

Star Formation Main Sequence in a Hierarchical Universe

ChangHoon Hahn¹, Jeremy L. Tinker², Andrew R. Wetzel^{3,4,5}

changhoon.hahn@lbl.gov

DRAFT --- 6654c90 --- 2018-01-25 --- NOT READY FOR DISTRIBUTION

ABSTRACT

motivation, methodology, impact. In observations star forming galaxies form a tight $\log M_*$ to $\log SFR$ relation referred to as the *star formation main sequence* (SFMS) out to $z \sim 2$. Beyond the evolution “along” this SFMS, however, the star formation histories of star forming galaxies have not been precisely characterized. The SFH of these galaxies govern SMF, SFMS, and also observed constraints on the stellar mass to halo mass relation.

By combining high-resolution cosmological N -body simulation with observed evolutionary trends of SF galaxies, we construct a model that tracks the evolution of star forming central galaxies over the redshift $z < 1$. Comparing this model

Observations find a remarkably small scatter in the stellar mass to halo mass relation. Somehow the star formation histories of galaxies must

According to observations, star forming galaxies form a tight $\log M_*$ to $\log SFR$ relation referred to as the “star formation main sequence” out to $z \sim 2$.

Subject headings: methods: numerical – galaxies: clusters: general – galaxies: groups: general – galaxies: evolution – galaxies: haloes – galaxies: star formation – cosmology: observations.

We find that by decreasing the timescale of stochasticity on a simple SFH model that traces the overall SFMS evolution does in fact decrease the scatter seen in the SMHMR. However, even with timescales less than XXXX, we cannot reproduce observations. Ultimately to reproduce observations, we need to add in assembly bias.

Checklist

- Check the correlation between halo growth rate with different t_{delay} and δt_{abias} with the total halo growth rate between $z \sim 0$ and $z \sim 1$.

¹Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, CA 94720

²Center for Cosmology and Particle Physics, Department of Physics, New York University, 4 Washington Place, New York, NY 10003

³TAPIR, California Institute of Technology, Pasadena, CA USA

⁴Carnegie Observatories, Pasadena, CA USA

⁵Department of Physics, University of California, Davis, CA USA

1. Introduction

- Motivate why we think SF galaxies evolve along the main sequence
- Discuss the current thought process on galaxy assembly bias
- Explain the limitation of SFH derivable from observations (Claire’s fisher matrix paper would be really good; ask her about the details)
- Observations also can’t provide detail host dark matter halo properties
- So the approach with combining observations with N-body (empirical modeling) is very effective in the context of the halo.
- Maybe talk about how the bigger context of why this is important?
- Why only centrals – because our current best understanding of satellites is that they quench after infall, so it doesn’t make sense to look at them
- our model goes from $z < 1$ because beyond that the observations are statistically meaningless.

2. Central Galaxies of SDSS DR7

We construct our galaxy sample following the sample selection of [Tinker et al. \(2011\)](#). We select a volume-limited sample of galaxies with $M_r 5\log(h) < 18$ and complete in $M_* > 10^{9.4} M_\odot$ from the NYU Value-Added Galaxy Catalog (VAGC; [Blanton et al. 2005](#)) of the Sloan Digital Sky Survey Data Release 7 (SDSS DR7; [Abazajian et al. 2009](#)) at $z \approx 0.04$. The stellar masses of these galaxies are estimated using the `kcorrect` code ([Blanton & Roweis 2007](#)) assuming a [Chabrier \(2003\)](#) initial mass function. The star formation of the galaxies are estimated spectroscopically using the specific star formation rates (SSFR) from the current release of the MPA-JHU spectral reductions¹ ([Brinchmann et al. 2004](#)). Generally speaking, $\text{SSFR} > 10^{-11} \text{yr}^{-1}$ are derived from H_α emission, $10^{-11} > \text{SSFR} > 10^{-12} \text{yr}^{-1}$ are derived from a combination of emission lines, and $\text{SSFR} < 10^{-12} \text{yr}^{-1}$ are based on D_n4000 (see discussion in [Wetzel et al. 2013](#)). We note that $\text{SSFR} < 10^{-12} \text{yr}^{-1}$ should only be considered upper limits to the true galaxy SSFR ([Salim et al. 2007](#)).

