The UKIRT was equipped with the WFCAM using *JK* filters, while Subaru was equipped with the Suprime-Cam with the *BVRi'z'* filters. All candidates were brighter than  $z' = 26$ . The UV rest frame magnitude range is  $-22.4 \leq M_{1500} \leq -20.6$ . The

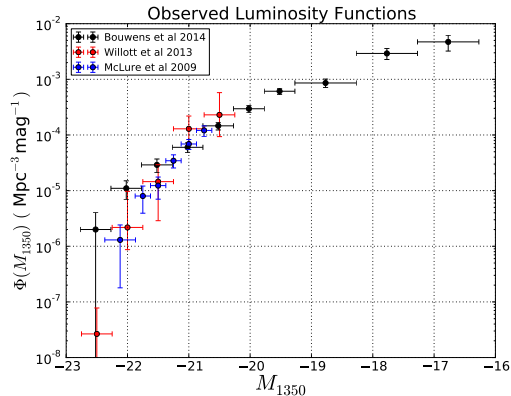


FIG. 1.— Observational data from Bouwens et al. (2015); McLure et al. (2009) and Willott et al. (2013).

LF was calculated using the maximum likelihood estimator of Schmidt (1968). Their analysis gave a total number of 104 LBG candidates in the redshift range  $5.7 \leq z \leq 6.3$ . They reported the following values for the Schechter function  $\phi^* = (1.8 \pm 0.5) \times 10^{-3} \text{Mpc}^{-3}$ ,  $M_{1500}^* = -20.04 \pm 0.12$  and  $\alpha = -1.71 \pm 0.11$ .

### 2.5. Dust Attenuation

Dust present in these high redshift star forming galaxies can attenuate the UV intrinsic luminosity. Although relatively uncertain compared to studies in the local universe, there are constraints on the extinction level in LBGs at  $z = 6$ .

The UV Spectral Slope  $\beta$  was introduced by Meurer et al. (1999) as a UV color to study dust attenuation in local starburst galaxies and extrapolating to high redshift galaxies. This index appears when a power law fitting is performed over the spectral flux  $f$  as function of the wavelength  $\lambda$ ;

$$f \propto \lambda^\beta.$$

The relation for ultra-violet attenuation at 1600Å they found is:

$$A_{1600} = 4.43 + 1.99\beta, \quad (1)$$

with  $A_{1600}$  in magnitude units.

Due LBGs have more similar spectra properties to local starburst galaxies rather than AGNs for example, we can assume that local calibration of  $\beta - A_{1600}$  can be applied also for LBGs, “The main requirement is that the data include fluxes in two broad bands or coarse spectra covering the rest-frame UV.”

Bouwens et al. (2012) uses the fluxes on different bands to estimate  $\beta$  on each LBG candidate found with  $z \sim 4-7$ . After in redshift groups they found a linear relation between  $\beta$  and the UV magnitude:

$$\langle \beta \rangle = \frac{d\beta}{dM_{UV}} (M_{UV,AB} + 19.5) + \beta_{M_{UV}=-19.5} \quad (2)$$

with  $\beta_{M_{UV}=-19.5} = -2.20$  and  $d\beta/dM_{UV} = -0.21$  at  $z = 5.9$ .

Smit et al. (2012) used the aforementioned relations to infer the corrected luminosity functions, i.e. dust-free luminosity functions, and the corrected SFR at  $z = 4$ .

Here we use the inverse relation, starting from the intrinsic or dust-free galaxy magnitude, obtaining the ob-

served magnitude:

$$M_{obs} = \begin{cases} \frac{M_{int} - 4.616}{1.259}, & \text{if } A > 0 \\ M_{int}, & \text{else} \end{cases}. \quad (3)$$

## 3. THEORETICAL AND NUMERICAL FRAMEWORK

### 3.1. The Abundance Matching Methodology

We use an abundance matching approach to find the relationship between dark matter halo mass and star formation rate. This approach has been used to link stellar and dark matter masses. Here we extend it to constraint star formation rate properties at high redshift.

