

Discovering the True Regulator of Stellar Mass

Andrew P. Hearin

10 February 2015

ABSTRACT

Empirical models of the galaxy-halo connection are built upon the assumption that some host halo property serves as the statistical regulator of the stellar mass M_* content of the halo. This is the core assumption that permits us to use N-body simulations in cosmological parameter inference. In this paper I introduce the weighted cumulative abundance technique (WCA), which allows *the choice of halo property itself* to be parameterized and continuously varied in a likelihood analysis. There are significant, systematic differences between empirical models based on different halo properties, and it has recently been shown that the associated systematic error known as *galaxy assembly bias* threatens the precision cosmology program. I describe how to use WCA to generalize abundance matching models as well as halo occupation distribution (HOD) models to make them more robust to assembly bias. I conclude by discussing the role that the WCA technique will play in future constraints on cosmology with large-scale structure.

1 INTRODUCTION

(i) The halo model is the premier tool to constrain dark energy with large-scale structure measurements.

(ii) How does that work? First you enumerate dn_h/dM_{vir} and $b_h(M_{\text{vir}})$ through a painful calibration exercise.

Next you write down a parameterized form of an occupation model such as $\Phi(M_*|M_{\text{vir}})$ or $\langle N_{\text{gal}} \rangle(M_{\text{vir}})$.

Finally, you use e.g., clustering and lensing predictions of the model to constrain both cosmological and galaxy-halo parameters.

(iii) However, if you base your model on different halo properties, your model predictions may be qualitatively distinct, because assembly bias is scale dependent. Halo mass M_{vir} is the most widespread, but its definition is based on spherical top-hat collapse, which is poorly motivated in detail. The impact of the choice of halo parameter has been poorly studied heretofore, because there has been no way to easily compare models based on different properties, and because precision demands have only recently reached this point.

(iv) This paper provides a new technique that allows us to construct and compare models that are identical in every way, but differ by the choice of halo property.

The technique also allows us to parameterize our choice of the halo property in a continuous fashion, so that we can marginalize over our uncertainty in the true halo property.

(v) The AM formulation of the generalization is in § 2, and the HOD/CLF formulation is in § 3.

2 ABUNDANCE MATCHING WITH A GENERALIZED HALO PROPERTY

2.1 Traditional Implementation

The basic principle of abundance matching is that there exists a single (sub)halo property which is in monotonic correspondence with the stellar mass of the galaxy residing at the center of the (sub)halo. Once the amount of scatter between this halo property and M_* is specified, the mapping between galaxies and dark matter halos is determined.

The standard implementation of this model begins by first selecting a particular halo property, such as peak historical mass M_{peak} . Next, one computes the cumulative number density of halos and subhalos as a function of M_{peak} , which we will write as $N_{\text{halo}}(> M_{\text{peak}})$. This calculation is extraordinarily simple. The (sub)halos are rank-ordered according to the value of M_{peak} , and the cumulative number density of the i^{th} -rank (sub)halo is given by

$$N_{\text{halo}}(> M_{\text{peak}}^i) = \frac{i}{V_{\text{box}}},$$

where V_{box} is the simulated volume.

The key insight behind abundance matching is that there exists a unique deterministic mapping $\bar{M}_*(M_{\text{peak}})$ that preserves the observed stellar mass function. Once the value of $N_{\text{halo}}(> M_{\text{peak}})$ is known for every (sub)halo, this mapping can be numerically calculated by finding the zero of the function

$$\mathcal{F}(M_*|x) \equiv N_{\text{halo}}(> x) - N_{\text{gal}}(> M_*), \quad (1)$$

where $N_{\text{gal}}(> M_*)$ is the observed cumulative stellar mass function. For any particular choice of (possibly M_{peak} -dependent) stochasticity, there also exists

a unique map $\bar{M}_*(M_{\text{peak}})$ that includes this scatter and preserves the observed $N_{\text{gal}}(> M_*)$. Determining this map and introducing stochasticity between M_* and M_{peak} can be accomplished by the deconvolution method described in Section 3.3.1 of Behroozi et al. (2010) or in Appendix A of Kravtsov (2013).

Conventionally, the upshot of the above calculation is the determination of the mapping $\bar{M}_*(M_{\text{peak}})$. This well-studied map is most commonly employed in the literature with convenient fitting functions such as those given in Behroozi et al. (2010); Moster et al. (2010); Kravtsov (2013).

2.2 Weighted Cumulative Abundance: application to abundance matching

Ok, so the core quantity that we need to do abundance matching with, say, V_{peak} , is just the value of $N_{\text{halo}}(> V_{\text{peak}})$ for every (sub)halo in our box. Computing this quantity is trivial,

$$N_{\text{halo}}(> V_{\text{peak}}^i) = i/L_{\text{box}}^3,$$

where i is the V_{peak} -based rank-order of the (sub)halos.