From our galaxy sample, we identify the central galaxies using the [Tinker et al. \(2011\)](#) halo-based group-finding algorithm, which is based on the [Yang et al. \(2005\)](#) algorithm and tested in [Campbell et al. \(2015\)](#). The algorithm assigns a probability of being a satellite,

¹<http://wwwmpa.mpa-garching.mpg.de/SDSS/DR7/>

P_{sat} , to each galaxy in the sample. Galaxies with $P_{\text{sat}} \geq 0.5$ are classified as satellites and $P_{\text{sat}} < 0.5$ are classified as centrals. In this paper we focus on central galaxies. With any group finding algorithm, galaxies are misassigned due to projection effects and redshift space distortions. The purity of the full central galaxy sample is $\sim 90\%$ with a completeness of $\sim 95\%$ (Tinker et al. 2017). Furthermore, Campbell et al. (2015) find that the algorithm robustly identifies red and blue centrals as a function of stellar mass, which is highly relevant to our analysis.

In the left panel of Figure 1, we plot the SFR- M_* distribution of the SDSS DR7 central galaxies. In the right panel, we plot the distribution of SSFR, $p(\log \text{SSFR})$, for galaxies with $10.6 < \log M_* < 10.8$ (stellar mass range highlighted on the left panel). Both panels of Figure 1 illustrate the bimodality in the galaxy sample. The SFR- M_* distribution also illustrates the correlation between SFR and M_* in star-forming galaxies *i.e.* the star-formation main sequence (SFMS).

3. Model: Simulated Central Galaxies

We’re interesting in constructing a model that tracks central galaxies and their star formation within the hierarchical growth of their host halos. This requires a cosmological N -body simulation that accounts for the complex dynamical processes that govern the host halos of galaxies. In this paper we use the high resolution N -body simulation from Wetzel et al. (2013) generated using the White (2002) TreePM code with flat Λ CDM cosmology ($\Omega_m = 0.274$, $\Omega_b = 0.0457$, $h = 0.7$, $n = 0.95$, and $\sigma_8 = 0.8$). From initial conditions at $z = 150$ generated from second-order Lagrangian Perturbation Theory, 2048^3 particles with mass of $1.98 \times 10^8 M_\odot$ are evolved in a $250\text{Mpc}/h$ box with a Plummer equivalent smoothing of $2.5\text{kpc}/h$. For a more detailed description of the simulation, we refer readers to Wetzel et al. (2013, 2014).

From the TreePM N -body simulation, ‘host halos’ are identified using the Friends-of-Friends (FoF) algorithm of Davis et al. (1985) with linking length of $b = 0.168$ times the mean inter-particle spacing. Within these host halos, Wetzel et al. (2013) identifies ‘subhalos’ as overdensities in phase space through a six-dimensional FoF algorithm (FoF6D White et al. 2010). The host halos and subhalos are then tracked across the 45 simulation outputs from $z = 10$ to 0 to build merger trees (Wetzel et al. 2009; Wetzel & White 2010). The most massive subhalos in newly-formed host halos at a given simulation output are defined as the ‘central’ subhalo. A central subhalo retains its ‘central’ definition until it falls into a more massive host halo, at which point it becomes a ‘satellite’ subhalo.

Each subhalo is assigned a M_{peak} , the maximum host halo mass that it ever had as a central subhalo. Using M_{peak} , we construct a galaxy catalog from the subhalos using subhalo abundance matching (SHAM; Conroy et al. 2006; Vale & Ostriker 2006; Yang et al. 2009; Wetzel et al. 2012; Leja et al. 2013; Wetzel et al. 2013, 2014; Hahn et al. 2017). In principle,

SHAM assumes a one-to-one mapping between subhalo M_{peak} and galaxy stellar mass M_* : $n(> M_{\text{peak}}) > n(> M_*)$ that preserves the rank ordering. In practice, we apply a 0.2 dex log-normal scatter in M at fixed M_{peak} based on observations of the stellar mass to halo mass relation (SMHMR; **bunch of SMHMR citations**). Gu et al. (2016) compile empirical constraints on the scatter of this stellar mass to halo mass relation ($\sigma_{\log M_*}$). Using the SHAM mapping, we can assign galaxy stellar mass to subhalos based on observed stellar mass functions (SMFs) at the redshifts of the simulation outputs (snapshots).

