



The most massive galaxies and black holes allowed by Λ CDM

Peter Behroozi, Joseph Silk

► To cite this version:

Peter Behroozi, Joseph Silk. The most massive galaxies and black holes allowed by Λ CDM. Monthly Notices of the Royal Astronomical Society, 2018, 477 (4), pp.5382 - 5387. 10.1093/mnras/sty945 . hal-01832892

HAL Id: hal-01832892

<https://hal.sorbonne-universite.fr/hal-01832892>

Submitted on 9 Jul 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

The most massive galaxies and black holes allowed by Λ CDM

Peter Behroozi¹★ and Joseph Silk^{2,3}

¹Department of Astronomy and Steward Observatory, University of Arizona, Tucson, AZ 85721, USA

²Institut d'Astrophysique, UMR 7095 CNRS, Université Pierre et Marie Curie, F-75014 Paris, France

³Department of Physics and Astronomy, Johns Hopkins University, Baltimore, MD 21218, USA

Accepted 2018 April 12. Received 2018 April 11; in original form 2016 September 12

ABSTRACT

Given a galaxy's stellar mass, its host halo mass has a lower limit from the cosmic baryon fraction and known baryonic physics. At $z > 4$, galaxy stellar mass functions place lower limits on halo number densities that approach expected Λ CDM halo mass functions. High-redshift galaxy stellar mass functions can thus place interesting limits on number densities of massive haloes, which are otherwise very difficult to measure. Although halo mass functions at $z < 8$ are consistent with observed galaxy stellar masses if galaxy baryonic conversion efficiencies increase with redshift, *JWST* (James Webb Space Telescope) and *WFIRST* (Wide-Field InfraRed Survey Telescope) will more than double the redshift range over which useful constraints are available. We calculate maximum galaxy stellar masses as a function of redshift given expected halo number densities from Λ CDM. We apply similar arguments to black holes. If their virial mass estimates are accurate, number density constraints alone suggest that the quasars SDSS J1044–0125 and SDSS J010013.02+280225.8 likely have black hole mass to stellar mass ratios higher than the median $z = 0$ relation, confirming the expectation from Lauer bias. Finally, we present a public code to evaluate the probability of an apparently Λ CDM-inconsistent high-mass halo being detected given the combined effects of multiple surveys and observational errors.

Key words: galaxies: haloes – early Universe.

1 INTRODUCTION

In the framework of Λ CDM, galaxies form at the centres of dark matter haloes (see Silk & Mamon 2012; Somerville & Davé 2015, for reviews). The ratio of galaxy stellar mass to halo mass has an absolute maximum at the cosmic baryon fraction ($f_b \sim 0.16$; Planck Collaboration et al. 2016a). In practice, stellar feedback processes limit the maximum fraction of baryons converted to $\lesssim 40$ per cent (Behroozi, Conroy & Wechsler 2010; Moster et al. 2010; Behroozi, Wechsler & Conroy 2013; Moster, Naab & White 2013, 2018) even when adopting a Salpeter (1955) initial mass function (IMF). At $z < 4$, this maximum fraction is never achieved for massive haloes ($M_h > 10^{12} M_\odot$), due to inefficient cooling (Lu et al. 2011) and feedback from supermassive black holes (SMBHs) (Silk & Rees 1998). At $z > 4$, however, comparisons of galaxy and halo number densities suggest that massive haloes can reach from 10 to 40 per cent typical integrated efficiencies in converting baryons into stars, again depending on assumptions for the IMF and luminosity–stellar mass conversions (Behroozi et al. 2013; Behroozi & Silk 2015; Finkelstein et al. 2015; Sun & Furlanetto 2016; Moster et al. 2018).

Conversely, an observed galaxy mass (M_*) places a lower limit on its host halo mass (M_h). Λ CDM alone implies that $M_h > M_*/f_b \sim 6.3M_*$, and known baryonic physics would give more stringent limits depending on the assumed maximum conversion efficiency. This fact has been used in Steinhardt et al. (2016) to argue that galaxy number densities at $z \sim 5$ –6 are already inconsistent with Λ CDM. Although we disagree with their assumptions (especially that the M_*/M_h ratio cannot increase at $z > 4$) and therefore also their conclusions, the basic principle that galaxy number densities constrain halo number densities is well established.

As galaxy number densities are consistent with halo number densities for redshifts $z \lesssim 8$ (Behroozi et al. 2013), we compute galaxy mass limits corresponding to expectations from typical Λ CDM baryon fraction limits over $7 < z < 20$, observable with future infrared space-based telescopes (e.g. *JWST*, the *James Webb Space Telescope*, and *WFIRST*, the *Wide-Field InfraRed Survey Telescope*).