The starting point for this method is a population of dark matter halos. To each halo we assign a UV luminosity according to the following parameterization.

$$L_{UV}(M) = L_0 M \left[ \left( \frac{M}{M_0} \right)^{-\beta} + \left( \frac{M}{M_0} \right)^\gamma \right]^{-1}, \quad (4)$$

where  $M$  is the DM mass,  $L_0$  is a normalization constant,  $M_0$  is the critical mass where the luminosity function has a slope change,  $\beta$  and  $\gamma$  are two slopes corresponding to the faint and bright end, respectively. We now that this equation has the same functional dependence suggested by Moster et al. (2010) to link stellar and dark matter masses.

Once each halo has a UV luminosity we also have the option to include an extinction correction in order to modify this intrinsic luminosity value. From these UV values we build the LF. The free parameters in Eq. (4) are determined by requiring that the abundance matching LF follows the observational constraints. Then we use the following relationship between UV luminosity and Star Formation Rate (Madau et al. 1998; Kennicutt 1998)

$$\text{SFR} (M_\odot \text{yr}^{-1}) = 1.4 \times 10^{-28} L_\nu (\text{erg s}^{-1} \text{Hz}^{-1}), \quad (5)$$

to finally link SFR with DM halo mass.

### 3.2. N-body Simulations and Halo Catalogs

Cosmological N-body simulations are the source of the dark matter halo populations. We use two different simulations to cover the wide dynamical range explored by the observations: Big MultiDark Planck (MDPL) and Planck Bolshoi (P-Bolshoi). Both simulations use 2013 Planck cosmology defined by the following parameters:  $\Omega_M = 0.307$ ,  $\Omega_B = 0.048$ ,  $\Omega_\Lambda = 0.730$ ,  $\sigma_8 = 0.829$ ,  $n_s = 0.96$  and  $H_0 = 67.8$ .

The MDPL run is a N-body dark matter only simulation based on the L-Gadget2 code. The simulated volume is a cubic box of  $1 \text{Gpc h}^{-1}$  on a side. It has  $3840^3$  dark matter particles mass of  $1.51 \times 9 M_\odot h^{-1}$ . The DMH Catalog at  $z = 6$  contains  $\sim 10.9 \times 10^7$  halos, to avoid incompleteness in the low mass end, halos with mass below  $10^{10.3} M_\odot h^{-1}$  are rejected. We split this large volume into 64 smaller cubic boxes of  $250 h^{-1} \text{Mpc}$  on a side (similar volume to the observations by Willott et al. (2013)) to study the influence of cosmic variance.

The P-Bolshoi simulation...

We obtained the data from the public database<sup>1</sup> (Riebe et al. 2013).

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

### 3.3. Fitting Model

The key element to connect SFR with halo mass from simulations is the observed UVLF. If there exist a function whom gives to each halo a UV luminosity then is posible to reproduce UVLF from DMH catalogs. The fitting parameters can be explored using a Markov Chain Monte-Carlo (MCMC) implementation. Once the parameters are found, the SFR-Halo Mass relation have been found.

### 3.4. MCMC

We used a Markov Chain Monte Carlo method to find the best parameters and its uncertainties over each one of the boxes and each observational dataset. The code was written on Python using the SciPy Library (Jones et al. 2001–) the IPython environment (Pérez & Granger 2007).

First of all, one of the four observational data sets (Bouwens, Finkelstein, Willott and McLure) is selected to be the model to fit. From the whole simulated halo catalog, a subsample with cubic box shape of  $250\text{Mpc h}^{-1}$  length is selected.

With an initial set of parameters ( $\alpha, \gamma, M_0$ , and  $L_0$ ) the UV luminosity for the halo is calculated as function of his mass according to equation 4. Then Luminosity is converted to magnitude units using:

$$M_{UV} = 51.82 - 2.5 \log_{10}(L_{UV}).$$

If we consider the dust attenuation in the equation 3 we have a dust-corrected UV Luminosity Function.

