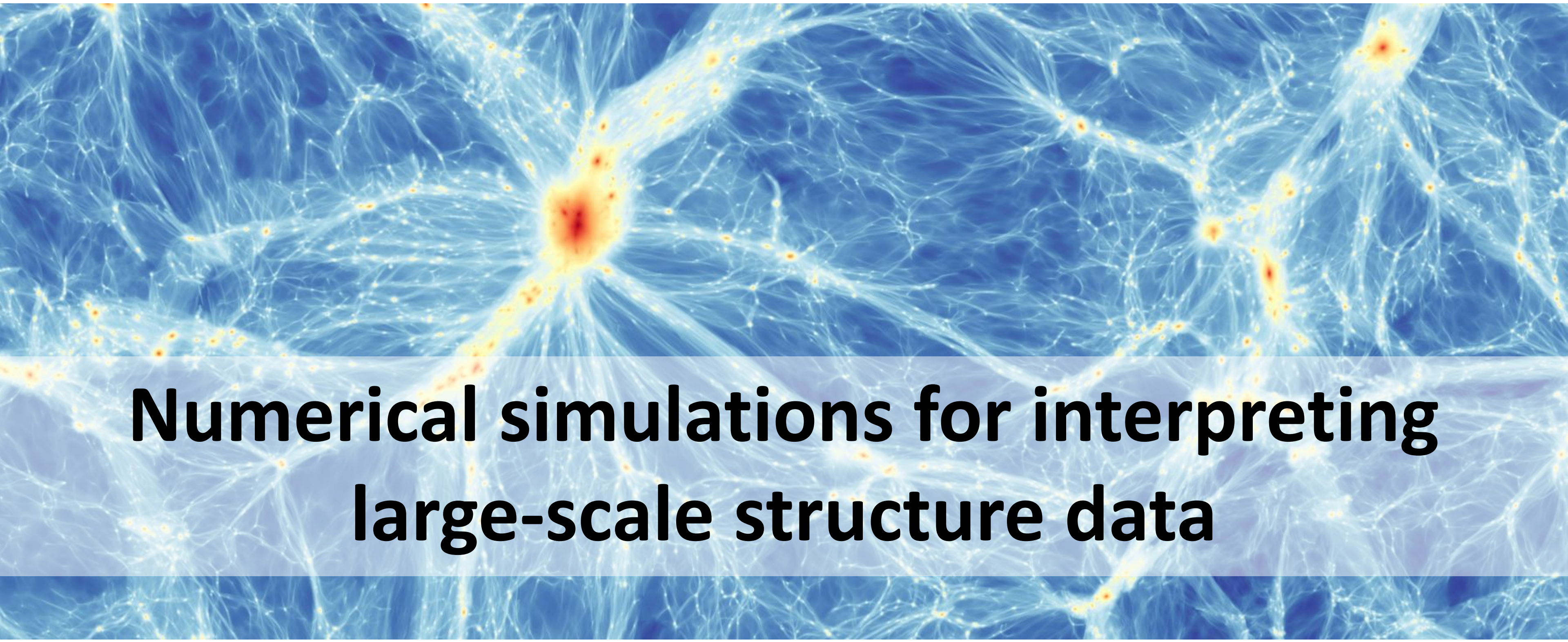


Numerical simulations for interpreting large-scale structure data

Raul E. Angulo

Donostia International Physics Center

Alexander von Humboldt Stiftung/Foundation

DIPC

ikerbasque Basque Foundation for Science

erc

Donostia International Physics Center

- **3 staff**
- **6 PhD**
- **6 Postdocs**

- **Galaxy formation**
- **Black Holes & AGN/Quasars**
- **Star and clusters formation**
- **Large-scale structure**

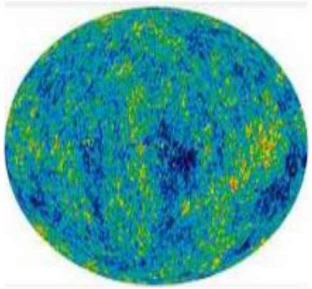
Our current understanding of structure formation in the Universe:

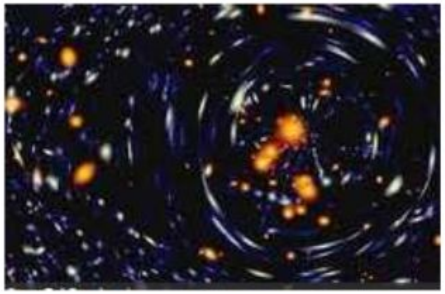
General Relativity

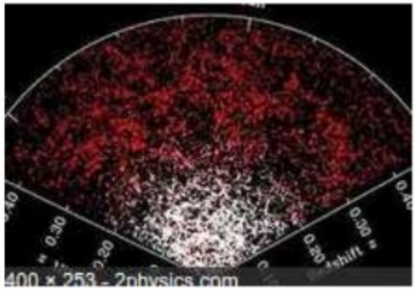
Dark Matter

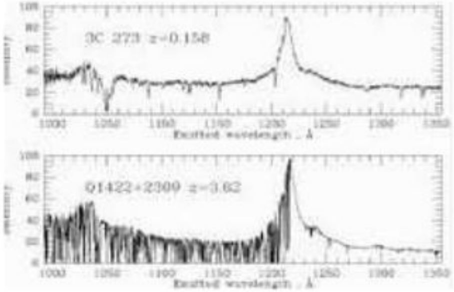
Dark Energy

Inflation









Despite the progress, there are questions about each of the fundamental ingredients of LCDM

Gravity Law

GR, Galileon, f(R)?

Accelerated Expansion

DE, coupled, w(z)?

Dark Matter

Temperature, Neutrinos, Quantum?

Initial Conditions Inflation

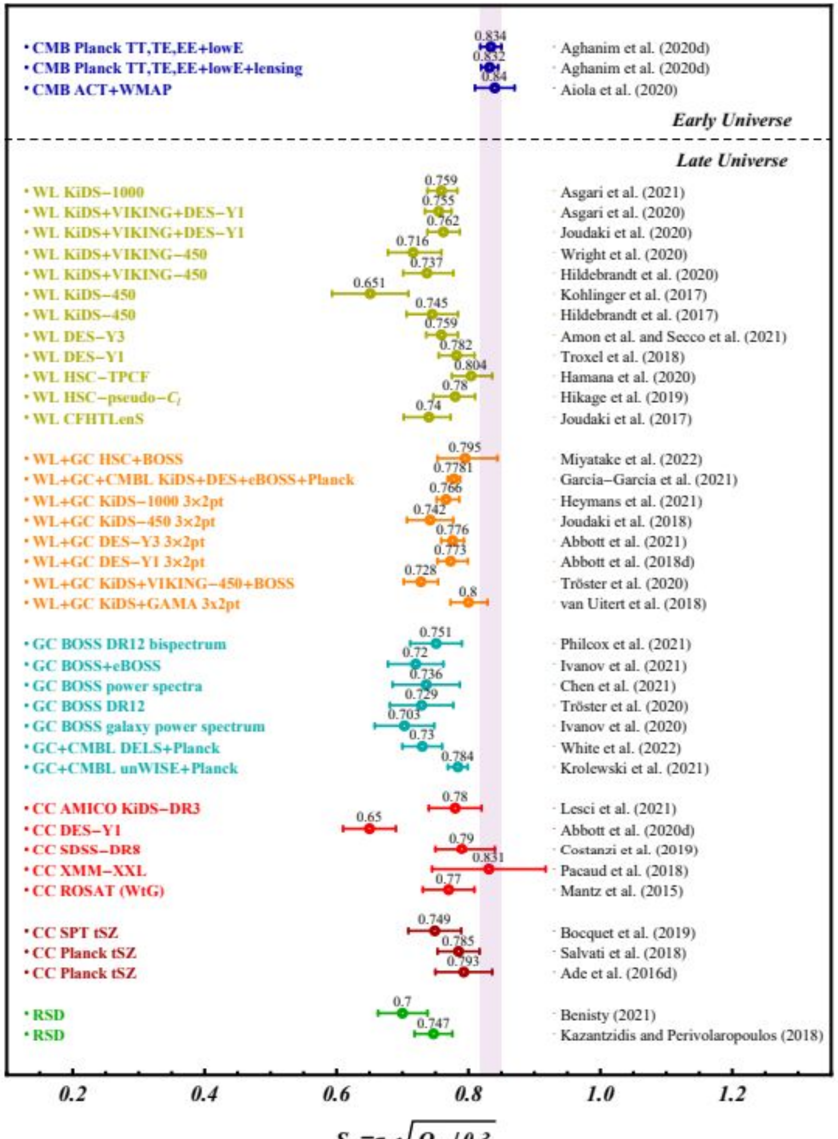
Single/multi-field?

