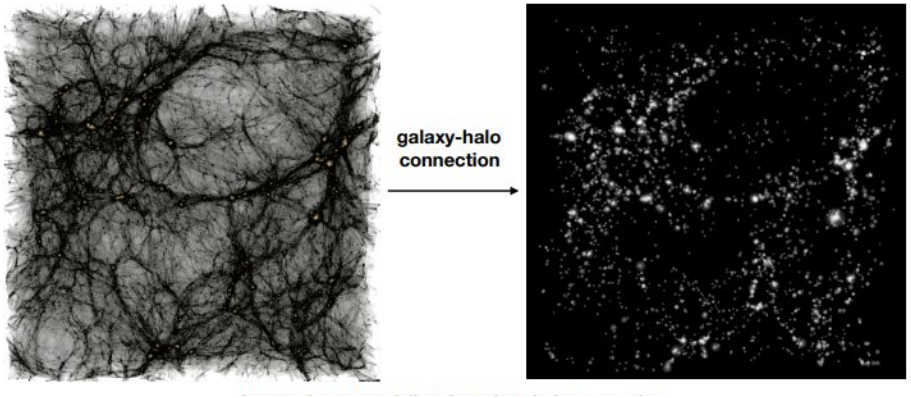
Tackling challenges in galaxy-dark matter connection modeling and secondary biases

Sihan Yuan

with Daniel Eisenstein, Lehman Garrison, Alexie Leauthaud,
Boryana Hadzhiyska, Hong Guo, and Sownak Bose

The importance of galaxy-dark matter connection

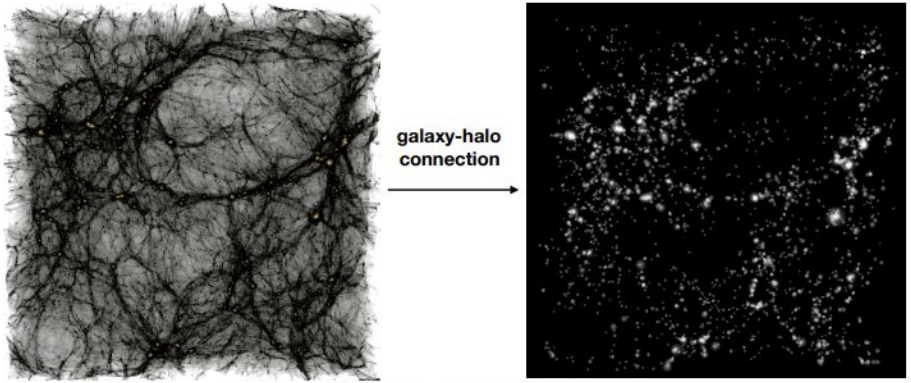


Wechsler & Tinker 2018

- Inferring cosmological parameters.
- Understanding the physics of galaxy formation.
- Probing the structure and properties of dark matter.

The march to small scales!

The various approaches



Wechsler & Tinker 2018

Approaches to modeling the galaxy-halo connection

physical models ← → empirical models

Hydrodynamical Simulations	Semi-analytic Models	Empirical Forward Modeling	Subhalo Abundance Modeling	Halo Occupation Models
Simulate halos & gas; Star formation & feedback recipes	Evolution of density peaks plus recipes for gas cooling, star formation, feedback	Evolution of density peaks plus parameterized star formation rates	Density peaks (halos & subhalos) plus assumptions about galaxy—(sub)halo connection	Collapsed objects (halos) plus model for distribution of galaxy number given host halo properties

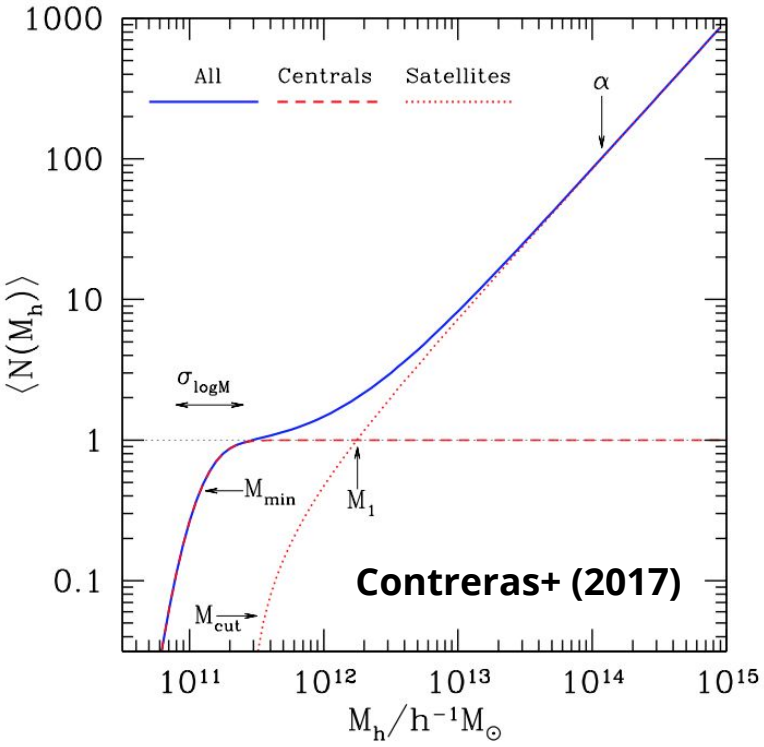
Expensive ← Computation cost → Cheap

The Halo Occupation Distribution (HOD) model

- The HOD abstracts dark matter into gravitationally bound halos, which are then analytically associated with galaxies.

- The HOD is essential for large high resolution surveys such as DESI.
  - Cheap enough to process very large simulation volumes.
  - Takes advantage of the subscale physics built into simulations.

