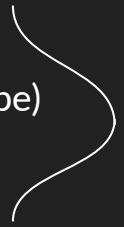
HOD in stages IV surveys

Romain Paviot

Context

Galaxy surveys in cosmology:

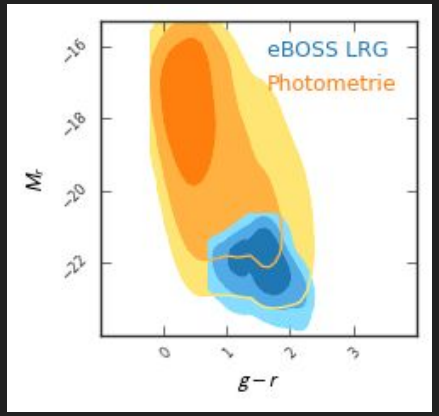
- **Photometric survey :** Angular information + photo-z
  → angular 3x2pt (correlation position-position, position-shape,shape-shape)
- **Spectroscopic survey** from target selection : Accurate-z
  → 2D clustering



Cosmological parameters: Analytical models / Forward modelling

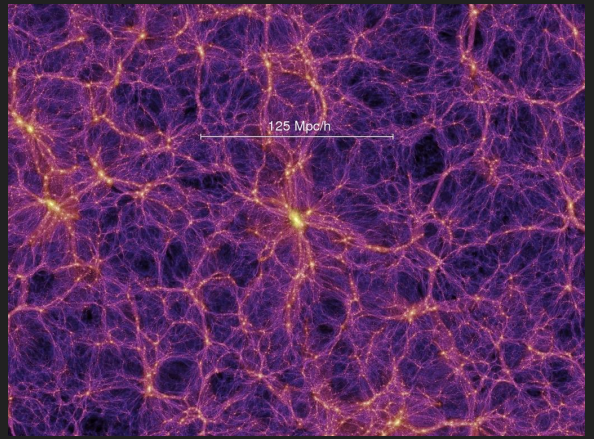
We need precise galaxy-halo connections model: perturbative (bias expansions) and empirical models.

Halo occupation and galaxies



Hydrodynamical simulations : Too much expensive to run
We need :

- N-Body simulation with large volume
- Empirical models which describe the galaxy-halo connection