No deviations have from the simplest LCDM model been detected so far

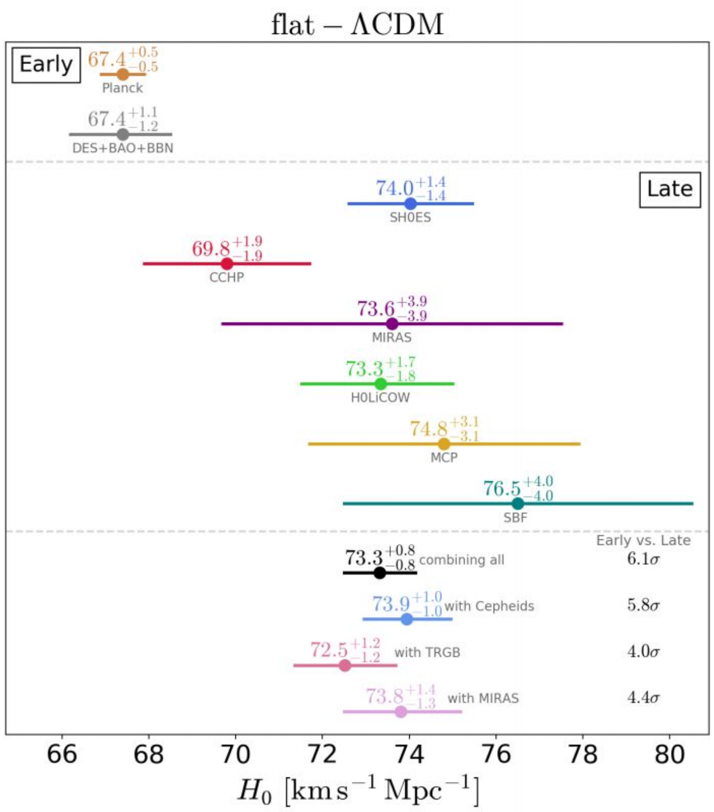
Ω_K	-0.0096 ± 0.0061
Σm_ν [eV]	< 0.241
$N_{\rm eff}$	$2.89^{+0.36}_{-0.38}$
$r_{0.002}$	< 0.101