The Halo Occupation Distribution (HOD) model



$$\langle N_{\rm gal}(M_{\rm halo})\rangle = \langle N_{\rm cen}(M_{\rm halo})\rangle + \langle N_{\rm sat}(M_{\rm halo})\rangle$$

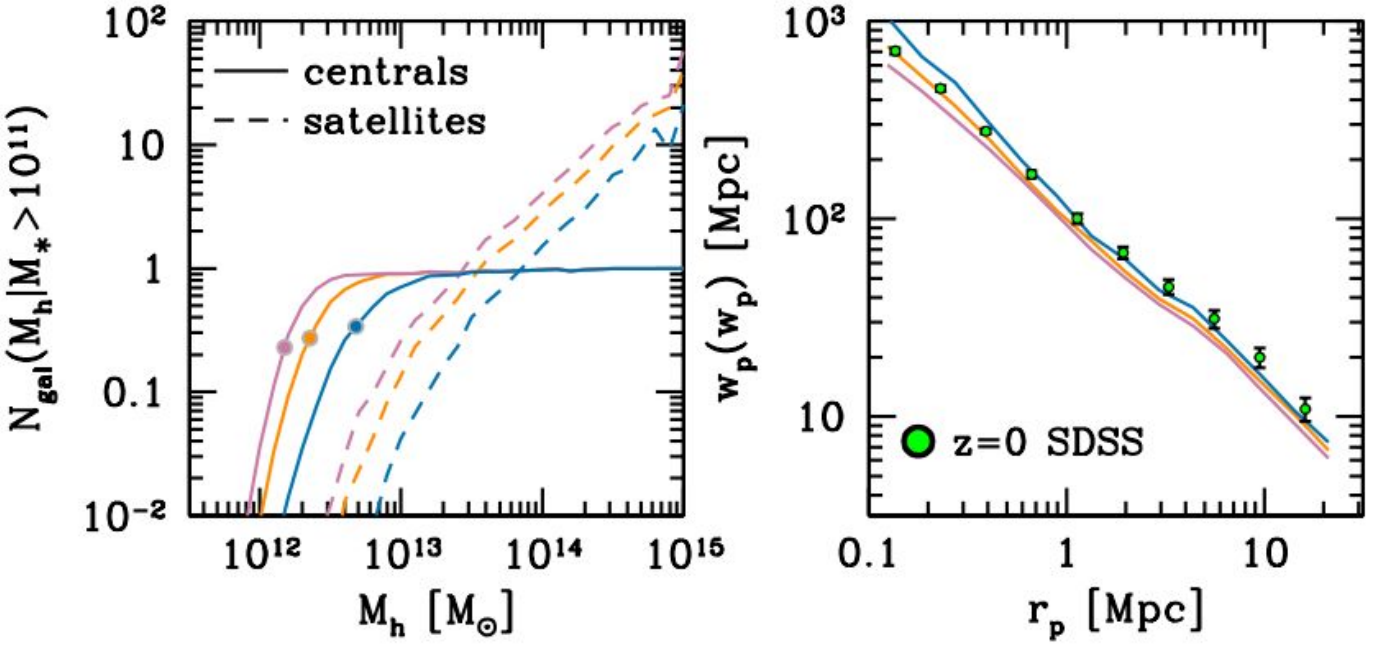
$$\langle N_{\rm cen}(M_{\rm halo})\rangle = \frac{1}{2}\left[1 + {\rm erf}\left(\frac{\log {\rm M}_{\rm halo} - \log {\rm M}_{\rm min}}{\sigma_{\log {\rm M}}}\right)\right]$$

$$\langle N_{\rm sat}(M_{\rm halo})\rangle = \left(\frac{M_{\rm halo} - M_{\rm cut}}{M_1}\right)^{\alpha}$$

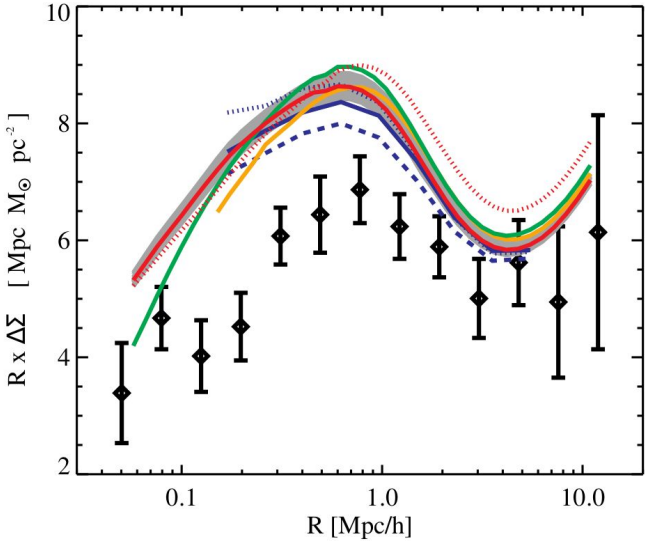
Jing, Mo & Börner (1998); Benson+ (2000); Peacock & Smith (2000); Berlind & Weinberg (2002), Zheng+ (2005) ....

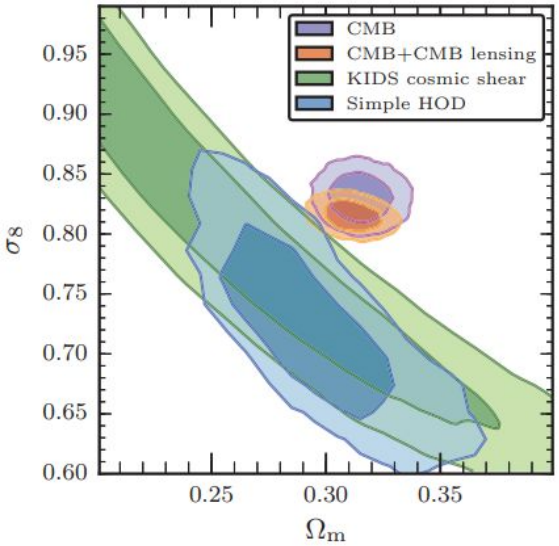
The Halo Occupation Distribution (HOD) model

Wechsler & Tinker 2018



The lensing tension

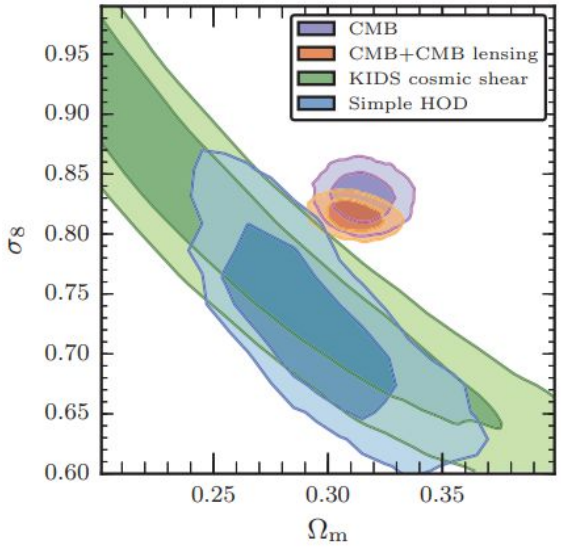




Leauthaud+17

Systematics? Modeling? Cosmology?