Similarly, useful physical thresholds can be calculated for SMBHs. The number density of a given quasar sample places a lower limit on the number density of their host haloes, which in turn limits the maximum average host halo mass in Λ CDM (see also Haiman & Loeb 2001). This then limits the maximum average host galaxy mass (via $M_* < f_b M_h$). Hence, given the number density

★ E-mail: behroozi@email.arizona.edu

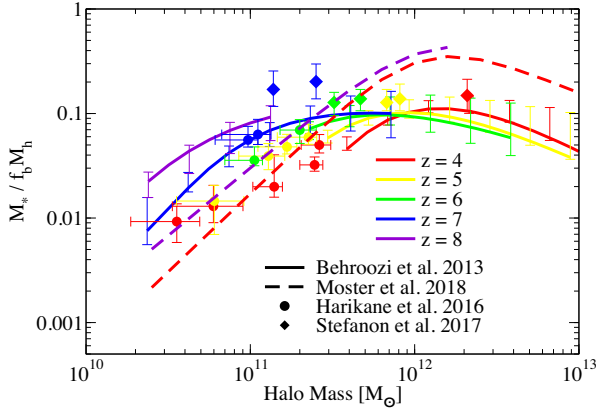


Figure 1. Median stellar mass–baryonic mass ratios at $z = 4$ – 8 reach up to 10–40 per cent. With a scatter of even 0.2 dex (as at low redshifts), it is plausible that individual galaxies can reach ratios near unity. Results have been converted to Salpeter (1955) IMF and to Planck cosmology where appropriate. Behroozi et al. (2013), Stefanon et al. (2017), and Moster et al. (2018) use abundance matching; Harikane et al. (2016) use halo occupation distribution modelling of angular correlation functions.

of black holes above a certain mass, we can derive a lower limit for their $M_\bullet/M_\star$ ratios without requiring observations of the host galaxy. The current claimed maximum $M_\bullet/M_\star$ ratio is 15 per cent (van den Bosch et al. 2012; Seth et al. 2014). In comparison, the highest median relations for the $M_\bullet/M_\star$ ratio at $z = 0$ (Kormendy & Ho 2013; Savorgnan & Graham 2016) give ~ 0.4 per cent for $M_{\text{bulge}} = 10^{11} M_\odot$ (after conversion to a Salpeter 1955 IMF).

Throughout, we assume a flat, Λ CDM cosmology with $\Omega_M = 0.309$, $\Omega_b = 0.0486$, $\sigma_8 = 0.816$, $h = 0.678$, $n_s = 0.967$, corresponding to the best-fitting Planck cosmology (Planck Collaboration et al. 2016a), as well as a Salpeter (1955) IMF. For halo masses, we use the virial overdensity definition of Bryan & Norman (1998).

2 METHODOLOGY

We adopt cumulative halo mass functions (Φ_h) from Behroozi et al. (2013) and define a redshift-dependent maximum cumulative galaxy mass function:

$$\Phi_{\star, \Lambda\text{CDM}}(M_\star, z) \equiv \Phi_h(M_\star/f_b, z) \quad (1)$$

where $\Phi_{\star, \Lambda\text{CDM}}$ is the Λ CDM upper limit on the true stellar mass function – equivalently, it is the expected collapsed gas mass function for Λ CDM. In practice, a given survey volume may contain a higher number density due to sample variance or observational errors, especially when multiple surveys are conducted. Appendix A offers a simple method and code to test whether a given outlier is significantly discrepant from the Λ CDM prediction.

Given that median stellar mass to baryonic mass ratios already reach up to 40 per cent at $z > 4$ (Fig. 1) with the assumption of Salpeter (1955) IMF, and given that measurements of scatter in stellar mass at a fixed halo mass are typically ~ 0.2 dex (e.g. Reddick et al. 2013), it is plausible that some galaxies could reach nearly 100 per cent baryonic conversion efficiencies modulo the effects of stellar mass loss (20–30 per cent; Conroy, Gunn & White 2009).

We also define two cumulative SMBH mass functions:

$$\Phi_{\bullet, \text{max}}(M_\bullet, z) \equiv \Phi_h(M_\bullet/(0.15 f_b), z) \quad (2)$$

$$\Phi_{\bullet, \text{median}}(M_\bullet, z) \equiv \Phi_h(M_\bullet/(0.004 f_b), z) \quad (3)$$