Tensions between the low-z universe and CMB expectations

Structure Parameter



Hubble Parameter



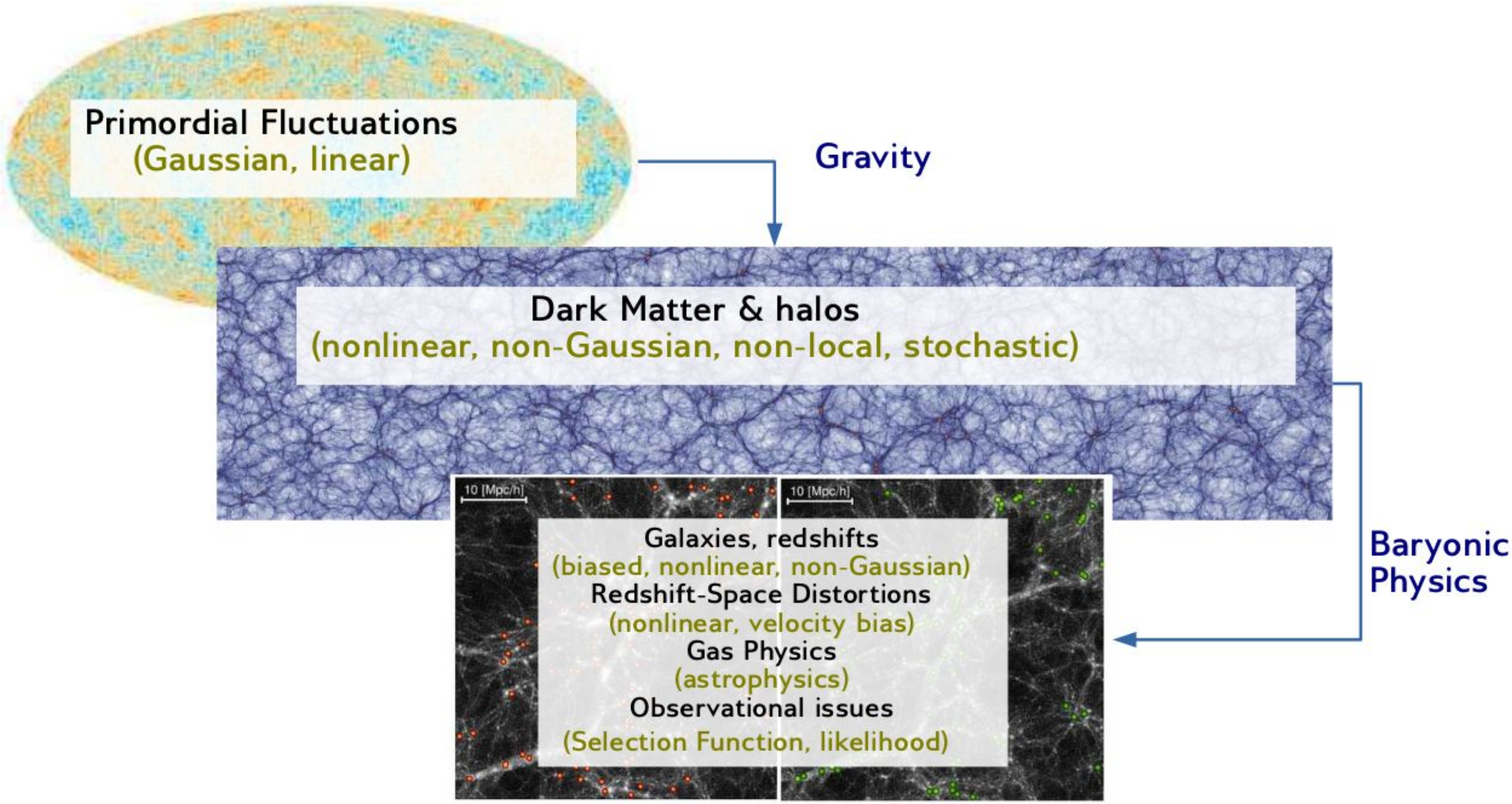
Verde et al 2019, Vagnozzi et al 2021

The era of precision cosmology for Large-scale structure data

2019-2030:

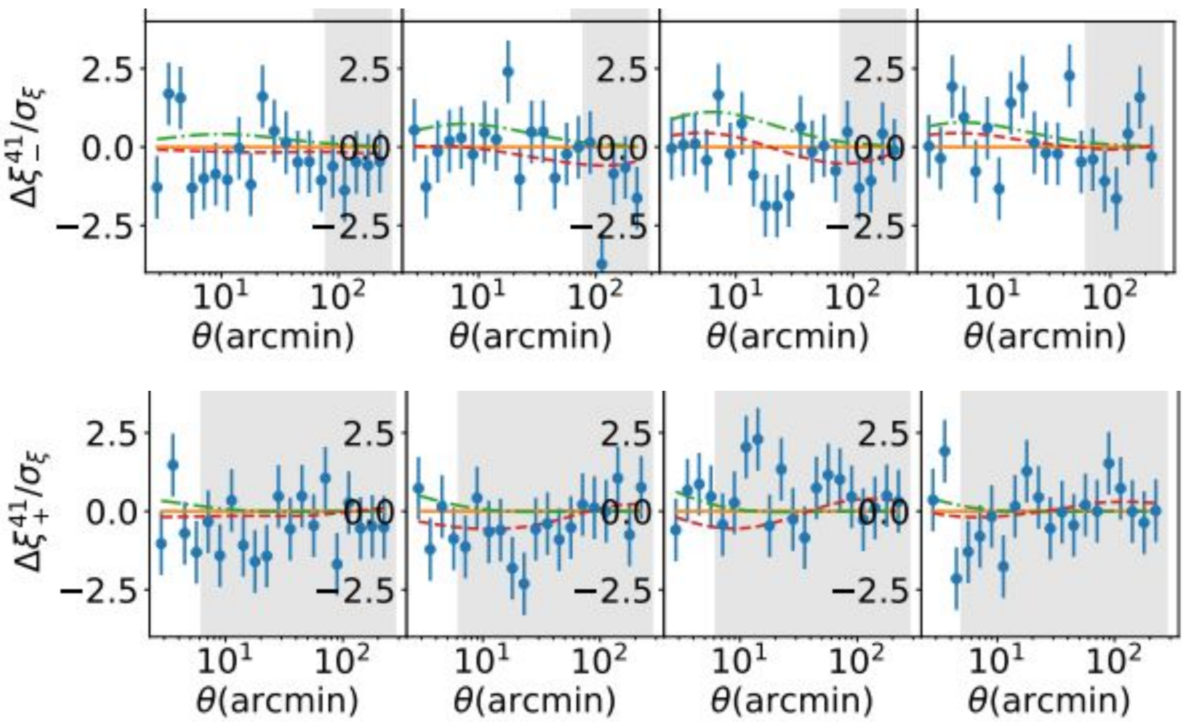
Lensing	Clustering
SZ-IM	Clusters

Exploiting current and future LSS data is a very hard problem.



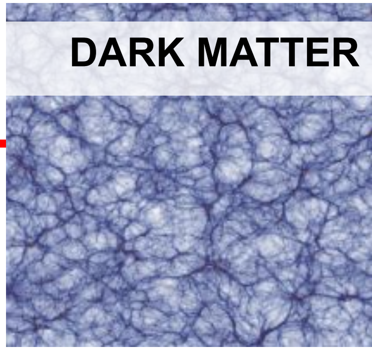
Cosmology from weak gravitational lensing

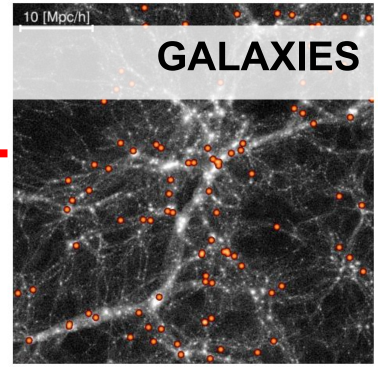
Current analyses are restricted by nonlinearities and gas physics

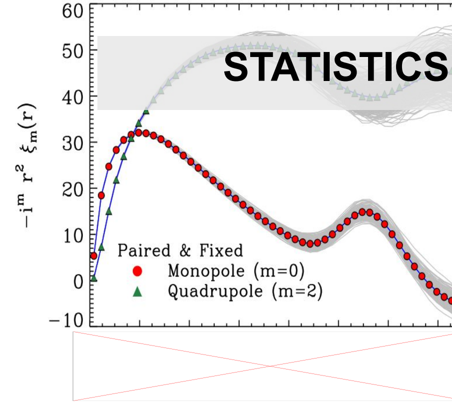


- Non linear matter power spectrum
- Galaxy Bias
- Baryonic effects
- Intrinsic Alignments