The lensing tension



Linder 2005:

$\sigma_8(\gamma)$

γ --- Growth index, derived from GR

Growth rate informs underlying gravity models.

The lensing tension can be significantly remedied with more realistic HOD models

(Yuan et al. 2020b)

Assembly bias

- *Assembly bias* is the phenomenon when the clustering of halos/galaxies at fixed mass depends on additional properties of the halo (i.e. other than mass)

Formation time
Concentration
Spin
Velocity dispersion
… and so on

- Early forming halos have more time to undergo mergers and tidal disruption.
- Leads to fewer but more luminous galaxies.

- Can significantly bias cosmology/HOD inference (Lange+2019).
- Observational evidence: Zentner+2014, Miyatake+2016.

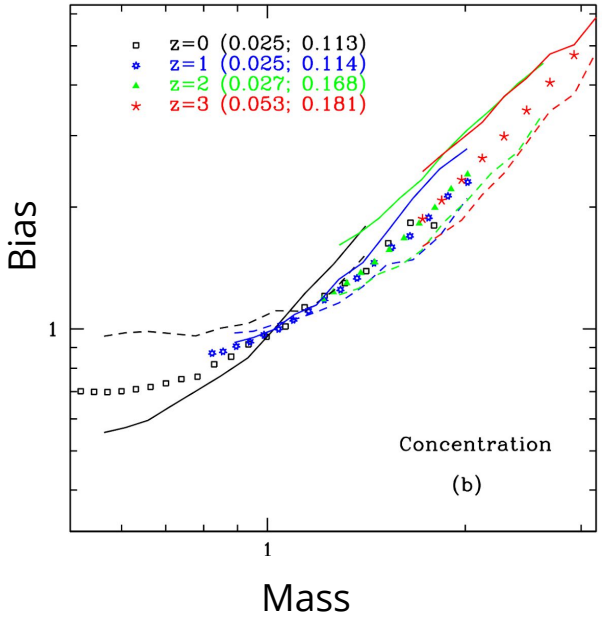
Assembly bias

- Halo concentration has been the most popular marker of assembly bias.

$$c = \frac{r_{\rm vir}}{r_s}$$

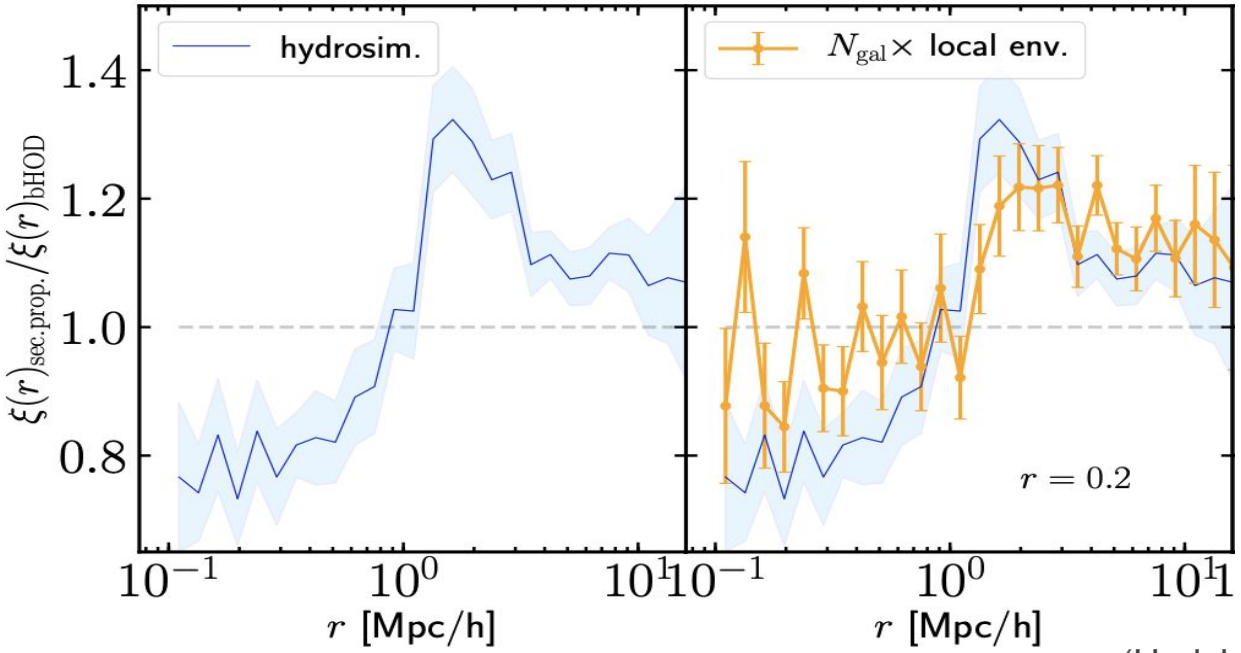
  - Concentration correlates well with formation time.

- In the HOD model: $P(N_g \mid M, c)$.