We use the SMF from Li & White (2009) at $z = 0.05$ and at higher redshifts interpolate between the Li & White (2009) SMF and the SMF from Marchesini et al. (2009) at $z = 1.6$. We choose the Li & White (2009) SMF because it is based on the same SDSS NYU-VAGC sample as our SDSS DR7 group catalog (Section 2). We choose the Marchesini et al. (2009) SMF, amongst others, because it produces interpolated SMFs that monotonically increase at $z < 1$. As noted in Hahn et al. (2017), at $z \approx 1$, the SMF interpolated between the Li & White (2009) and Marchesini et al. (2009) SMFs is consistent with more recent measurements from Muzzin et al. (2013) and Ilbert et al. (2013). At each snapshot, we independently use SHAM to assign galaxy M_* . This way, we not only track the evolution of subhalos, but also the the galaxies’ M_* . With the 45 snapshots outputs from our simulation, we can in principle track the central galaxies back to $z \sim 10$. However, we restrict ourselves to snapshots at $z \lesssim 1$, where we have the most statistically meaningful observations. We next describe how we select star forming central galaxies in our model and initialize them. **TBD: Perhaps mention in appendix how we test different SMF assumptions**

3.1. Selecting Star Forming Central Galaxies

In our model, we’re interested in tracking the evolution of the SFR and stellar mass of SF central galaxies. To construct such a model, we first need to select star-forming galaxies from the central galaxies in our simulation described earlier in this section. Since we later compare our model to observation, our selection is based on $f_{\text{SFMS}}^{\text{cen}}(M_*)$, the fraction of central galaxies within the star-forming main sequence, measured from the SDSS DR7 VAGC (Section 2). Below, we describe how we derive this $f_{\text{SFMS}}^{\text{cen}}(M_*)$ and use it to select star-forming central galaxies in our model. Afterwards we describe how we initialize the SFRs and M_* of these galaxies in our model.

Often in the literature, an empirical color-color or SFR- M_* cut that separates the two main modes (red/blue or star-forming/quiescent) in the distribution is chosen to classify galaxies (*e.g.* Baldry et al. 2006; Blanton & Moustakas 2009; Drory et al. 2009; Peng et al. 2010; Moustakas et al. 2013; Hahn et al. 2015). The red/quiescent or blue/star-forming fractions derived from this sort of classification, by construction, depend on the choice of cut and neglect the transitioning galaxies *i.e.* the galaxies in the “green valley”. For our $f_{\text{SFMS}}^{\text{cen}}(M_*)$ measurement, we instead use a method from Tjitske et al. (in prep), which is

based on fitting the SFMS from the SFR- M_* distribution

The SFMS fitting scheme first divides the SDSS DR7 VAGC central galaxy sample (Section 2) into stellar mass bins of width $\Delta \log M_* = 0.2$ dex. We then fit the SSFR distribution of each bin using Gaussian mixture models (GMMs) with 1 – 3 components using the expectation-maximization algorithm (EM; Dempster et al. 1977; Neal & Hinton 1998). We restrict ourselves to models with a maximum of 3 components for the three possible galaxy classifications: quiescent, star-forming, and green valley populations. From the three GMMs, we select the model with the lowest Bayesian Information Criteria (BIC Schwarz 1978). The Gaussian component of this GMM with mean $\log \text{SSFR} > -11$ is identified as the SFMS. In the rare cases when more than one GMM component has mean $\log \text{SSFR} > -11$, we compare the weights of the components. If the weight of one component is less than a third of the other, we take the component with the higher weight to represent the SFMS. Otherwise, we omit the stellar mass bin altogether. The weight of the SFMS GMM component provides an estimate of $f_{\text{SFMS}}^{\text{cen}}$ for the given stellar mass bin. In the right panel of Figure 1, we plot the SFMS GMM component (blue shaded region) of the $p(\log \text{SSFR})$ for the SDSS DR7 central galaxies within $10.6 < \log M_* < 10.8$. The SFMS constitutes $f_{\text{SFMS}}^{\text{cen}} = 0.21$ of the SDSS central galaxies in this stellar mass bin.