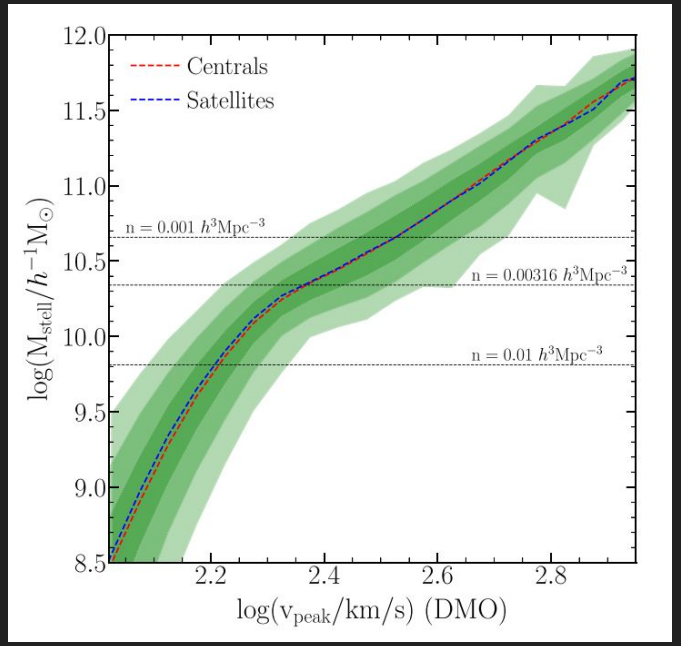


The Millennium simulation

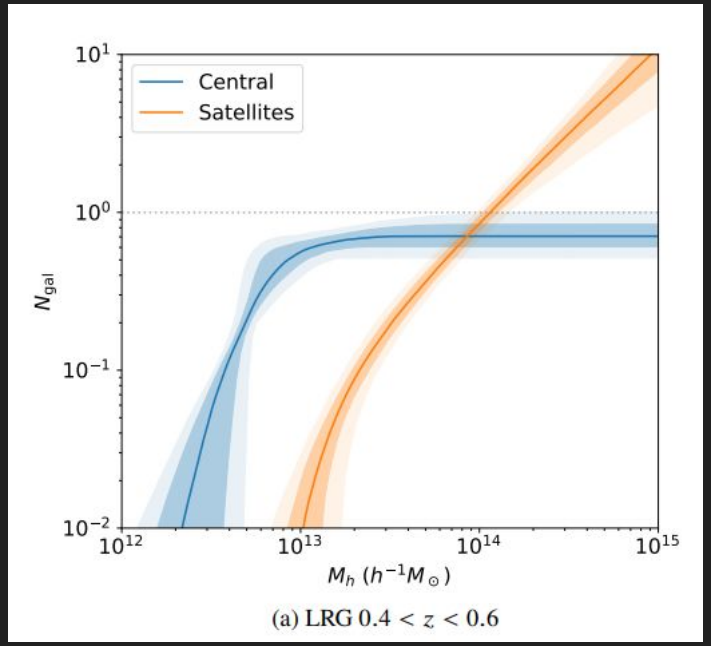
Two approaches :

- Reproduce galaxy observables (Luminosity, COLOURS, b(M))
  surveys: KiDS, DES, Euclid photometry -> DIFFICULT
- Reproduce galaxy clustering of a specific population
  surveys : (e)BOSS, DESI, Euclid spectroscopy -> EASIER

Halo Occupation Distribution (HOD) model and SubHalo Abundance Matching (SHAM)



Contreras +21

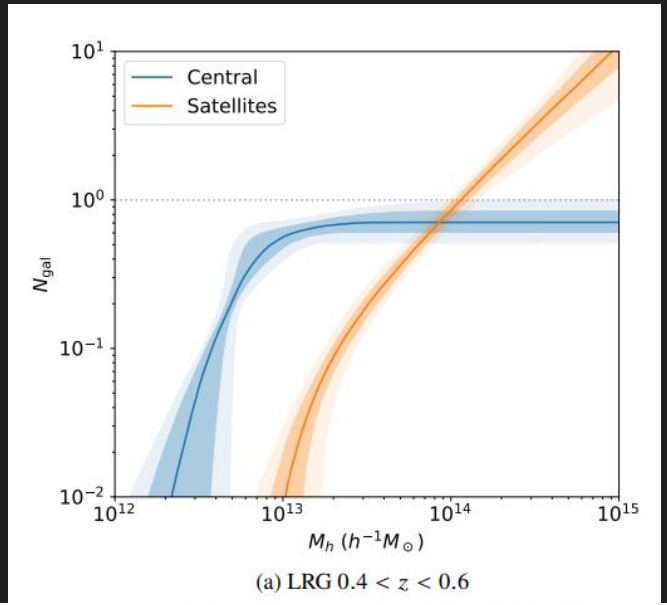


Yuan +23

eSHAM : relates galaxy properties to halo properties (formation time, vpeak)

HOD : populate dark-matter halo given some probability distribution of occupation