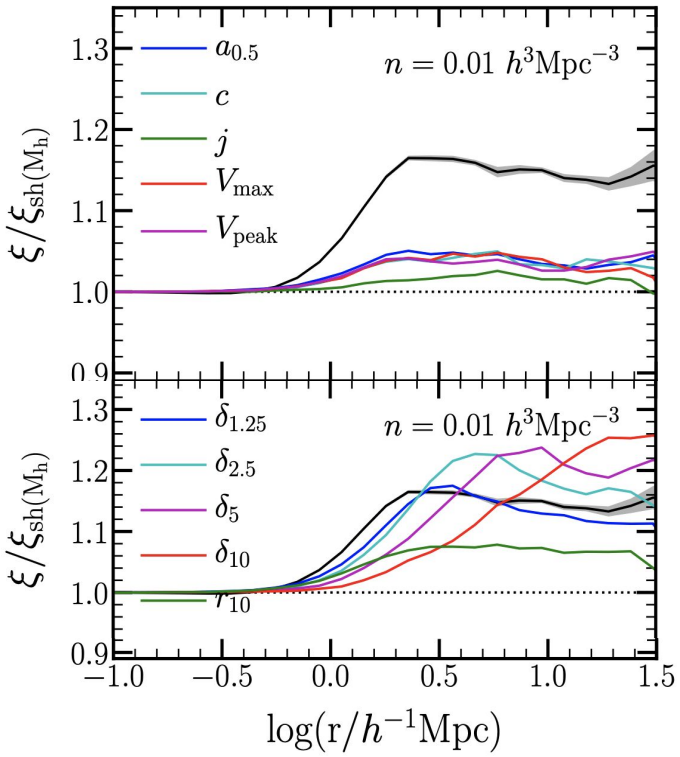
Gao & White 2007

Environment-based secondary bias



(Hadzhiyska et al 2020)

Environment-based secondary bias



(Xu+2020)

Yuan et. al. 2018, 2020b

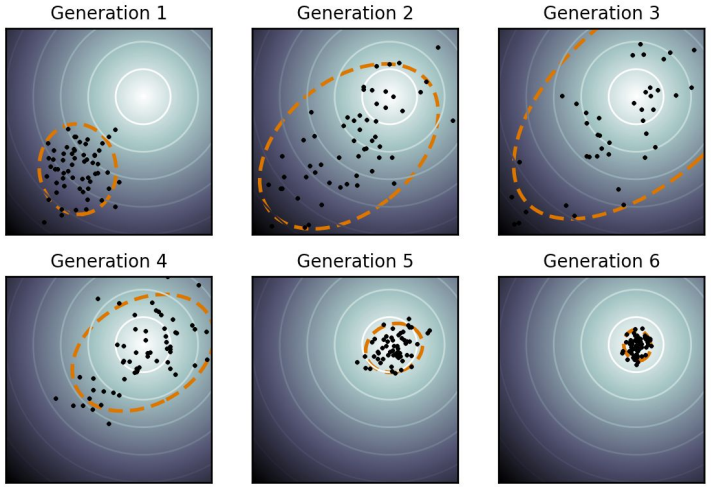
Extended HOD including secondary biases

- Vanilla parameters. $\rightarrow [M_{\rm cut}, M_1, \sigma, \alpha, \kappa]$
- Generalized parameters:
  - Assembly bias based on concentration and environment. $\rightarrow [A, A_e]$
  - Satellite radial distribution parameters. $\rightarrow [s, s_p]$
  - Velocity bias parameters for centrals and satellites. $\rightarrow [\alpha_c, s_v, s_r]$

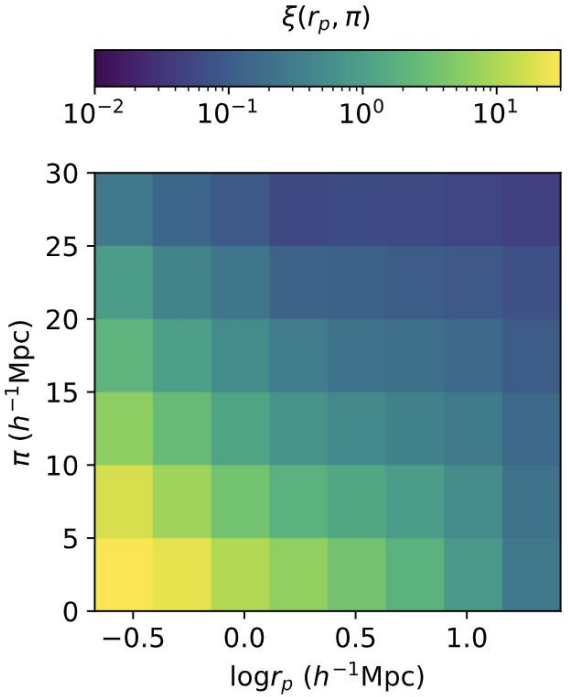
Yuan et. al. 2020b

Fitting the BOSS clustering with extended HOD

- Data:
  - BOSS CMASS galaxies within 0.46 < z < 0.61 (DR12).
  - Fiber-collision corrected.
- Algorithm:
  - Evolutionary global optimization routine (CMAES).



(wikipedia)

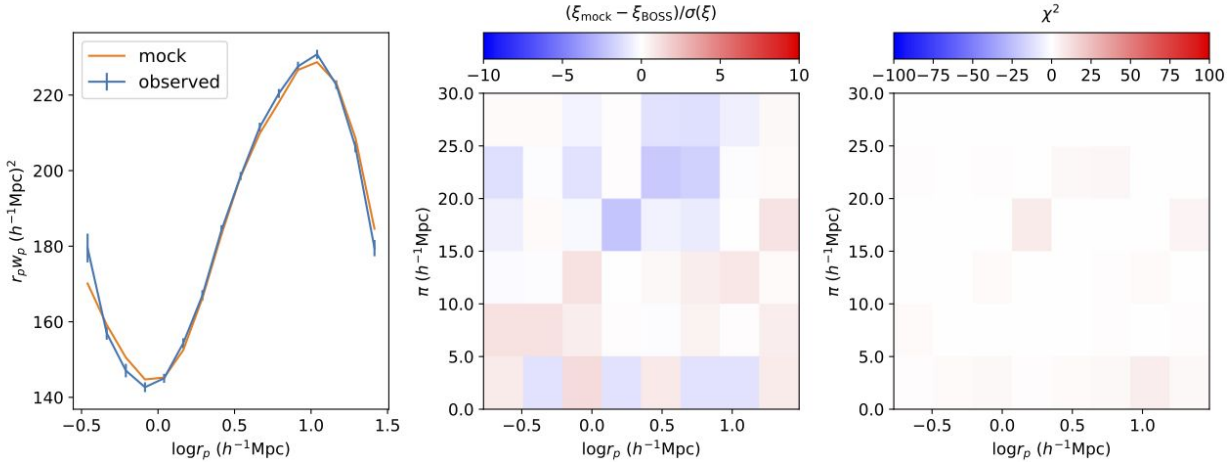


Yuan et. al. 2020b

Fitting the BOSS redshift-space 2PCF

- Best fit: □^{2} = 50 (d.o.f = 37).
- Preference for including both secondary biases:
  - Include A: ΔBIC = 21.
  - Include Ae: ΔBIC = 17.
  - Combined: ΔBIC = 36.