If $\Phi_{\bullet, \text{obs}}$ exceeds $\Phi_{\bullet, \text{max}}$, the black holes must have observed $M_\bullet/M_\star$ ratios of > 15 per cent, regardless of galaxy formation physics. If $\Phi_{\bullet, \text{obs}}$ exceeds $\Phi_{\bullet, \text{median}}$, then some of the SMBHs must have observed $M_\bullet/M_\star$ ratios above the $z = 0$ median relation, again regardless of galaxy formation physics. Specifically, one expects at least a fraction

$$f_\uparrow \equiv \frac{\Phi_{\bullet, \text{obs}} - \Phi_{\bullet, \text{median}}}{\Phi_{\bullet, \text{obs}}} \quad (4)$$

of the SMBHs to have higher-than-median $M_\bullet/M_\star$ ratios. This is not necessarily a surprising finding; indeed, any $M_\bullet$ -selected sample will have at least half of the sample higher than the median relation. However, a high $f_\uparrow$ could suggest evolution in the $M_\bullet/M_\star$ relation or in the scatter in that relation (including observational errors) to high redshifts.

3 RESULTS

For arbitrary future surveys, Fig. 2 also shows cumulative number density thresholds as a function of galaxy stellar mass from $z = 4$ to $z = 20$, including a comparison to the massive galaxies in Stefanon et al. (2015) and Oesch et al. (2016) and to the stellar mass functions in Song et al. (2016). For Oesch et al. (2016), we use their estimated search volume of 1.2×10^6 Mpc. As in Fig. 1, more massive galaxies tend to reach higher $M_\bullet/M_h$ ratios, so their number densities more closely approach the expected collapsed gas mass function from Λ CDM.

Similar to CANDELS (Koekemoer et al. 2011; Grogin et al. 2011) with *Hubble*, a future *JWST* survey may probe galaxy cumulative number densities down to $n_l \sim 10^{-6} \text{ Mpc}^{-3}$. *WFIRST* has a ~ 100 times larger field of view, so it may reach cumulative number densities of $n_w \sim 10^{-8} \text{ Mpc}^{-3}$. For these two threshold densities, we plot $\Phi_{\star, \Lambda\text{CDM}}^{-1}(n, z)$ [i.e. the stellar mass $M_\star(z)$ at which $\Phi_{\star, \Lambda\text{CDM}}(M_\star(z), z) = n$] in Fig. 2. For comparison, we also plot threshold stellar masses for galaxies at $z = 7$ – 8 from extrapolations of Song et al. (2016)¹ and the $z = 11.1$ galaxy from Oesch et al. (2016). Finally, we calculate the expected largest galaxy in the Universe – i.e. using a cumulative number density $n(\geq z)$ such that only one object should exist in the volume of the observable Universe at all redshifts $\geq z$. These volumes are potentially accessible with all sky surveys such as SPHEREx (Doré et al. 2014). We note that finding a larger galaxy is possible as a result of both sample variance and observational errors (see Appendix A).

Threshold masses for black holes are shown in Fig. 3 and compared to the most massive known quasars and blazars at $z > 5$ (Jiang et al. 2007; Willott et al. 2010; Volonteri et al. 2011; De Rosa et al. 2011; Mortlock et al. 2011; Wu et al. 2015; Wang et al. 2015). As bright quasars are detectable in large-area photometric surveys (e.g. the SDSS, Pâris et al. 2014, and the CFHQS, Willott et al. 2007), we calculate mass thresholds at cumulative number densities of $10^{-9.5} \text{ Mpc}^{-3}$ and $10^{-11} \text{ Mpc}^{-3}$. If their virial mass estimates are correct, two quasars, SDSS J1044–0125 (Jiang et al. 2007) and SDSS J010013.02+280225.8 (Wu et al. 2015) likely have $M_\bullet/M_\star$ ratios larger than the $z = 0$ relation even if their host galaxies have 100 per cent $M_\bullet/M_b$ ratios. Because blazars have uncertain beaming

¹Extrapolations for Song et al. (2016) were derived from the posterior distribution of their Schechter function parameters, kindly provided by M. Song.