That’s all the standard jam. But now suppose we want to conduct AM using a property x that is “between” V_{peak} and M_{peak} . We begin in the normal way by computing the numerical value of $N_{\text{halo}}(> M_{\text{peak}})$ and $N_{\text{halo}}(> V_{\text{peak}})$ for every (sub)halo in the box, exactly as above. Next we define the following *weighted cumulative abundance*:

$$N_{\text{halo}}(> x) \equiv w N_{\text{halo}}(> V_{\text{peak}}) + (1 - w) N_{\text{halo}}(> M_{\text{peak}}), \quad (2)$$

where w is any number between 0 and 1. This gives us a numerical value of $N_{\text{halo}}(> x)$ for every (sub)halo in the box. Now we just use Eq. 1 to assign M_* to every (sub)halo using AM based on our newly defined property x .

Notice that we never even assigned values of x to the (sub)halos! So this trick entirely circumvents the problem that M_{peak} and V_{peak} do not even have the same units. **Upshot:** we can allow w to be a continuously varying parameter in an MCMC, allowing us to marginalize over our uncertainty about the “true” halo property in the AM-prediction of M_* . This provides a natural extension to the abundance matching parameter scan explored in Reddick et al. (2013).

3 THE HOD/CLF WITH A GENERALIZED HALO PROPERTY

Now let’s build new families of HOD models $P(N_{\text{gal}}|x)$ based on a generalized host halo property x that’s “between” M_{vir} and V_{max} .

(i) Write down your favorite parameterization for $P(N_{\text{gal}}|M_{\text{vir}})$.

(ii) Compute $N_{\text{halo}}(> M_{\text{vir}})$ for every halo.¹

¹ Again, this is computed as i/L_{box}^3 , except this time neglecting subhalos from all aspects of the calculation, as is usual for this type of model.

The outcome of this step is to provide a bijective mapping between $N_{\text{halo}}(> M_{\text{vir}}) \leftrightarrow M_{\text{vir}}$.

(iii) Use the mapping provided by Step (2) to recast

$$P(N_{\text{gal}}|M_{\text{vir}}) \Rightarrow P(N_{\text{gal}}|N_{\text{halo}}(> M_{\text{vir}})).$$

This is accomplished by function composition:

$$\langle N_{\text{gal}} \rangle(M_{\text{vir}}) \Rightarrow \langle N_{\text{gal}} \rangle(N_{\text{halo}}(> M_{\text{vir}})).$$

This doesn’t change our original model in the slightest; we’re now just using a new label $N_{\text{halo}}(> M_{\text{vir}})$ in our equations.

(iv) Compute $N_{\text{halo}}(> V_{\text{max}})$ for every halo, giving us $N_{\text{halo}}(> V_{\text{max}}) \leftrightarrow V_{\text{max}}$.

(v) By replacing all appearances of $N_{\text{halo}}(> M_{\text{vir}})$ with $N_{\text{halo}}(> V_{\text{max}})$, we construct a new HOD model $P(N_{\text{gal}}|V_{\text{max}})$ that is the “ V_{max} -equivalent” of our original model $P(N_{\text{gal}}|M_{\text{vir}})$.

(vi) More than this, we can also construct entirely new models using

$$N_{\text{halo}}(> x) \equiv w N_{\text{halo}}(> V_{\text{max}}) + (1 - w) N_{\text{halo}}(> M_{\text{vir}}).$$

Those models will be the “ x -equivalent” HOD model to our original. Again, at no point do we ever need to compute the value of x .

This gives us a recipe to explore entirely new models that are in every way equivalent to our original $P(N_{\text{gal}}|M_{\text{vir}})$, but instead belong to a continuous family of halo properties spanned by V_{max} and M_{vir} .

4 CONCLUSION

Going beyond vir is not just a matter of taste. The danger is to rule out GR. This is a serious danger. Currently, no guarantees. You just hope you’ve chosen the right property, based on very little at all. Mostly heuristic arguments that are known to fail in detail.

REFERENCES

- Behroozi P. S., Conroy C., Wechsler R. H., 2010, ApJ , 717, 379
- Kravtsov A. V., 2013, ApJL , 764, L31
- Moster B. P., Somerville R. S., Maubetsch C., van den Bosch F. C., Macciò A. V., Naab T., Oser L., 2010, ApJ , 710, 903
- Reddick R. M., Wechsler R. H., Tinker J. L., Behroozi P. S., 2013, ApJ , 771, 30