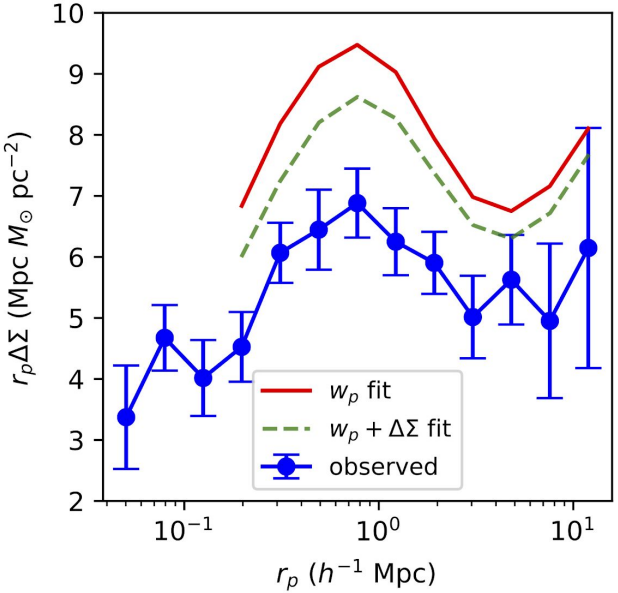
Best fit:
A = -0.7± 0.2
Ae = 0.04 ± 0.01



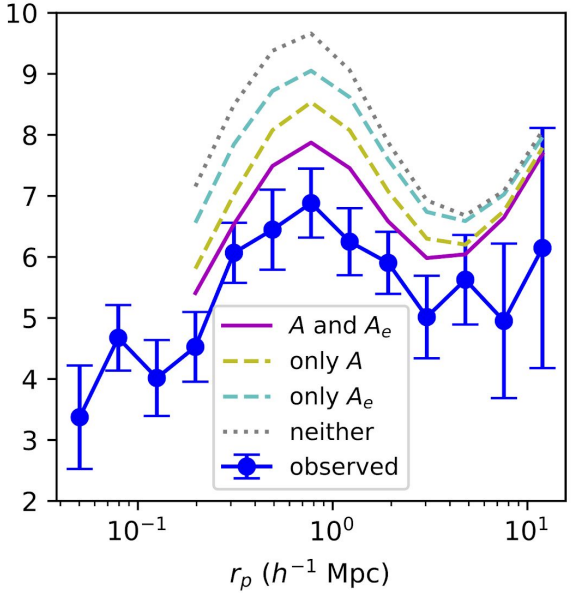
Yuan et. al. 2020a, b

The lensing prediction

Previous predictions

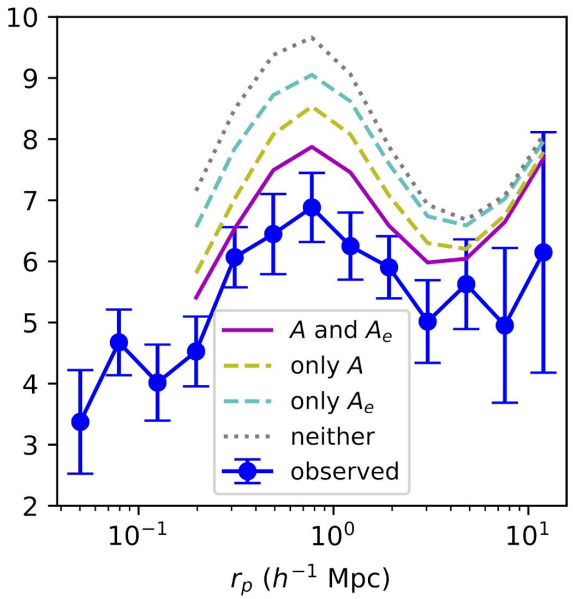


New predictions



Yuan et. al. 2020b

A path towards resolving the lensing tension?



Average halo mass per galaxy:

- No secondary bias:
  4.1e13 Msun,
- Include A:
  3.6e13 Msun,
- Include Ae:
  3.7e13 Msun,
- Include both:
  3.3e13 Msun.

The LOS structure of the 2PCF is pushing galaxies into lower mass halos.

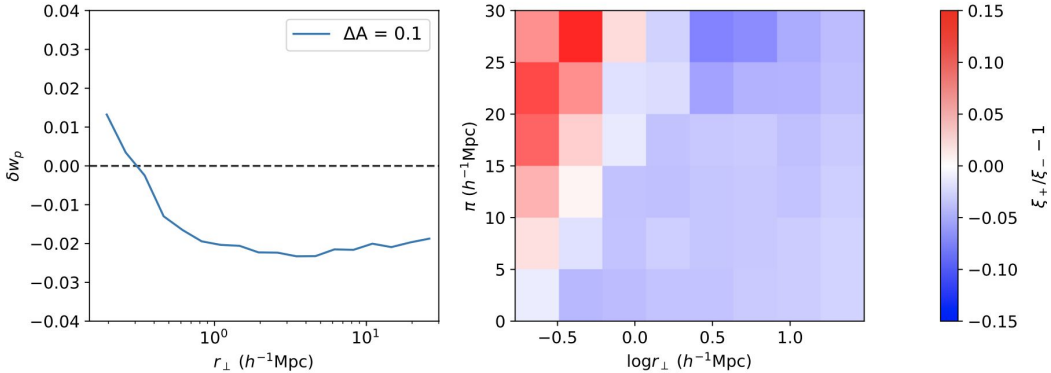
Yuan et. al. 2020b

What is exactly driving these secondary biases?