Next, with the $f_{\text{SFMS}}^{\text{cen}}$ values spanning the different stellar mass bins, we fit $f_{\text{SFMS}}^{\text{cen}}$ as a function of $\log M_*$. Using a fiducial stellar mass of $\log M_{\text{fid}} = 10.5$, we derive the following best-fit

$$f_{\text{SFMS,bestfit}}^{\text{cen}}(M_*) = -0.627(\log M_* - 10.5) + 0.354. \quad (1)$$

We note that this $f_{\text{SFMS,bestfit}}^{\text{cen}}(M_*)$ is in good agreement with the $f_{\text{Q}}^{\text{cen}}(M_*; z \sim 0)$ fit from Hahn et al. (2017). For each central galaxy in our simulation, we assign a probability of it being on the SFMS, using $f_{\text{SFMS,bestfit}}^{\text{cen}}(M_*)$ with M_* at $z \sim 0$ assigned through SHAM. Based on these probabilities, we randomly identify central galaxies from our simulation as star-forming at $z \sim 0$.

At this point we make the assumption that once a SF galaxy quenches its star formation, it remains quiescent. This means that all of the central galaxies that we place on the SFMS at $z \sim 0$, were also on the SFMS at $z \sim 1$ — the initial redshift of our model. From SHAM we have the initial M_* s of our SF centrals at $z \sim 1$. Their initial SFRs, we can get from observations of the SFMS. The SFMS in the literature, however, have significant discrepancies. In Speagle et al. (2014), they compile SFRs of the SFMS from 25 studies in the literature, each with different methods of deriving galaxy SFR and M_* . Even after the calibration, the SFRs of the SFMS from different analyses at $z \sim 1$ for $M_* = 10^{10.5} M_\odot$ vary by more than a factor of 2 (see Figure 2 of Speagle et al. 2014). This is particularly concerning since the amplitude and slope of the SFMS significantly dictate the SFR and M_* evolution of SF centrals in our model. In order to deal with this uncertainty, we use the following prescription for the SFMS SFR: $\log \text{SFR}_{\text{MS}}(M_*, z)$.

At $z \sim 0$, the central galaxies of the SDSS DR7 group catalog provide an anchoring measurement of the SFMS. We therefore fix $\log \text{SFR}_{\text{MS}}(M_*, z = 0)$ to $\log \text{SFR}(M_*)$ of the SDSS central galaxy SFMS. To obtain the SFMS $\log \text{SFR}(M_*)$ of the SDSS centrals, we again take the SFMS GMM components, used to estimate $f_{\text{SFMS}}^{\text{cen}}$, at each stellar mass bin and linearly fit their mean $\log \text{SFR}$ as a function of M_* . For $z > 0$, we include two free parameters to dictate the redshift evolution of the amplitude and slope of the SFMS $\text{SFR}-M_*$ relation: A_z and m_z , respectively.

$$\log \text{SFR}_{\text{MS}}(M_*, z) = \quad (2)$$

Later in our analysis, we set the following priors for A_z and m_z : **prior**. These priors are conservatively chosen to encompass the best-fit SFMS from [Speagle et al. \(2014\)](#) and also measurements from [Moustakas et al. \(2013\)](#); [Lee et al. \(2015\)](#).

3.2. Evolving along the Main Sequence

- talk about how we assign the initial SFR using the SHAM M^*
- Talk about the SFR and M_* prescriptions
- parameterization of SFMS
- explicitly talk about the free parameters of the model.
- priors for β_M and β_z encompass the observable constraints
- talk about inference using ABC

4. Results

4.1. The duty cycle of star formation

- Figure that illustrates the fit to observables
- Figure of sigma M star as a function of duty cycle compared to observations

4.2. The need for a galaxy assembly bias

- discuss how t_{duty} is not enough to be consistent with σ_{M^*} .
- first clarify what you mean by galaxy assembly bias
- discuss implementation of galaxy assembly bias

- Figure (pedagogical) of dlogSFR versus dMh dt for different correlation amounts
- Figure of different tdelay and dtbias
- Figure of σM_{star} as a function of duty cycle and realistic dt bias and t delay

5. Discussion

5.1. Rethinking the Main Sequence?

- Test the SMHMR for Louis’s SFHs

6. Summary

A. $z \sim 1$ observations

Much of the results presented in this paper are based on comparison between our model and observations at $z \sim 0.$. Our model is initialized at $z \sim 1$. Therefore, in this section we test some of the choices we make in our initializations.