The luminosity function is constructed as an histogram of the magnitudes normalized by the volume of the catalog. Each observed luminosity function has a different bin range.

Once having the LF, we compare our LF against observed LF. The error function we consider is the sum of the square difference over each bin, divided by the observational data uncertain.

$$\chi^2 = \sum_{i=0}^n \frac{(x_{i,obs} - x_{i,fit})^2}{2\sigma_i^2}$$

We worked on the logarithmic space of luminosities to have a good fitting on six decades.

The likelihood will have this property:

$$\mathcal{L} \propto \exp\left(-\frac{\chi^2}{n}\right)$$

, where  $n$  is the number of degrees of freedom. We have the maximum likelihood when the error is the minimum.

Each MCMC step will have a small variation of the parameters, the UV luminosity and magnitude are calculated again for each halo, we have a new error  $\chi_{new}^2$  and a new likelihood  $\mathcal{L}_{new}$ .

Following the Metropolis method, we compare the new and the old likelihood:

$$R = \frac{\mathcal{L}_{new}}{\mathcal{L}_{old}} = \exp(\chi_{new}^2 - \chi_{old}^2)$$

If  $R \geq 1$ , then we immediate accept the new set of parameters and start the next MCMC step.

Else, we have a chance to keep the new set. When  $R < 1$ , we compare with a uniformly random number

$p$  in the range  $[0, 1]$ . if  $R > p$  we accept the new set of parameters and start the next MCMC step. Else we reject the new set and start again with the old set.

We performed 100.000 effective MCMC steps, plus 10.000 thermalization steps over each box. We repeated for the same box without consider dust attenuation.

Then we perform the same method over the 64 boxes and the resting three data sets.

The restrictions imposed over the parameters where  $0 \leq \alpha \leq 2.0$  and  $\gamma \geq 0$ .

Finally, the UV luminosity for each galaxy can be directly related with SFR according to Madau et al. (1998). This model is accurate within the range of  $10^8 - 10^9\text{yr}$  (Kennicutt 1998). The relation between UV luminosity and Star Formation Rate is given by:

$$\text{SFR} (M_{\odot}\text{yr}^{-1}) = 1.4 \times 10^{-28} L_{\nu} (\text{erg s}^{-1}\text{Hz}^{-1}) \quad (6)$$

### 3.5. The Luminosity Model

In this model we have made two assumptions:

1. Each halo in the catalog hosts one galaxy. There are not empty halos, also none of halos has two or more galaxies.
2. The UV luminosity of each galaxy is function of one variable: the mass of the DMH in which is located.

Our model is a four parameter function. Each galaxy has a luminosity given by:

$$L = L_0 M \left[ \left( \frac{M}{M_0} \right)^{-\beta} + \left( \frac{M}{M_0} \right)^{\gamma} \right]^{-1} \quad (7)$$

where  $M$  is the hosting DMH mass,  $L_0$  is a normalization constant,  $M_0$  is the critical mass where the luminosity function has a slope change,  $\beta$  and  $\gamma$  are the slopes. This equation has a similar fashion to the mass to light relation (van den Bosch et al. 2003) and the mean relation between stellar mass of a galaxy and the mass of its halo used by Moster et al. (2010).

There exists more sophisticated models as Lee et al. (2009) shown, with scattering in the DM Mass-Stellar Mass, but also including some randomization in star forming histories, starburst episodes (duty cycle); galaxies have not synchronization on the beginning of star forming stage, also this stage may be time limited. It results in the existence of massive galaxies in the studied volume, but with no emission in the UV continuum due their duty cycle may has not started as well it may ended. Also may be present a normal distribution of the luminosity around the expected values.