Projected 2PCF

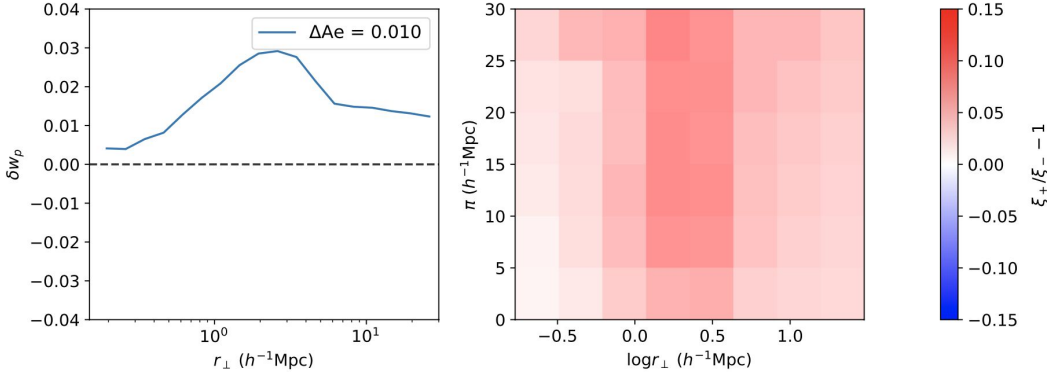
Redshift-space 2PCF

Concentration-based assembly bias **A**:



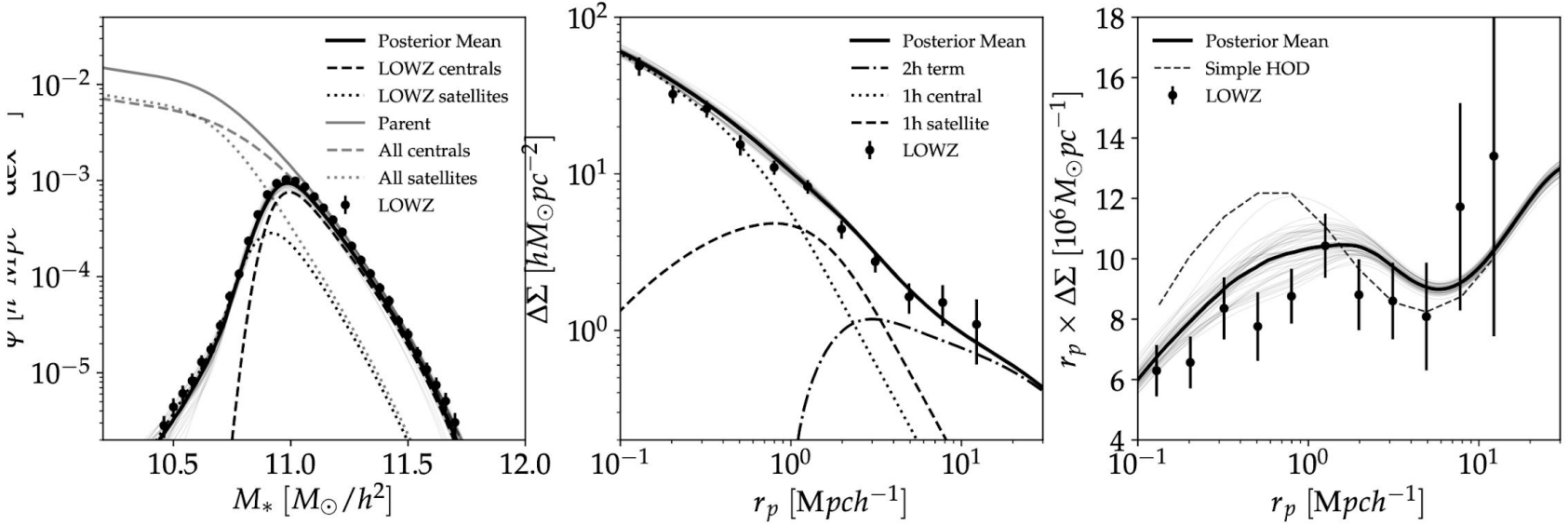
A < 0:
colder halos.

Environment-based secondary bias **Ae**:



Ae > 0:
Extended profiles.

Other solutions to “lensing is low”?