Chen+ 2022



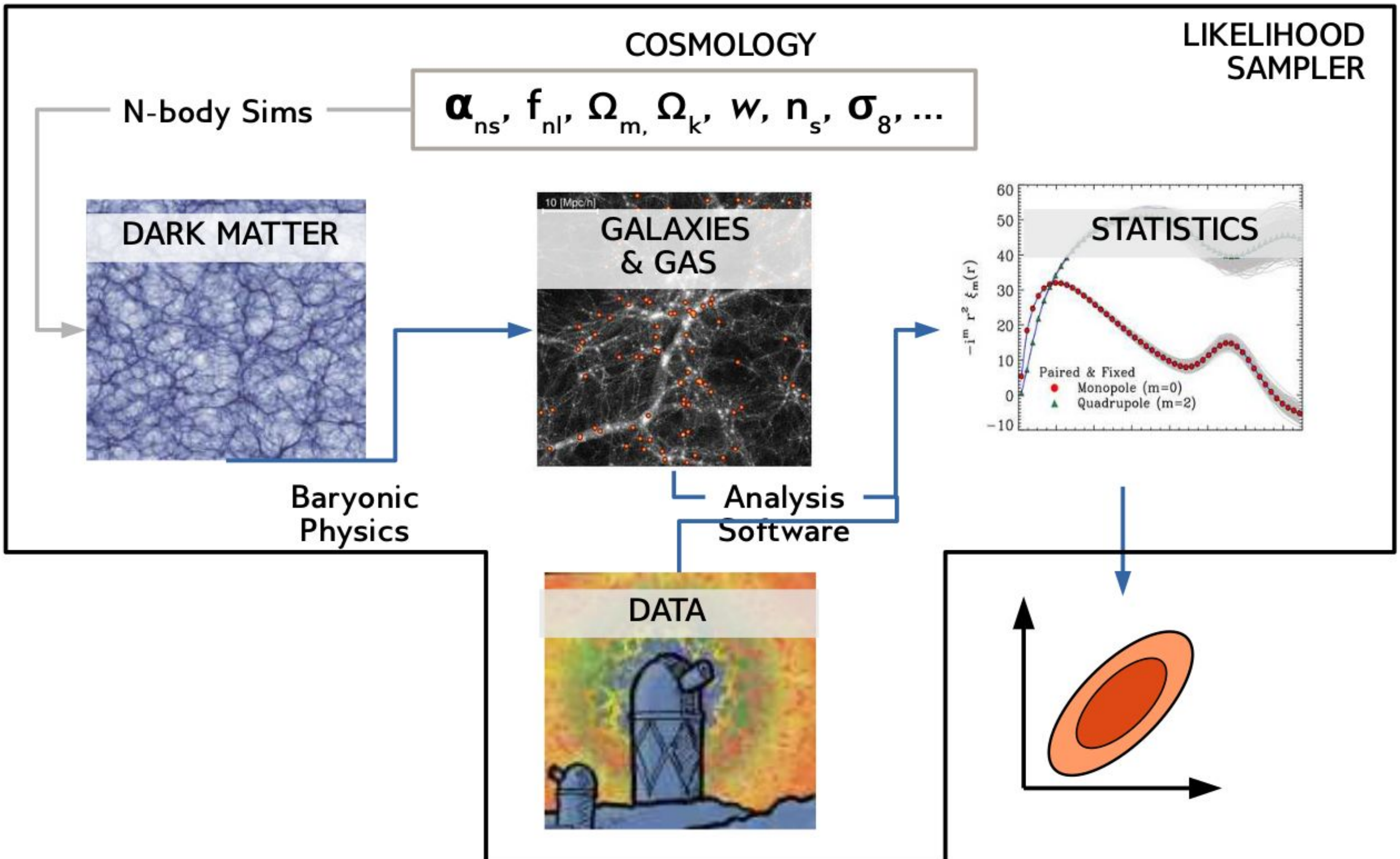




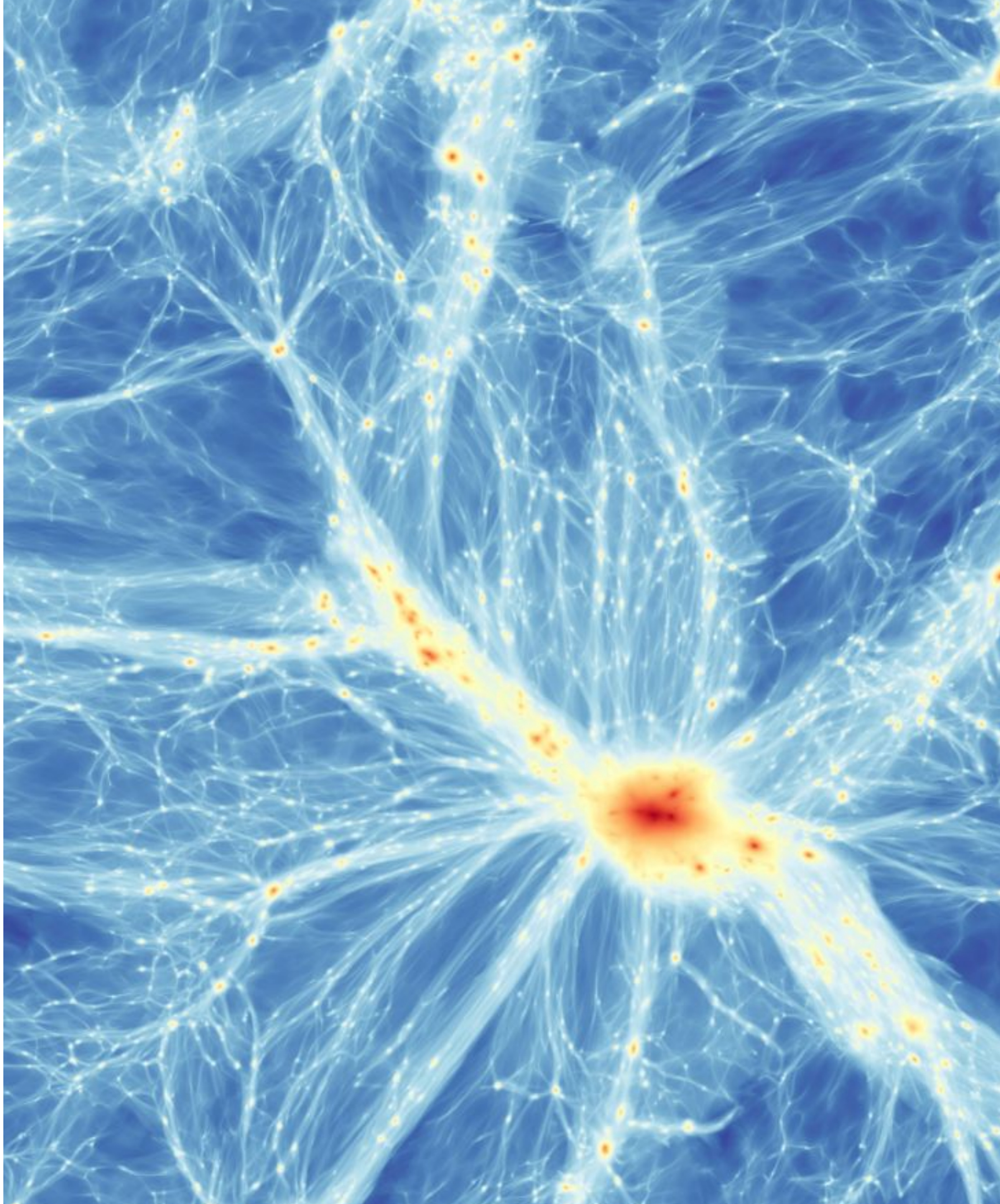
How do we unlock the full power of late-universe measurements?