## 4. RESULTS

### 4.1. Willott

We performed MCMC runnings with 100.000 steps over the 64 small cubic boxes (of  $250 h^{-1}\text{Mpc}$  side length) fitting the LF to the Willott observational data, comparing the two cases; with and without dust attenuation. (figures 3 and 4). We used the Likelihood Ratio criterion ( $\mathcal{LR} = 0.5$ ) to define the  $1\sigma$  confidence interval for our parameters.

$M_0$  is quite similar in both cases (they are compatible within the error bars), the turnover point corresponds to

Parameter	Dust Att.		No-Dust Att.	
	MCMC	C.Var.	MCMC	C.Var.
$\log_{10}(L_0/L_\odot)$	$18.07^{+0.10}_{-0.06}$	0.03	$17.81^{+0.12}_{-0.08}$	0.02
$\log_{10}(M_0/M_\odot)$	$11.19^{+0.67}_{-0.03}$	0.07	$11.12^{+0.58}_{-0.24}$	0.06
$\beta$	$1.480^{+0.11}_{-1.28}$	0.18	$1.44^{+0.16}_{-1.35}$	0.18
$\gamma \times 0.1$	$4.028^{+3.01}_{-0.82}$	0.58	$5.18^{+2.36}_{-0.59}$	0.46

TABLE 1  
BEST FIT PARAMETERS TO THE WILLOTT DATA WITH AND WITHOUT DUST ATTENUATION OVER 64 SMALL BOXES. MEAN BEST VALUE ESTIMATED WITH MCMC, MEAN  $1\sigma$  CONFIDENCE INTERVAL AND COSMIC VARIANCE (C.VAR.)

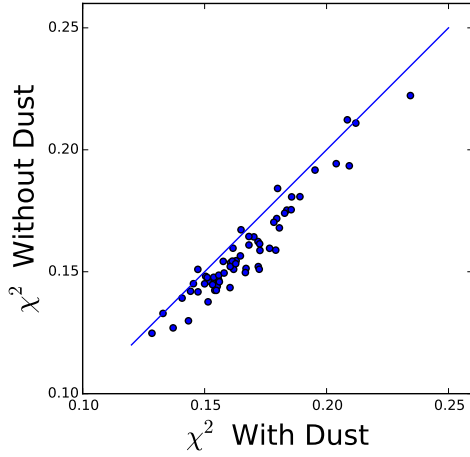


FIG. 2.— Best fit to Willott comparison between the two models. Each point represent  $\chi^2$  calculated over each small-box. The solid line represents the ratio 1:1

the same mass.  $\gamma$  and  $L_0$  shown a significative difference in booth cases.  $\beta$  is hard to constraint in booth cases. The paramerer was limited to vary in the range form 0.0 to 1.6.  $1\sigma$  region covers the whole range.

The UV luminosity model (eqn. 4) that we have chosen can be divided in two regimes; high mass regime (with  $M > M_0$ ) and low mass regime (with  $M < M_0$ ).

The observational dataset from Willott are in the high mass regime with one point in the low mass regime. It makes makes hard to impose restrictions over  $\beta$ , but the other three parameters can be well defined.

We also compare the likelihood of the two cases on each individual small box. The Dust Attenuation model is more accurate than the No-Dust Attenuation model in most of the cases as is shown in the figure 2.

To study cosmic variance effects, we compared the best fit parameters of each box and its likelihood value. We found that cosmic variance effects are less significative in best fit parameters than MCMC parameter estimation itself.

## 5. DISCUSSION

## 6. CONCLUSIONS

## ACKNOWLEDGMENTS

Acknowledgments...