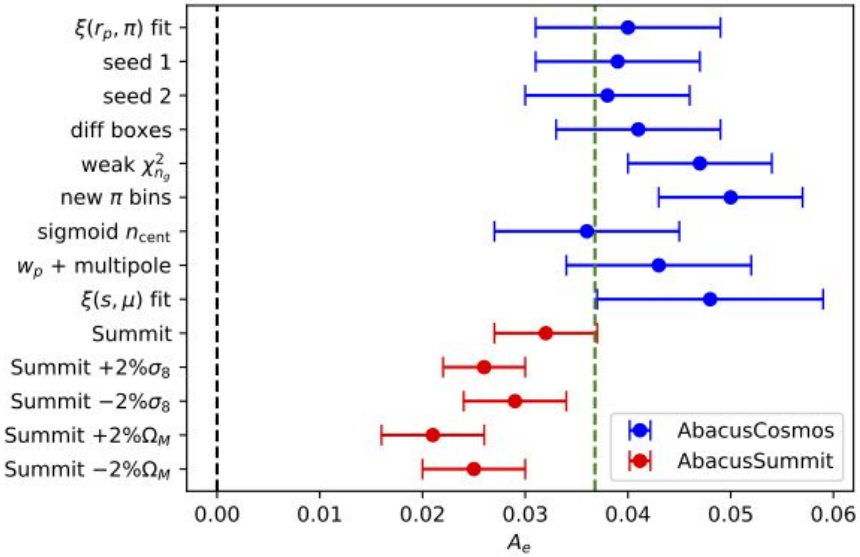


Issues: 1) satellite fraction 80% :(

Zu Ying 2020

Yuan et. al. 2020b

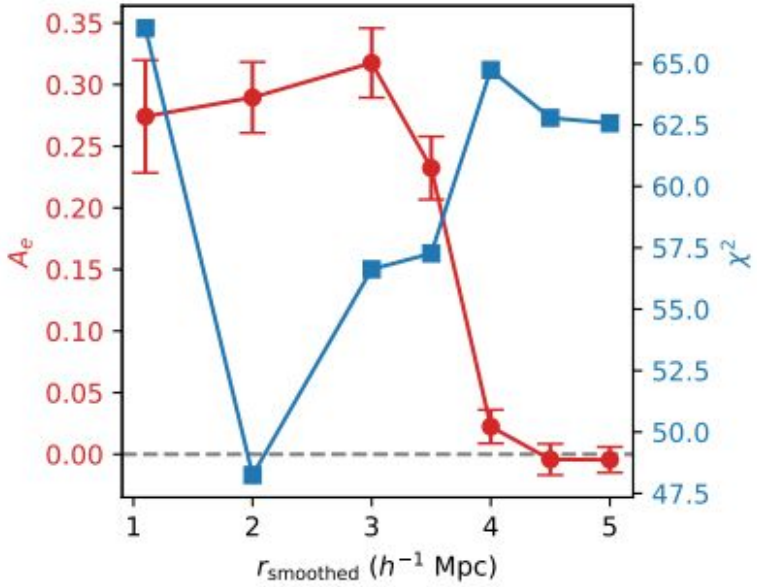
A positive detection of Ae



- A consistent detection of Ae across all fits:
- Ae might depend on cosmology (need more testing).

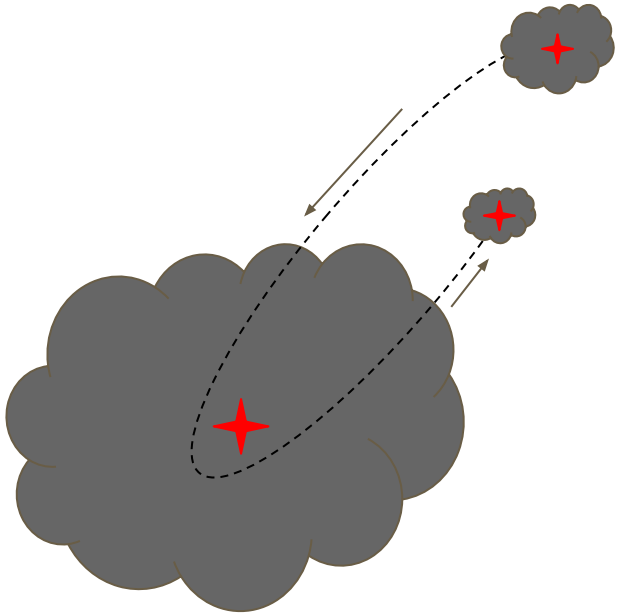
Yuan et. al. 2020b

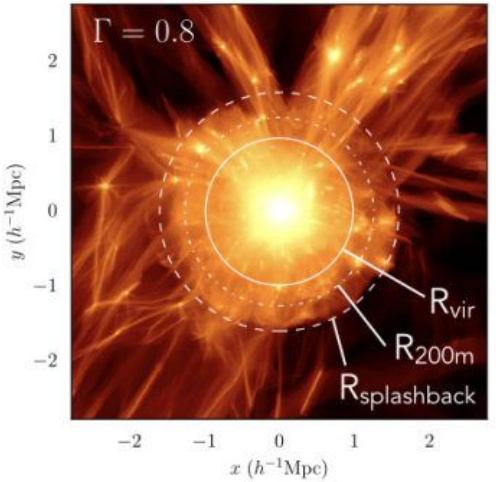
The dependence of Ae on environment definition



Yuan et. al. 2020b

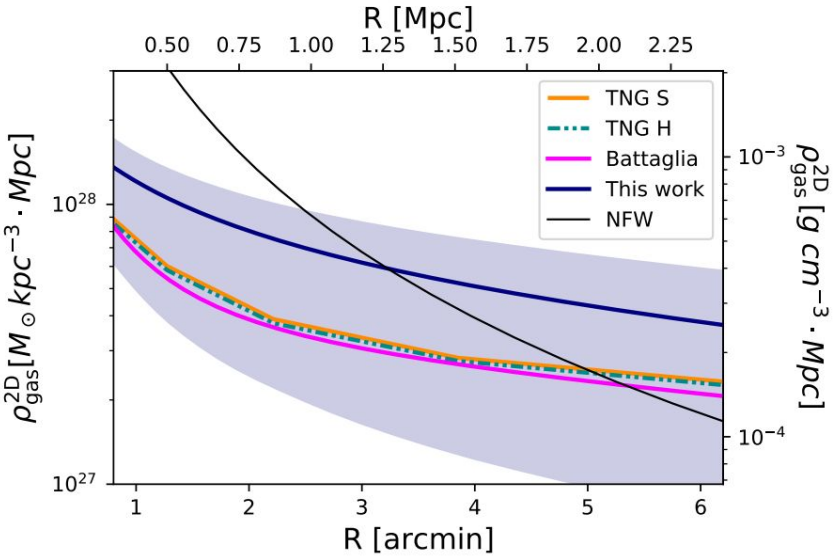
Splashback can explain Ae

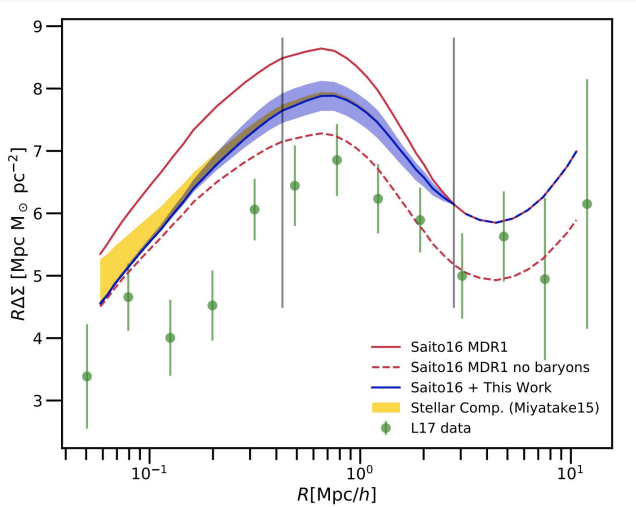




Credit: Benedikt Diemer

Other evidence of extended baryon profiles





Amodeo et al. 2020

Current work: AbacusHOD (DESI-HOD)