Common HOD assumption



(a) LRG $0.4 < z < 0.6$

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

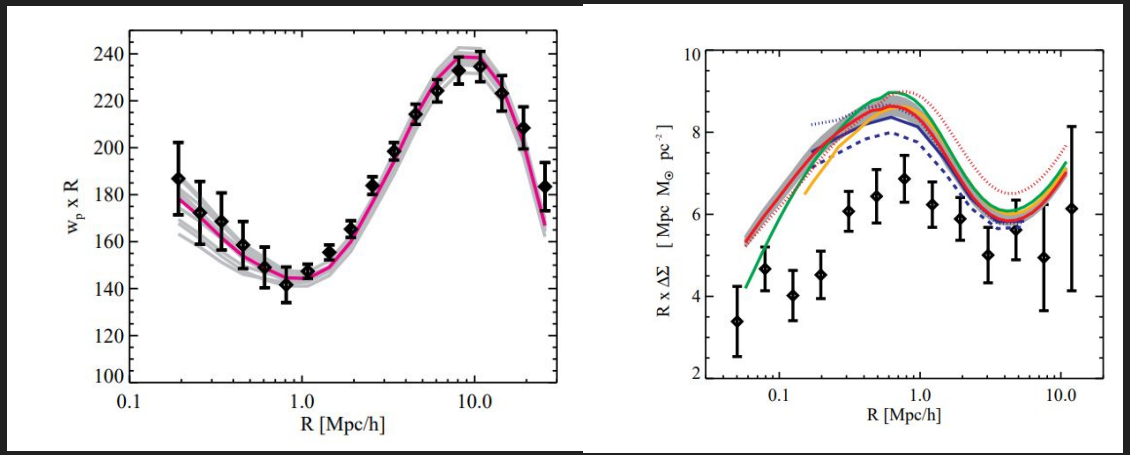
$$\langle N_s(M_h)\rangle = \langle N_c(M_h)\rangle \left(\frac{M_h}{M_1}\right)^{\alpha}$$

Zheng +07

- Halo occupation depends solely on halo mass
- Central are placed at halo centre of potential
- Satellites follow a Poisson distribution
- Satellites radial profiles follow standard spherical NFW profiles (potentially rescaled)
- Satellite velocities draw from DM velocity dispersion (+ eventual infall velocities)
- No baryonic feedback taken into account

Halo Occupation Distribution in spec-z surveys

HOD have been widely used in BOSS (White +10, Zhai +17, Avila +20) and DESI (Yuan +22, Rocher +23) and reproduce well projected and 2D clustering. However, it has found in Leauthaud +16 to over-predict galaxy-matter cross correlations.



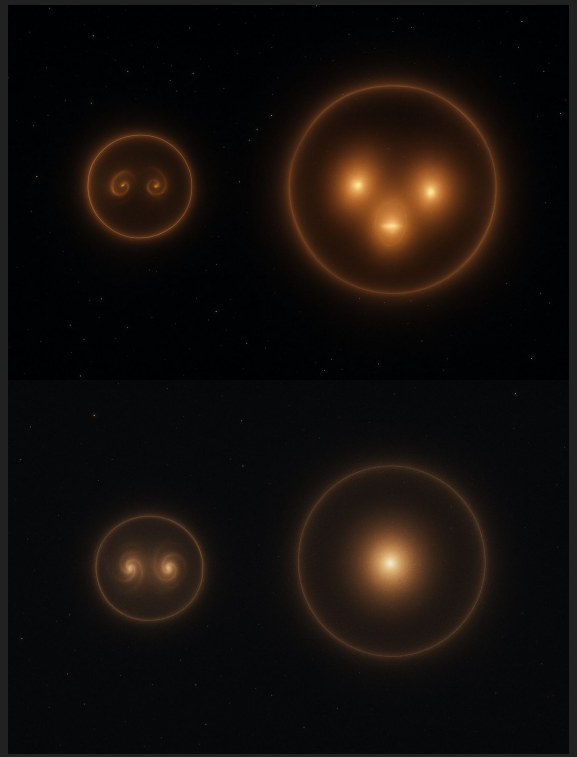
Leauthaud +16

HOD had to be improve.
(only for GGL!)