The observational datasets were retrieved using

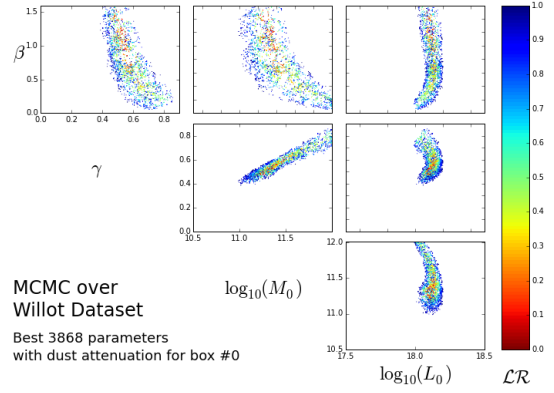


FIG. 3.— Parameter dispersion fitting the Willott with the Dust Attenuation model.  $1\sigma$  is defined by the likelihood ratio between 0.0 (red) and 0.5 (green)

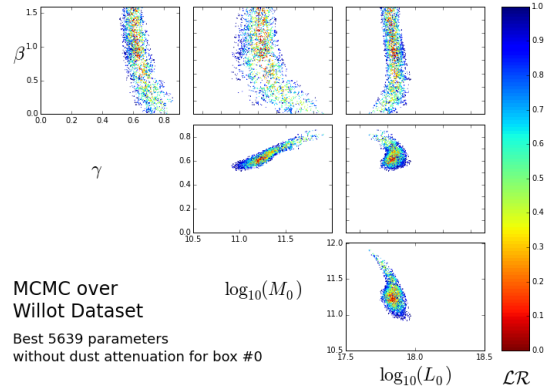


FIG. 4.— Parameter dispersion fitting the Willott with the No-Dust Attenuation model.  $1\sigma$  is defined by the likelihood ratio between 0.0 (red) and 0.5 (green)

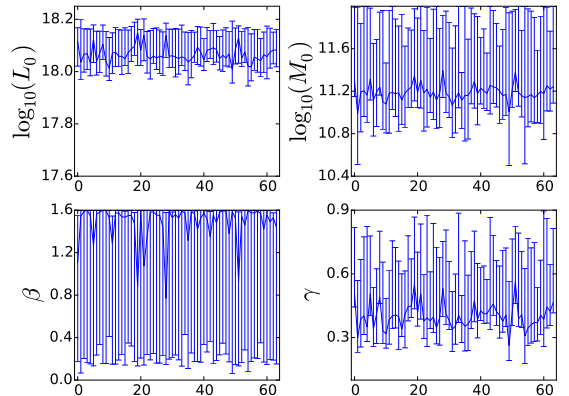


FIG. 5.— Individual small-box parameter estimation with dust attenuation: Best fit values to the Willott data set (solid line) and  $1\sigma$  confidence interval using likelihood ratio  $\mathcal{LR} = 0.5$ . The  $x$  axis corresponds to the box number.

GAVO-DEXTER<sup>2</sup>.

<sup>2</sup> <http://dc.zah.uni-heidelberg.de/dexter/ui/ui/custom>

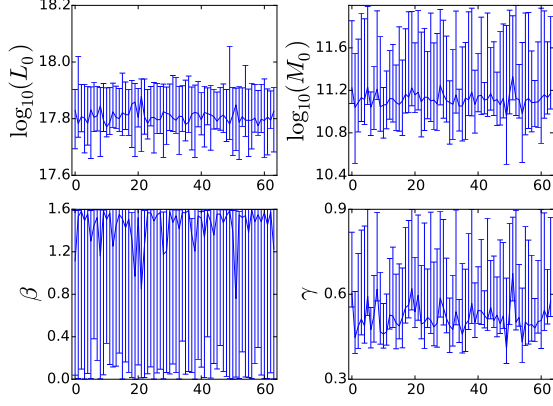


FIG. 6.— Individual small-box parameter estimation with no-dust attenuation: Best fit values to the Willott data set (solid line) and  $1\sigma$  confidence interval using likelihood ratio  $\mathcal{LR} = 0.5$ . The  $x$  axis corresponds to the box number.