- Highly Efficient
  - **Performance**: 80ms/tracer to populate a 2Gpc box on a 32 core desktop. Scalable to more cores and nodes.
  - **Optimizations**:
    - Mass-dependent subsampling.
    - Preloading halo and particle data in memory.
    - Memory in-place implementation with numba parallel.
- Feature Rich
  - HOD extensions: secondary biases, velocity biases, satellite distribution flexibilities, RSD.
  - Interface with Abacus merger tree outputs.
  - Includes fast (0.1 second) 2-point calculators.
- Multi-tracer (LRG, ELG, QSO)
  - ELG/QSO HOD based on Shadab's eBOSS fits.
  - Performance is 2x slower when ELG/QSO is enabled.

To dos

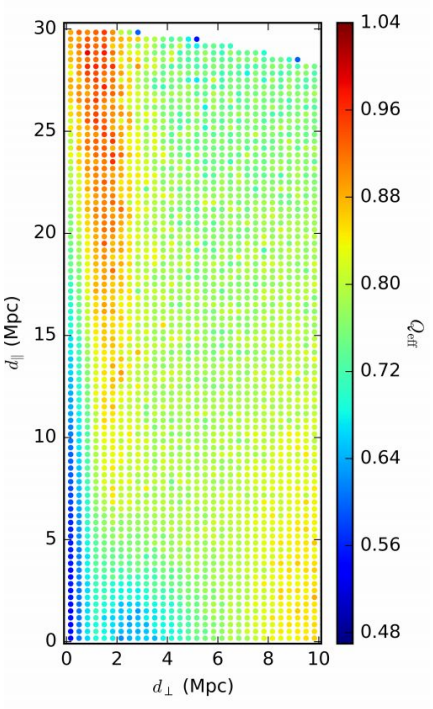
- Incorporating extended baryon profiles:
  - Leveraging the merger tree to re-merge splashback halos.
  - Develop a generalized NFW profile feature?

Applications

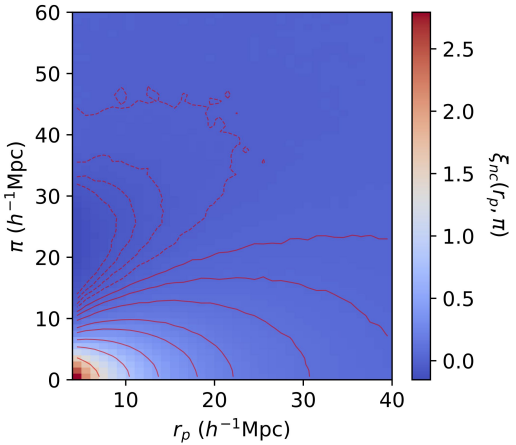
- DESI HOD fits (1% and then full sample) and cosmology+HOD analysis.
- Abacus Emulator:
  - 155 different cosmologies.
  - Marginalize over the extended HOD space and emulate the Bayesian evidence.
- More robust assembly bias fits. (DESI? Novel statistics?)

Novel statistics?

- Squeezed 3PCF (Yuan et al 2017, 2018).



- SZ-marked group correlation function (work in progress).

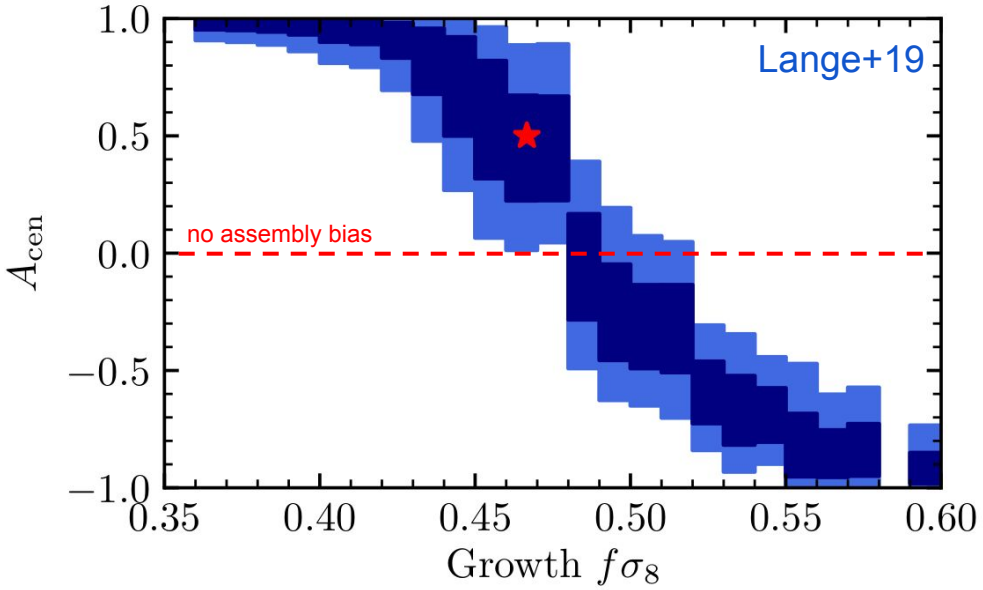


Summary

- We find strong evidence for extended HOD models with secondary biases.
- The corresponding lensing prediction is significantly more consistent with data than previous predictions.
- Redshift-space clustering offers a lot more information.
- The secondary biases have interesting implications ranging from dark matter structure, to CGM, to GR.
- Currently developing a fast extended HOD that will enable robust HOD and cosmology analyses, especially within DESI.

Credit: Johannes Lange

A negative detection of A



The concentration-based assembly bias is degenerate with sigma_8.