Galaxy and halo assembly bias

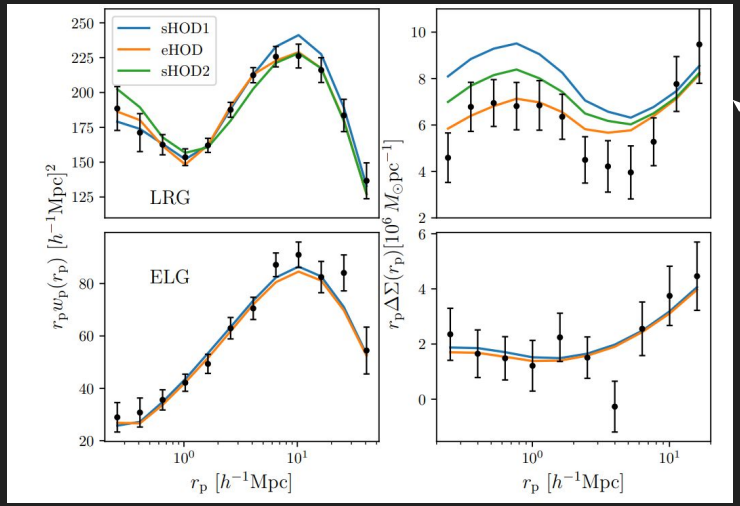
halo A.B : Halo occupation/clustering depends on properties other than halo mass.
galaxy A.B : Galaxy clustering depends on properties other than halo mass.
Properties: halo formation time, local environments, local anistropries (Zehavi +18, Hadzhiyska +22a,b)

Some examples :
- Early-forming halos are more likely to host central galaxies at lower halo mass.
- At fixed halo mass, older halos tend to host more massive galaxies (opposite for satellites)
-Star-forming galaxies formation is anti-correlated with the presence of nearby massive clusters.



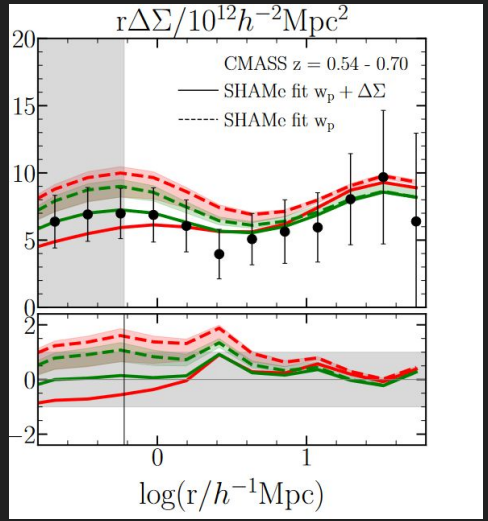
Galaxy and halo assembly bias

HOD model of GGL signal can be improved by implementing assembly bias.



Paviot +24

Can be improved by modifying satellite occupation models



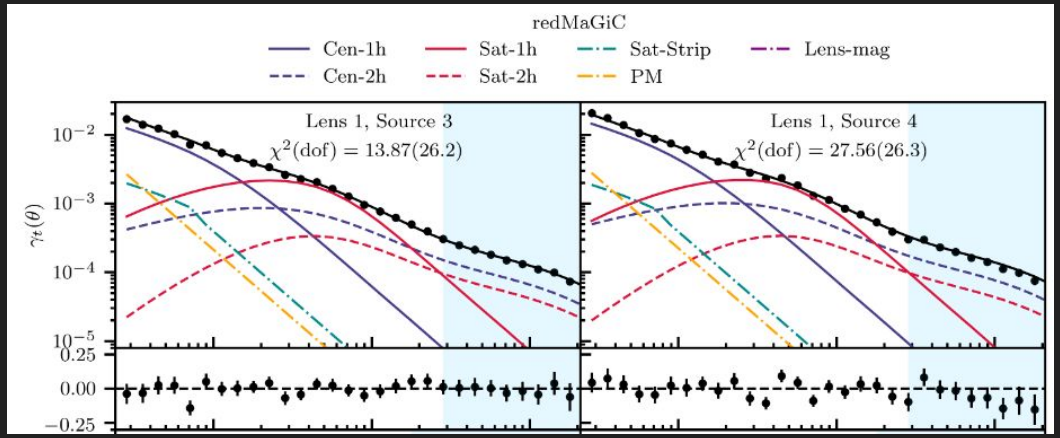
eSHAM, Contreras+23b

$$\langle N_s(M_h)\rangle = \langle N_c(M_h)\rangle \left(\frac{M_h}{M_1}\right)^{\alpha}$$