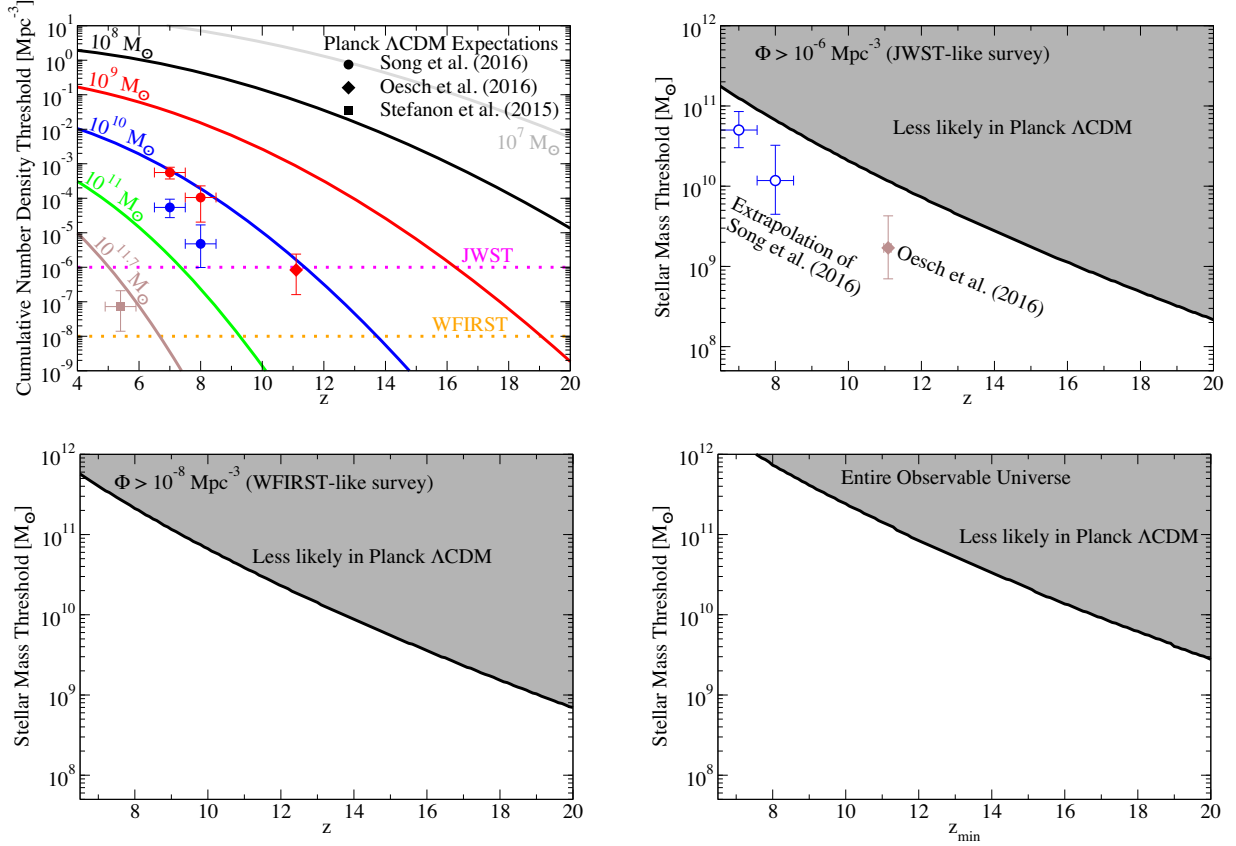


Figure 2. Top-left panel: Cumulative number density thresholds as a function of stellar mass and redshift; observed galaxy cumulative number densities are expected to be below these thresholds in Planck Λ CDM, subject to sample variance and observational errors (see Appendix A). Colored solid lines correspond to different stellar mass thresholds; the brown line corresponds to the $10^{11.7} M_\odot$ mass estimated for the massive galaxy in Stefanon et al. (2015). Colored dotted lines correspond to expected number density limits for the JWST and WFIRST missions. Top-right panel: threshold stellar masses for a cumulative number density of $\Phi = 10^{-6} \text{ Mpc}^{-3}$. If a survey found that galaxies with stellar masses larger than the black line had a cumulative number density higher than 10^{-6} Mpc^{-3} with significant confidence (see Appendix A), it would rule out Λ CDM. Bottom-left panel: same, for a cumulative number density threshold of $\Phi = 10^{-8} \text{ Mpc}^{-3}$. Bottom-right panel: same, for the entire observable Universe (i.e. all sky survey with $z_{\min} < z < \infty$).