- Test impact of $z \sim 1$ SMF
- Test impact of $z \sim 1$ $\sigma_{\log M_{\star}}$

Acknowledgements

Louis Abramson

REFERENCES

- Abazajian, K. N., Adelman-McCarthy, J. K., Agüeros, M. A., et al. 2009, [The Astrophysical Journal Supplement Series](#), 182, 543
- Baldry, I. K., Balogh, M. L., Bower, R. G., et al. 2006, [Monthly Notices of the Royal Astronomical Society](#), 373, 469
- Blanton, M. R., & Moustakas, J. 2009, [Annual Review of Astronomy and Astrophysics](#), 47, 159
- Blanton, M. R., & Roweis, S. 2007, [The Astronomical Journal](#), 133, 734
- Blanton, M. R., Schlegel, D. J., Strauss, M. A., et al. 2005, [The Astronomical Journal](#), 129, 2562
- Brinchmann, J., Charlot, S., White, S. D. M., et al. 2004, [Monthly Notices of the Royal Astronomical Society](#), 351, 1151

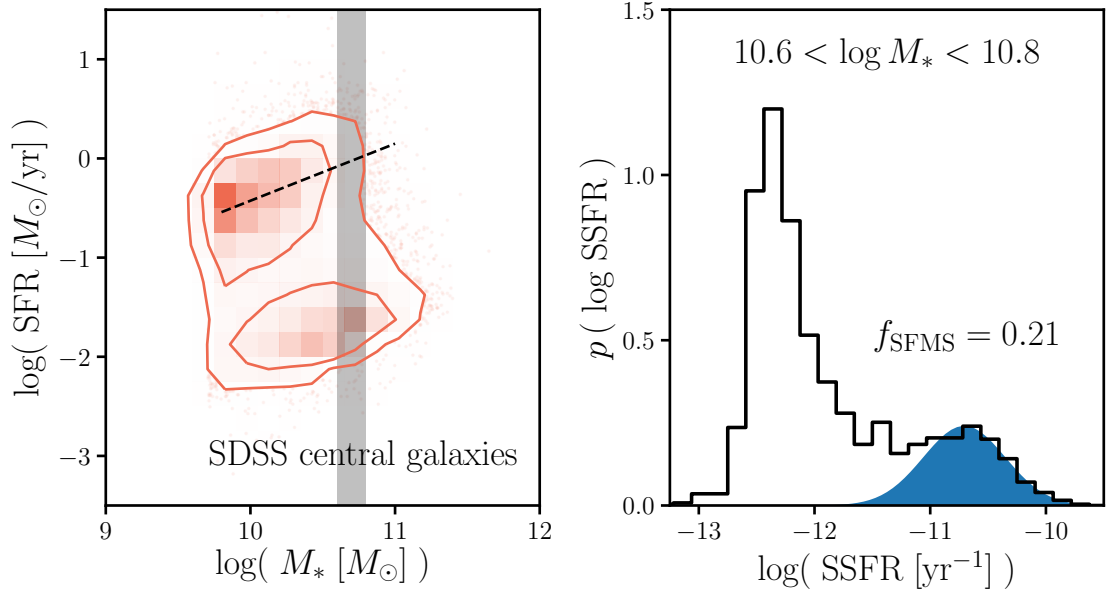


Fig. 1.— Central galaxies in the SDSS DR7 group catalog. *Left:* . *Right:* Fitting of the SFMS.