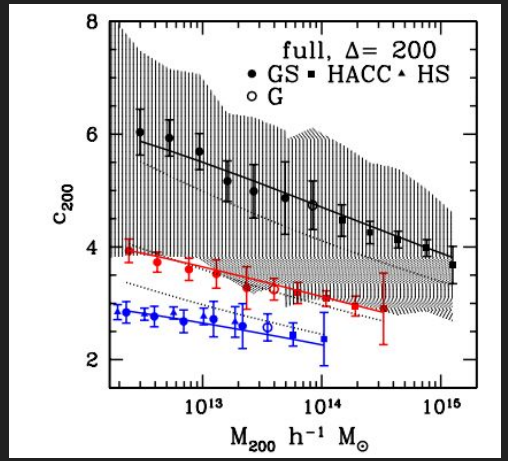
No baryonic feedback in each model!

Halo Occupation Distribution in photo-z surveys

HOD models has been extensively used in Stage III such as KiDS or DES as a semi-analytical approach to reproduce galaxies statistics (GCs, GGL, IA)



DES: Zacharegkas +22



Bhattacharya +13

Example of GGL fit of DES RedMaGIC sample. To perform this fit, one needs to asume : a HOD, b(M), c(M) to integrate over halo mass

$$P_{\rm gm}^{\rm c1h}(k,z) = \frac{1}{\rho_m \bar{n}_g} \int dM_h \frac{dn}{dM_h} \times M_h \langle N_c(M_h) \rangle u_{\rm dm}(k|M_h)$$

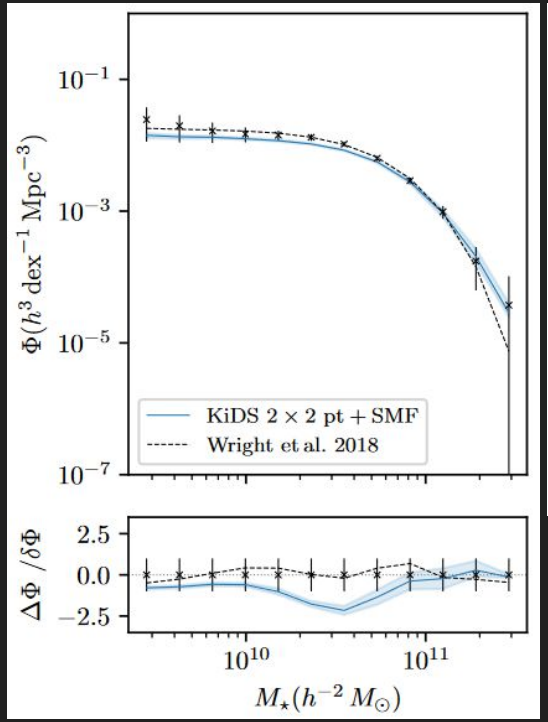
Halo Occupation Distribution in photo-z surveys

$$\Phi(M_\star|M) = \Phi_c(M_\star|M) + \Phi_s(M_\star|M).$$

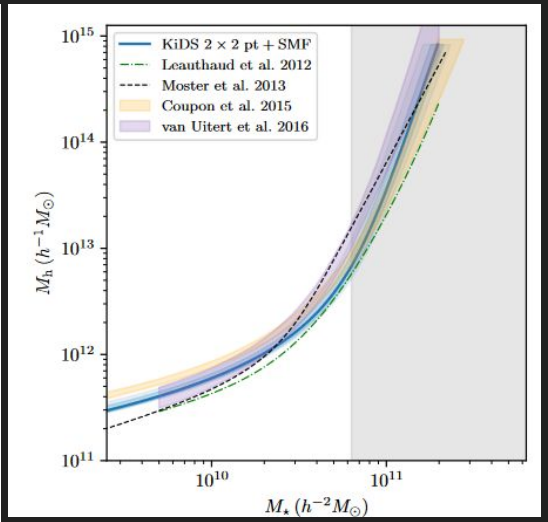
$$\Phi_c(M_\star|M) = \frac{1}{\sqrt{2\pi}\ln(10)\,\sigma_c M_\star}\exp\left[-\frac{\log(M_\star/M_c^*)^2}{2\,\sigma_c^2}\right]$$

$$\Phi_s(M_\star|M) = \frac{\phi_s^*}{M_s^*}\left(\frac{M_\star}{M_s^*}\right)^{\alpha_s}\exp\left[-\left(\frac{M_\star}{M_s^*}\right)^2\right]$$