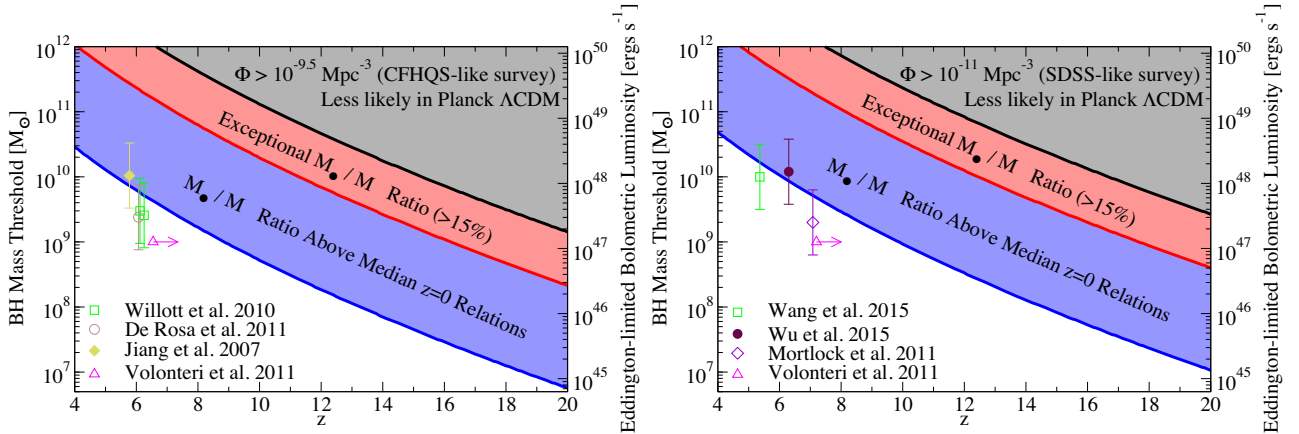


Figure 3. Left-hand panel: threshold black hole masses for a cumulative number density of $\Phi = 10^{-9.5} \text{ Mpc}^{-3}$. If a survey found black holes with masses above the red line and cumulative number densities above $10^{-9.5} \text{ Mpc}^{-3}$, those black holes would exceed the current record $z = 0$ black hole mass–stellar mass ratio. Similarly, if a survey found black holes with masses above the blue line and cumulative number densities above $10^{-9.5} \text{ Mpc}^{-3}$, those black holes would exceed all current determinations of the median $z = 0$ black hole mass–stellar mass ratio. Right-hand panel: same, for a cumulative number density threshold of $\Phi = 10^{-11} \text{ Mpc}^{-3}$. Solid data points indicate the most massive black holes found to date in surveys of the respective volume; open data points indicate less-massive black holes.

corrections, we show lower limits assuming $\Gamma = 5$ from Fig. 5 in Volonteri et al. (2011).

4 DISCUSSION AND CONCLUSIONS

Several factors limit attempts to rule out Λ CDM with galaxy or black hole masses, including observational errors and multiple comparisons (see Appendix A). For galaxies, there are significant uncertainties in converting luminosity to stellar mass (Conroy et al. 2009; Behroozi et al. 2010). Besides systematic offsets, these also induce Eddington/Malmquist bias (Eddington 1913; Malmquist 1922) that artificially inflates the number densities of massive galaxies (Behroozi et al. 2010; Caputi et al. 2011; Grazian et al. 2015). Also present are uncertainties in photo- z codes and priors, which can similarly inflate massive galaxy counts (Stefanon et al. 2015). Just as problematic are multiple peaks in the posterior distribution of z , as for the massive galaxy in Stefanon et al. (2015). We also note the relatively high-lensing optical depth at $z > 8$, which further boosts the apparent number of massive galaxies (Mason et al. 2015). For a full discussion of other sources of systematic error affecting the stellar mass–halo mass relation, we refer readers to Behroozi et al. (2010).

Black hole masses are also subject to many uncertainties (see Peterson 2014, for a review), and virial masses in particular may be overestimates (Shankar et al. 2016). Selecting the largest black hole from a sample with uncertain masses also imposes Eddington bias. Nonetheless, our limits agree with other approaches that infer large black hole mass–stellar mass ratios (Targett, Dunlop & McLure 2012; Venemans et al. 2016), which are expected due to selecting for luminous, massive black holes (Lauer et al. 2007; Volonteri & Stark 2011). We note in passing that blazars are also subject to the same number density constraints; however, estimates of their number densities are made more complicated due to uncertain beaming factors (Ghisellini et al. 2009).

Even so, it is exciting that the highest stellar masses observed in Fig. 2 are so close to the limits expected for Λ CDM. This suggests that high-redshift galaxy surveys will give lower bounds on the evolution of the halo mass function at $z > 8$, which is otherwise very difficult to measure. Combined with constraints on primordial non-Gaussianities and dark matter from faint galaxies (Habouzit et al. 2014; Governato et al. 2015), *JWST* and *WFIRST* will place very interesting limits on early Universe cosmology. For SMBHs, Fig. 3 provides a simple estimate of whether a given $M_{\bullet}$ requires an anomalously high $M_{\bullet}/M_{\star}$ ratio, potentially bolstering the case for follow-up observations.