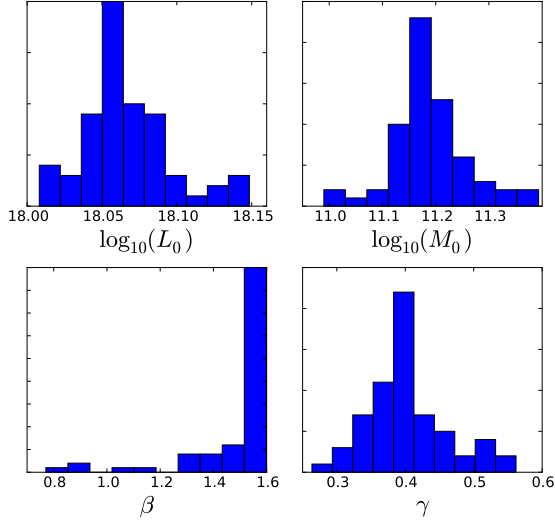


FIG. 7.— Best fit parameter distribution due to cosmic variance with Willott data in the dust attenuation model.

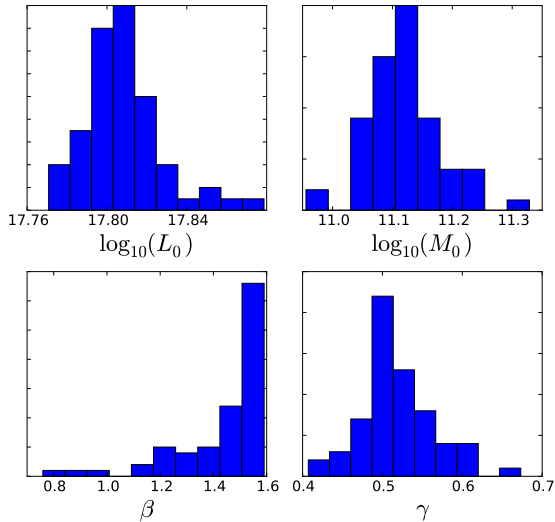


FIG. 8.— Best fit parameter distribution due to cosmic variance with Willott data in the no-dust attenuation model.

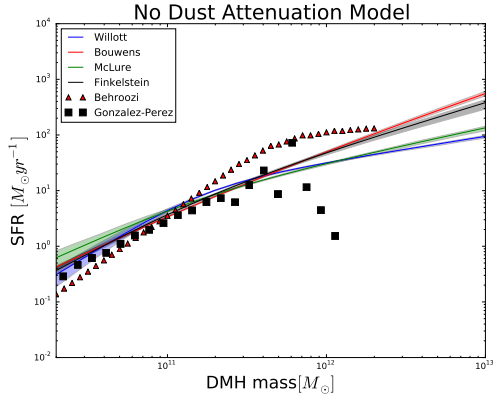


FIG. 9.— Star formation rate as function of the dark matter halo mass without dust attenuation. Solid lines represents the mean SFR value over the small boxes within 50% shaded region. Comparison with the GALFORM semi-analytical model (Gonzalez-Perez et al. 2014) and a implementation of abundance matching model (Behroozi et al. 2013).

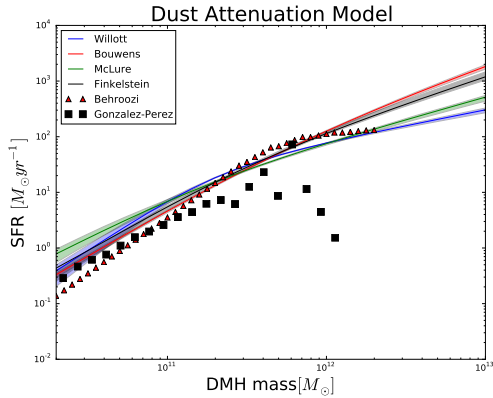


FIG. 10.— Star formation rate as function of the dark matter halo mass with dust attenuation. Solid lines represents the mean SFR value over the small boxes within 50% shaded region. Comparison with the GALFORM semi-analytical model (Gonzalez-Perez et al. 2014) and a implementation of abundance matching model (Behroozi et al. 2013).