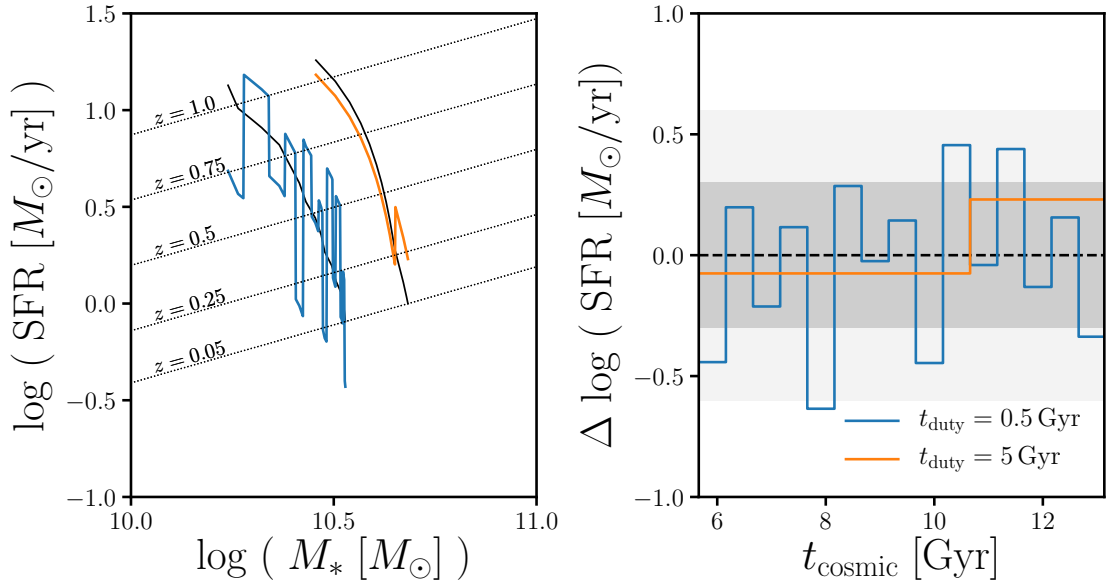


Fig. 2.— Pedagogical figure that illustrates how star forming central galaxies in our model evolve along the SFMS.