Finally, we cite examples of ‘unusual’ physics that could be invoked if future observations cross the thresholds outlined here. We refer to ideas beyond usual prescriptions for supernova feedback and active galactic nuclei (AGNs) quenching that limit star formation in current cosmological and zoom-in simulations. Specifically, we mention positive feedback from AGN (as a precursor to the negative feedback observed via AGN-driven massive gas outflows; Gaibler et al. 2012; Ishibashi & Fabian 2012; Silk 2013; Wagner et al. 2016), examples of which are beginning to be found (Zinn et al. 2013; Cresci et al. 2015; Salomé, Salomé & Combes 2015), and a significant duty cycle of hyper-Eddington accretion, increasingly invoked to solve SMBH growth problems at high redshift (Jiang, Stone & Davis 2014; Volonteri, Silk & Dubus 2015; Inayoshi, Haiman & Ostriker 2016). These processes may be able to increase the ratio of stellar mass or black hole mass to total baryonic mass up to the limit imposed by Λ CDM (i.e. the cosmic baryon fraction). Unusual physics that allows accelerated halo growth in

overdense regions (e.g. non-Gaussianities as in Pillepich, Porciani & Hahn 2010, although standard models are now strongly limited by Planck Collaboration et al. 2016b) could also result in overmassive galaxies and black holes that exceed standard Λ CDM limits.

ACKNOWLEDGEMENTS

We thank Robert Antonucci, Andrea Cimatti, Alister Graham, Ryan Trainor, Marta Volonteri, and the anonymous referee for many extremely helpful comments that greatly improved this paper. PB was partially supported by program number HST–HF2–51353.001–A, provided by the National Aeronautics and Space Administration (NASA) through a Hubble Fellowship grant from the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Incorporated, under NASA contract NAS5-26555. JS acknowledges support from the European Research Council Project No. 267117 (DARK) hosted by Pierre et Marie Curie University (UPMC) - Paris 6, France.

REFERENCES

- Behroozi P. S., Silk J., 2015, *ApJ*, 799, 32
- Behroozi P. S., Conroy C., Wechsler R. H., 2010, *ApJ*, 717, 379
- Behroozi P. S., Wechsler R. H., Conroy C., 2013, *ApJ*, 770, 57
- Bryan G. L., Norman M. L., 1998, *ApJ*, 495, 80
- Caputi K. I., Cirasuolo M., Dunlop J. S., McLure R. J., Farrah D., Almaini O., 2011, *MNRAS*, 413, 162
- Conroy C., Gunn J. E., White M., 2009, *ApJ*, 699, 486
- Cresci G. et al., 2015, *ApJ*, 799, 82
- De Rosa G., Decarli R., Walter F., Fan X., Jiang L., Kurk J., Pasquali A., Rix H. W., 2011, *ApJ*, 739, 56
- Doré O. et al., 2014, preprint ([arXiv:1412.4872](https://arxiv.org/abs/1412.4872))
- Dunn O. J., 1958, *Ann. Math. Stat.*, 29, 1095
- Eddington A. S., 1913, *MNRAS*, 73, 359
- Finkelstein S. L. et al., 2015, *ApJ*, 810, 71
- Gaibler V., Khochfar S., Krause M., Silk J., 2012, *MNRAS*, 425, 438
- Ghisellini G., Foschini L., Volonteri M., Ghirlanda G., Haardt F., Burlon D., Tavecchio F., 2009, *MNRAS*, 399, L24
- Governato F. et al., 2015, *MNRAS*, 448, 792
- Grazian A. et al., 2015, *A&A*, 575, A96
- Grogin N. A. et al., 2011, *ApJS*, 197, 35
- Habouzit M., Nishimichi T., Peirani S., Mamon G. A., Silk J., Chevillard J., 2014, *MNRAS*, 445, L129
- Haiman Z., Loeb A., 2001, *ApJ*, 552, 459
- Harikane Y. et al., 2016, *ApJ*, 821, 123
- Inayoshi K., Haiman Z., Ostriker J. P., 2016, *MNRAS*, 459, 3738
- Ishibashi W., Fabian A. C., 2012, *MNRAS*, 427, 2998
- Jiang L., Fan X., Vestergaard M., Kurk J. D., Walter F., Kelly B. C., Strauss M. A., 2007, *AJ*, 134, 1150
- Jiang Y.-F., Stone J. M., Davis S. W., 2014, *ApJ*, 796, 106
- Koekemoer A. M. et al., 2011, *ApJS*, 197, 36
- Kormendy J., Ho L. C., 2013, *ARA&A*, 51, 511
- Lauer T. R., Tremaine S., Richstone D., Faber S. M., 2007, *ApJ*, 670, 249
- Lu Y., Kereš D., Katz N., Mo H. J., Fardal M., Weinberg M. D., 2011, *MNRAS*, 416, 660
- Malmquist K. G., 1922, *Meddelanden fran Lunds Astronomiska Observatorium Serie I*, 100, 1
- Mason C. A. et al., 2015, *ApJ*, 805, 79
- Mortlock D. J. et al., 2011, *Nature*, 474, 616
- Moster B. P., Somerville R. S., Maubetsch C., van den Bosch F. C., Macciò A. V., Naab T., Oser L., 2010, *ApJ*, 710, 903
- Moster B. P., Naab T., White S. D. M., 2013, *MNRAS*, 428, 3121
- Moster B. P., Naab T., White S. D. M., 2018, *MNRAS*, 477, 1822
- Oesch P. A. et al., 2016, *ApJ*, 819, 129
- Pâris I. et al., 2014, *A&A*, 563, A54