Dvornik +23: Conditional stellar mass function

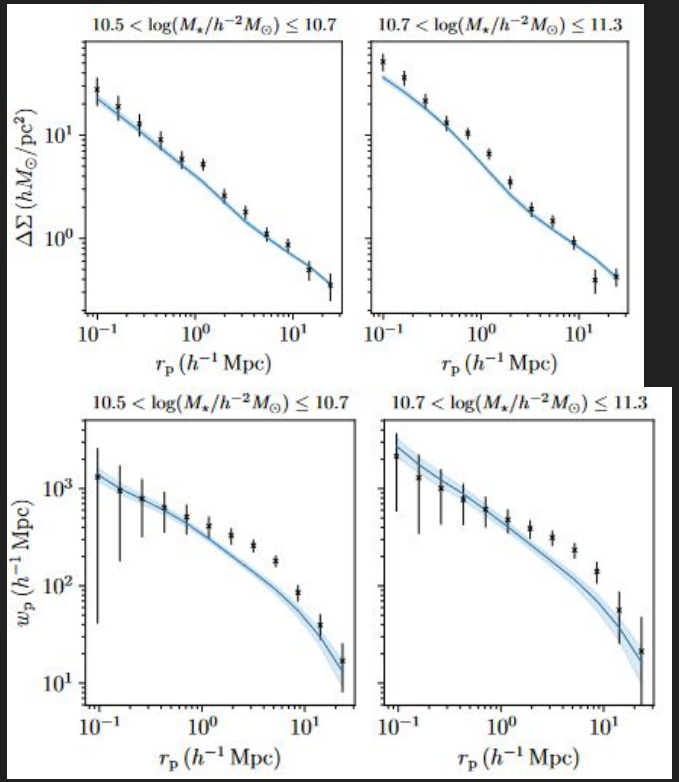


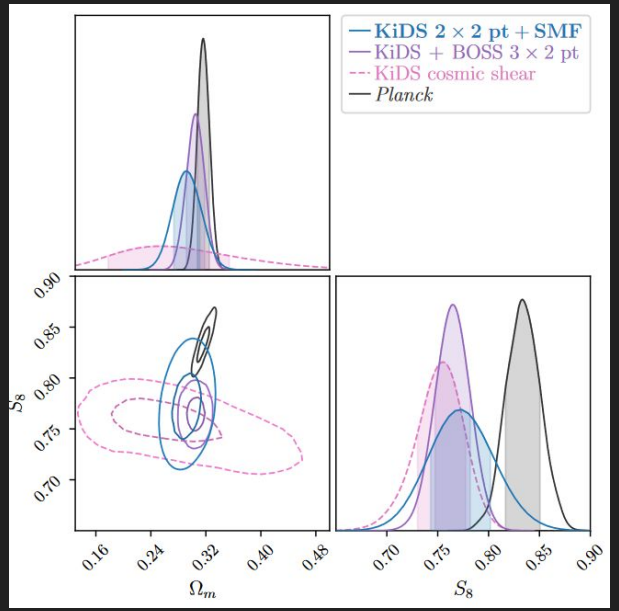
Dvornik +23



Dvornik +23

Halo Occupation Distribution in photo-z surveys

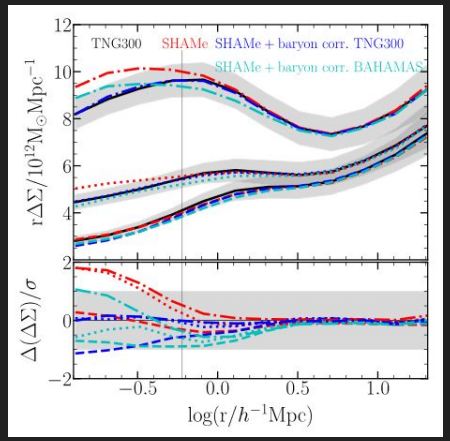




Dvornik +23

Can be used to constraints cosmology!