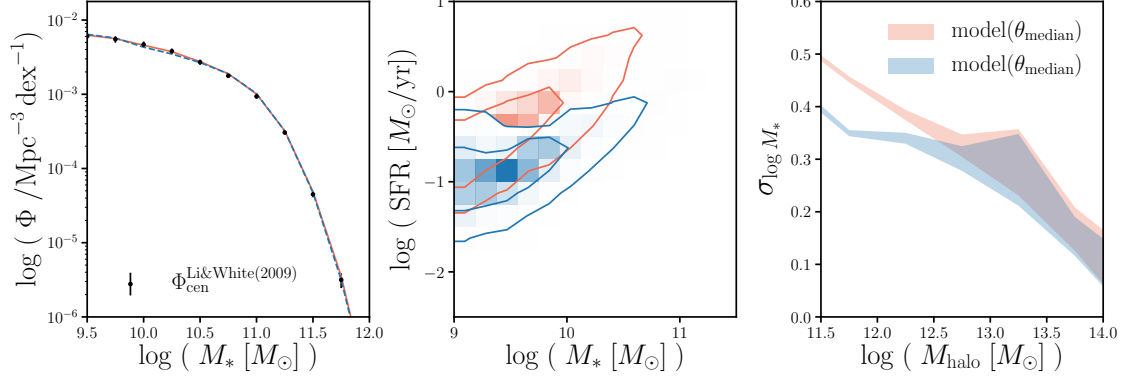


Fig. 3.—

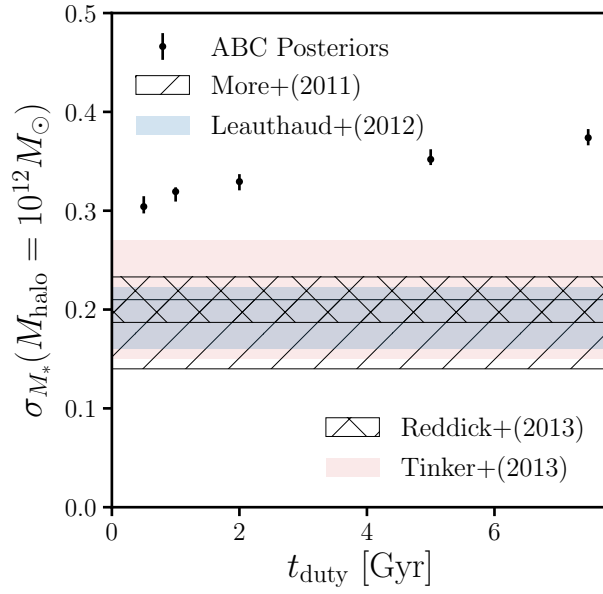


Fig. 4.—

- Campbell, D., van den Bosch, F. C., Hearin, A., et al. 2015, [Monthly Notices of the Royal Astronomical Society](#), 452, 444
- Chabrier, G. 2003, [Publications of the Astronomical Society of the Pacific](#), 115, 763
- Conroy, C., Wechsler, R. H., & Kravtsov, A. V. 2006, [The Astrophysical Journal](#), 647, 201
- Davis, M., Efstathiou, G., Frenk, C. S., & White, S. D. M. 1985, [The Astrophysical Journal](#), 292, 371
- Dempster, A. P., Laird, N. M., & Rubin, D. B. 1977, [Journal of the Royal Statistical Society. Series B \(Methodological\)](#), 39, 1
- Drory, N., Bundy, K., Leauthaud, A., et al. 2009, [The Astrophysical Journal](#), 707, 1595
- Gu, M., Conroy, C., & Behroozi, P. 2016, [The Astrophysical Journal](#), 833, 2
- Hahn, C., Tinker, J. L., & Wetzel, A. R. 2017, [The Astrophysical Journal](#), 841, 6
- Hahn, C., Blanton, M. R., Moustakas, J., et al. 2015, [The Astrophysical Journal](#), 806, 162
- Ilbert, O., McCracken, H. J., Le Fèvre, O., et al. 2013, [Astronomy and Astrophysics](#), 556, A55
- Lee, N., Sanders, D. B., Casey, C. M., et al. 2015, [The Astrophysical Journal](#), 801, 80
- Leja, J., van Dokkum, P., & Franx, M. 2013, [The Astrophysical Journal](#), 766
- Li, C., & White, S. D. M. 2009, [Monthly Notices of the Royal Astronomical Society](#), 398, 2177
- Marchesini, D., van Dokkum, P. G., Förster Schreiber, N. M., et al. 2009, [The Astrophysical Journal](#), 701, 1765
- Moustakas, J., Coil, A. L., Aird, J., et al. 2013, [The Astrophysical Journal](#), 767, 50
- Muzzin, A., Marchesini, D., Stefanon, M., et al. 2013, [The Astrophysical Journal](#), 777, 18
- Neal, R. M., & Hinton, G. E. 1998, in [Learning in Graphical Models](#), NATO ASI Series (Springer, Dordrecht), 355
- Peng, Y.-j., Lilly, S. J., Kovač, K., et al. 2010, [The Astrophysical Journal](#), 721, 193
- Salim, S., Rich, R. M., Charlot, S., et al. 2007, [The Astrophysical Journal Supplement Series](#), 173, 267
- Schwarz, G. 1978, [The Annals of Statistics](#), 6, 461
- Speagle, J. S., Steinhardt, C. L., Capak, P. L., & Silverman, J. D. 2014, [The Astrophysical Journal Supplement Series](#), 214, 15
- Tinker, J., Wetzel, A., & Conroy, C. 2011, ArXiv e-prints, 1107, arXiv:1107.5046
- Tinker, J. L., Hahn, C., Mao, Y.-Y., & Wetzel, A. R. 2017, arXiv:1705.08458 [astro-ph], [arXiv:1705.08458 \[astro-ph\]](#)
- Vale, A., & Ostriker, J. P. 2006, [Monthly Notices of the Royal Astronomical Society](#), 371, 1173
- Wetzel, A. R., Cohn, J. D., & White, M. 2009, [Monthly Notices of the Royal Astronomical Society](#), 395, 1376
- Wetzel, A. R., Tinker, J. L., & Conroy, C. 2012, [Monthly Notices of the Royal](#)

- [Astronomical Society](#), 424, 232
- Wetzel, A. R., Tinker, J. L., Conroy, C., & van den Bosch, F. C. 2013, [Monthly Notices of the Royal Astronomical Society](#), 432, 336
- . 2014, [Monthly Notices of the Royal Astronomical Society](#), 439, 2687
- Wetzel, A. R., & White, M. 2010, [Monthly Notices of the Royal Astronomical Society](#), 403, 1072
- White, M. 2002, [The Astrophysical Journal Supplement Series](#), 143, 241
- White, M., Cohn, J. D., & Smit, R. 2010, [Monthly Notices of the Royal Astronomical Society](#), 408, 1818
- Yang, X., Mo, H. J., & van den Bosch, F. C. 2009, [The Astrophysical Journal](#), 695, 900
- Yang, X., Mo, H. J., van den Bosch, F. C., & Jing, Y. P. 2005, [Monthly Notices of the Royal Astronomical Society](#), 356, 1293