- Peterson B. M., 2014, *Space Sci. Rev.*, 183, 253
 Pillepich A., Porciani C., Hahn O., 2010, *MNRAS*, 402, 191
 Planck Collaboration, 2016a, *A&A*, 594, A13
 Planck Collaboration, 2016b, *A&A*, 594, A17
 Reddick R. M., Wechsler R. H., Tinker J. L., Behroozi P. S., 2013, *ApJ*, 771, 30
 Salomé Q., Salomé P., Combes F., 2015, *A&A*, 574, A34
 Salpeter E. E., 1955, *ApJ*, 121, 161
 Savorgnan G. A. D., Graham A. W., 2016, *MNRAS*, 457, 320
 Seth A. C. et al., 2014, *Nature*, 513, 398
 Shankar F. et al., 2016, *MNRAS*, 460, 3119
 Silk J., 2013, *ApJ*, 772, 112
 Silk J., Mamon G. A., 2012, *Res. Astron. Astrophys.*, 12, 917
 Silk J., Rees M. J., 1998, *A&A*, 331, L1
 Somerville R. S., Davé R., 2015, *ARA&A*, 53, 51
 Song M. et al., 2016, *ApJ*, 825, 5
 Stefanon M. et al., 2015, *ApJ*, 803, 11
 Stefanon M., Bouwens R. J., Labbé I., Muzzin A., Marchesini D., Oesch P., Gonzalez V., 2017, *ApJ*, 843, 36
 Steinhardt C. L., Capak P., Masters D., Speagle J. S., 2016, *ApJ*, 824, 21
 Sun G., Furlanetto S. R., 2016, *MNRAS*, 460, 417
 Targett T. A., Dunlop J. S., McLure R. J., 2012, *MNRAS*, 420, 3621
 Trenti M., Stiavelli M., 2008, *ApJ*, 676, 767
 van den Bosch R. C. E., Gebhardt K., Gültekin K., van de Ven G., van der Wel A., Walsh J. L., 2012, *Nature*, 491, 729
 Venemans B. P., Walter F., Zschaechner L., Decarli R., De Rosa G., Findlay J. R., McMahon R. G., Sutherland W. J., 2016, *ApJ*, 816, 37
 Volonteri M., Stark D. P., 2011, *MNRAS*, 417, 2085
 Volonteri M., Haardt F., Ghisellini G., Della Ceca R., 2011, *MNRAS*, 416, 216
 Volonteri M., Silk J., Dubus G., 2015, *ApJ*, 804, 148
 Wagner A. Y., Bicknell G. V., Umemura M., Sutherland R. S., Silk J., 2016, *Astron. Nachr.*, 337, 167
 Wang F. et al., 2015, *ApJ*, 807, L9
 Willott C. J. et al., 2007, *AJ*, 134, 2435
 Willott C. J. et al., 2010, *AJ*, 140, 546
 Wu X.-B. et al., 2015, *Nature*, 518, 512
 Zinn P.-C., Middelberg E., Norris R. P., Dettmar R.-J., 2013, *ApJ*, 774, 66
 Šidák Z., 1967, *J. Am. Stat. Assoc.*, 62, 626

APPENDIX A: ESTIMATING OUTLIER PROBABILITY

When a halo is found in a survey with mass M_h , it is often labelled ‘too massive for Λ CDM’ if the expected number of haloes is less than some threshold ϵ :

$$V \Phi(m > M_h) \equiv V \int_{M_h}^{\infty} \phi(m) dm < \epsilon \quad (\text{A1})$$

where V is the survey volume and $\phi(m)$ is the survey volume-averaged halo number density (per unit mass) for the adopted Λ CDM cosmology.