Where are the bayrons??

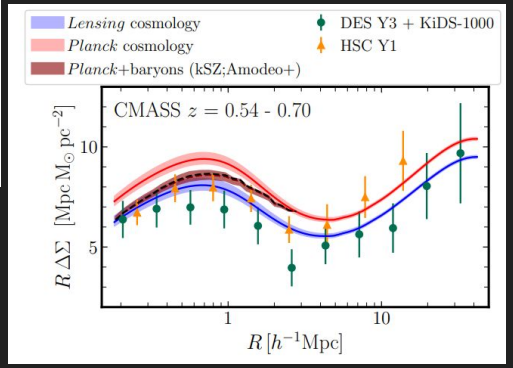


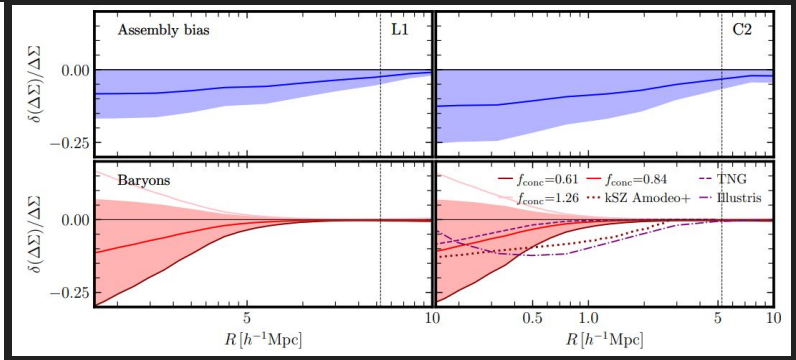
eSHAM, Contreras+23b

In Zacharegkas +22: No baryon, NFW halo rescaling

In Dvornik +23: No baryon, NFW halo rescaling

"Our adoption here of separate concentration-mass relations for dark matter haloes and satellite galaxies provides enough flexibility in the model to capture the uncertain impact of baryon feedback(Debackere +21, Amon+22)"

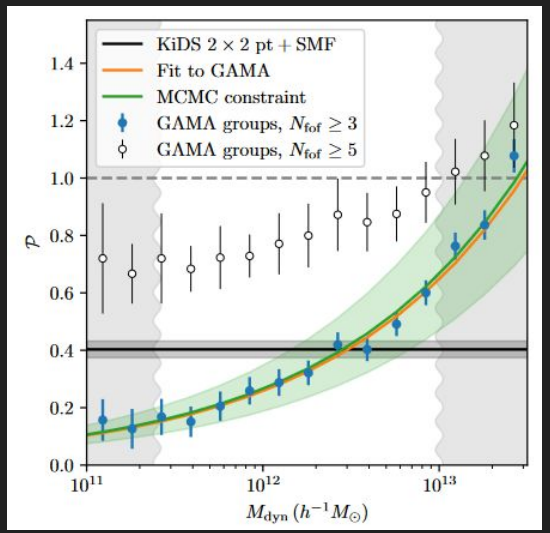




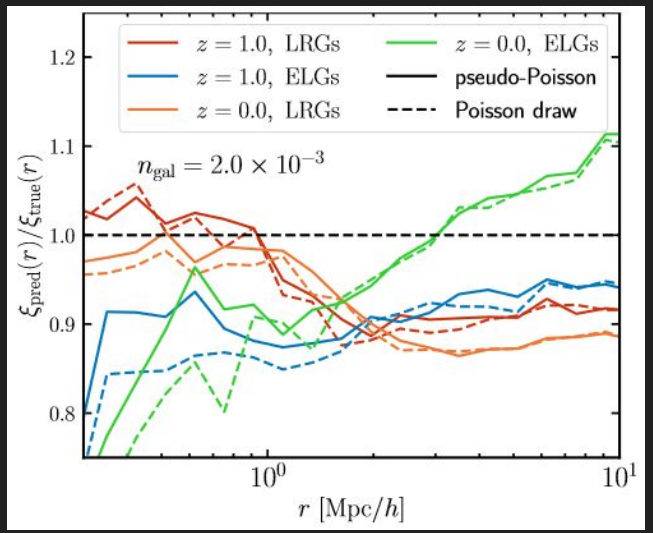
Amon +22, Amodeo +21

From Debackere +21 : Cluster counts: Weak lensing mass unbiased if concentration left free, but biased cosmological parameters.

or not?