## REFERENCES

- Behroozi, P. S., Wechsler, R. H., & Conroy, C. 2013, *ApJ*, 770, 57
- Bouwens, R. J., Illingworth, G. D., Oesch, P. A., Franx, M., Labbé, I., Trenti, M., van Dokkum, P., Carollo, C. M., González, V., Smit, R., & Magee, D. 2012, *ApJ*, 754, 83
- Bouwens, R. J., Illingworth, G. D., Oesch, P. A., Trenti, M., Labbé, I., Bradley, L., Carollo, M., van Dokkum, P. G., Gonzalez, V., Holwerda, B., Franx, M., Spitler, L., Smit, R., & Magee, D. 2015, *ApJ*, 803, 34
- Finkelstein, S. L., Ryan, Jr., R. E., Papovich, C., Dickinson, M., Song, M., Somerville, R., Ferguson, H. C., Salmon, B., Giavalisco, M., Koekemoer, A. M., Ashby, M. L. N., Behroozi, P., Castellano, M., Dunlop, J. S., Faber, S. M., Fazio, G. G., Fontana, A., Grogin, N. A., Hathi, N., Jaacks, J., Kocevski, D. D., Livermore, R., McLure, R. J., Merlin, E., Mobasher, B., Newman, J. A., Rafelski, M., Tilvi, V., & Willner, S. P. 2014, *ArXiv e-prints*
- Gonzalez-Perez, V., Lacey, C. G., Baugh, C. M., Lagos, C. D. P., Helly, J., Campbell, D. J. R., & Mitchell, P. D. 2014, *MNRAS*, 439, 264
- Jones, E., Oliphant, T., Peterson, P., et al. 2001–, *SciPy: Open source scientific tools for Python*
- Kennicutt, Jr., R. C. 1998, *ARA&A*, 36, 189
- Lee, K.-S., Giavalisco, M., Conroy, C., Wechsler, R. H., Ferguson, H. C., Somerville, R. S., Dickinson, M. E., & Urry, C. M. 2009, *ApJ*, 695, 368
- Madau, P., Pozzetti, L., & Dickinson, M. 1998, *ApJ*, 498, 106
- McLure, R. J., Cirasuolo, M., Dunlop, J. S., Foucaud, S., & Almaini, O. 2009, *MNRAS*, 395, 2196
- Meurer, G. R., Heckman, T. M., & Calzetti, D. 1999, *ApJ*, 521, 64
- Moster, B. P., Somerville, R. S., Maubetsch, C., van den Bosch, F. C., Macciò, A. V., Naab, T., & Oser, L. 2010, *ApJ*, 710, 903
- Pérez, F., & Granger, B. E. 2007, *Computing in Science and Engineering*, 9, 21
- Riebe, K., Partl, A. M., Enke, H., Forero-Romero, J., Gottlöber, S., Klypin, A., Lemson, G., Prada, F., Primack, J. R., Steinmetz, M., & Turchaninov, V. 2013, *Astronomische Nachrichten*, 334, 691
- Schmidt, M. 1968, *ApJ*, 151, 393
- Smit, R., Bouwens, R. J., Franx, M., Illingworth, G. D., Labbé, I., Oesch, P. A., & van Dokkum, P. G. 2012, *ApJ*, 756, 14
- Steidel, C. C., Giavalisco, M., Pettini, M., Dickinson, M., & Adelberger, K. L. 1996, *ApJ*, 462, L17
- van den Bosch, F. C., Yang, X., & Mo, H. J. 2003, *MNRAS*, 340, 771

Willott, C. J., McLure, R. J., Hibon, P., Bielby, R., McCracken, H. J., Kneib, J.-P., Ilbert, O., Bonfield, D. G., Bruce, V. A., & Jarvis, M. J. 2013, *AJ*, 145, 4