The true significance is always weaker than ϵ would imply; this is due to both observational errors and multiple comparisons. For extremely rare objects, sample variance from large-scale modes is extremely subdominant to variance from Poisson statistics (see e.g. the cosmic variance calculator in Trenti & Stiavelli 2008); hence, we exclude the former effect from our estimate here. For a single survey, the Poisson chance of observing a halo of mass M_h or larger is

$$P(M > M_h) = 1 - \exp \left[-V \int_{M_h}^{\infty} \left(\int_0^{\infty} \phi(m) P(M|m) dm \right) dM \right] \quad (\text{A2})$$

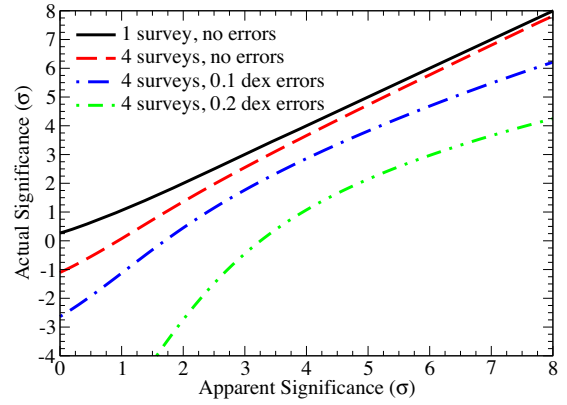


Figure A1. Even assuming standard Λ CDM, observational errors and multiple comparisons mean that it is common to find ‘overmassive’ haloes. For a single survey with no errors, the apparent significance (equation A1) corresponds almost exactly to the true significance. When multiple surveys are conducted, it is common for at least one survey to detect a halo that is slightly overmassive compared to expectations for its volume. However, even small amounts of observational error can result in significantly disparate true versus apparent significances, due to the steepness of the halo mass function.

where m is the true halo mass, M is the observed halo mass estimate, and $P(M|m)$ is the probability density of observing a halo mass M for a true underlying halo mass m .

With multiple surveys, the chance increases that one of the surveys will have an ‘outlier’ even in a standard Λ CDM universe. For the true probability of an ϵ -outlier (according to the definition in equation A1) occurring in at least one of the surveys, we can use the Dunn-Šidák assumption (i.e. multiple fully independent surveys; Šidák 1967) to estimate:

$$P(\epsilon) = 1 - \prod_{i=1}^n \left[1 - P_i \left(M > \Phi_i^{-1} \left(\frac{\epsilon}{V_i} \right) \right) \right] \quad (\text{A3})$$

where P_i is the equivalent of equation A2 for the i th survey, V_i is the i th survey’s volume, and Φ_i^{-1} is the inverse halo cumulative number density for the i th survey. In the regime where $P(\epsilon) < 0.05$, this is nearly identical to the Bonferroni limit ($P(\epsilon) \leq \sum_i P_i$; Dunn 1958).

To encourage correcting for these effects, we developed a public code implementing equations A2–A3.² As an example, we compute the relationship between true outlier significance and apparent (equation A1) significance, assuming that an overmassive object were to be found in at least one of the following surveys:

Description	Area (deg ²)	Redshifts
SDSS-like	14555	0–0.4
SPT-like	2500	0–1.5
EUCLID-like	15000	0–6
LSST-like	20000	0–3

We show results for a range of log-normal observational errors in Fig. A1, presented in terms of more familiar σ units.³ For a single

²<https://bitbucket.org/pbehozi/lcdm-probability>

³Formally, $\sigma \equiv \sqrt{2} \operatorname{erf}^{-1}(2\epsilon - 1)$.

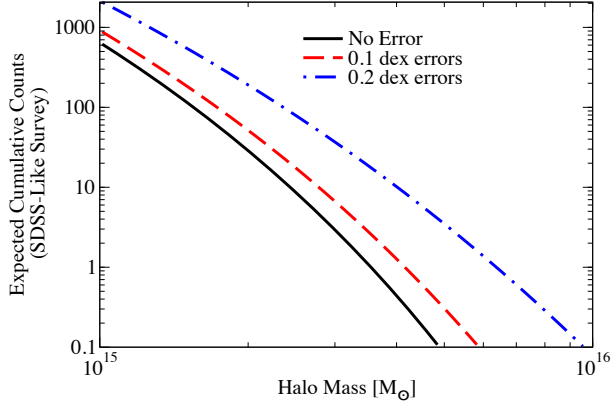


Figure A2. Effect of log-normal observational errors on total halo number counts in an SDSS-like photometric survey.

survey with no errors, the apparent significance (equation A1) is

close to the true significance (equation A2), with the difference arising from the skewness of the Poisson distribution (e.g. an object with average number density of 0.5 per unit volume will be found in less than half of all such unit volumes, because some volumes will have multiple objects). Adding multiple surveys as in the table above increases the chance of detecting a massive object in one of the surveys, as expected. Observational errors result in a very strong effect that is more pronounced for larger surveys. The effect is analogous to a point spread function blurring a sharp image; convolving the steep halo mass function with the observational error distribution results in a shallower falloff (e.g. Fig. A2) and therefore an inflated number density of observed massive haloes compared to the underlying true number density (see also Eddington bias; Eddington 1913).

This paper has been typeset from a $\text{\LaTeX}$ file prepared by the author.