Dvornik +23



Hadzhiyska +23b

Traditional Poisson distribution for satellites is incorrect.

Sub-poisson at low halo mass/ super-poisson at high halo mass.

Can impact clustering predictions of ELGs.

How can we model all sky?

The Mice and *Euclid* Flagship simulation (Fosalba +13, Castender +23)
The diffstuff project (Diffmah; Hearin +21, Diffstar; Alarcon +23, DSPS; Hearin +23)
Goal : Reproduce every observables at once.

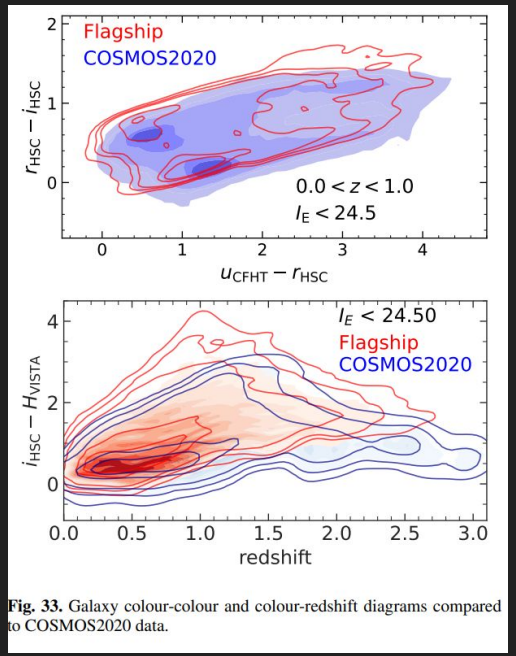
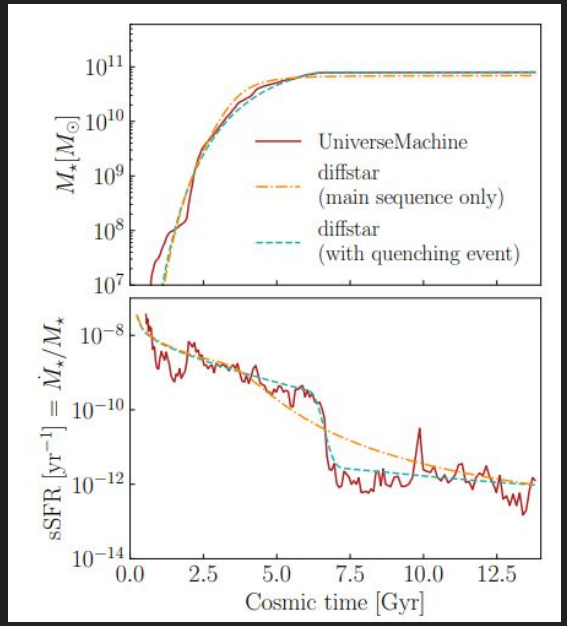


Fig. 33. Galaxy colour-colour and colour-redshift diagrams compared to COSMOS2020 data.



Conclusion: What's ahead of us?

HOD of colour-magnitude selected samples have intensively used and tested with BOSS/DESI data.

⟶ Robust method to test theoretical GCs analysis pipelines.

This apply as well to GGL, and to some extend IA (Van Alfven +24)

For magnitude limits surveys, the picture is not the same for now:
For ex, Dvornik +23 model *has not been* tested against sims.
Need accurate luminosity function of galaxies + accurate predictions of GCs/GGL/IA in colour-magnitude space.

Path toward this: *Euclid* Flagship simulation (Castander +24) and the diffstuff program.
Where Stage IV surveys will help: Common coverage between *Euclid*/LSST/DESI will provide un-precedent statistical precision to improve calibrations.