Simulations as LSS theory

Employing precise and physical structure formation models



GRAVITY AND DARK MATTER FIELD



The Vlassov-Poisson (aka collsionless Boltzmann)

Simplest case:DM is initially smooth, cold, classical, and collisionless

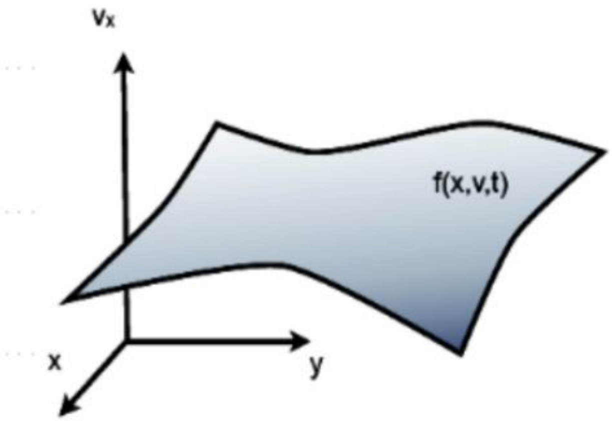
The Vlassov-Poisson Equation

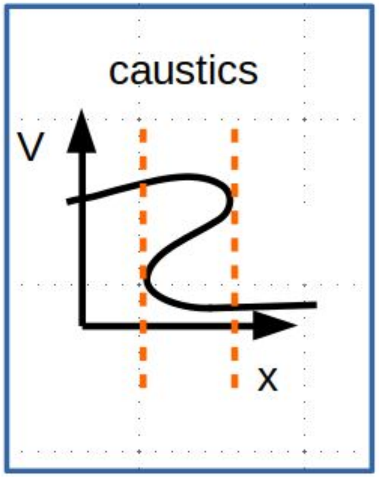
$$0 = \frac{\mathrm{d}f}{\mathrm{d}t} = \frac{\partial f}{\partial t} + \frac{\mathbf{v}}{a^2} \cdot \frac{\partial f}{\partial \mathbf{v}} - \frac{\partial f}{\partial \mathbf{x}} \frac{\partial \Phi}{\partial \mathbf{x}}$$

$$\nabla^2 \Phi = \frac{4\pi G}{a} \int f \mathrm{d}^3 v$$

CDM Sheet Properties

- → phase-space is conserved along characteristics
- → It can never tear
- → It can never intersect





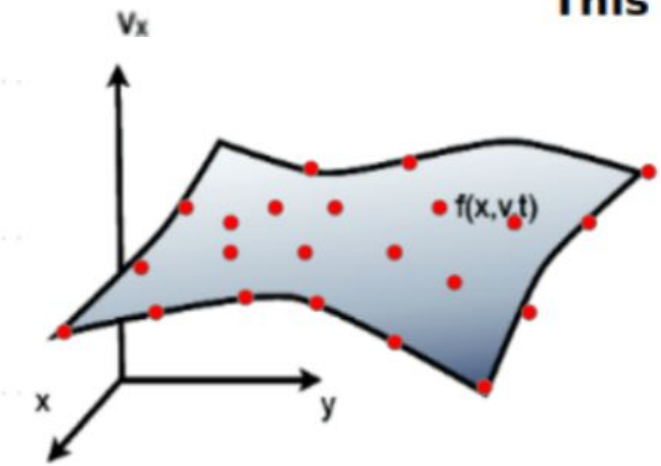
Solving Vlassov-Poisson via MC sampling

Currently, the most accurate solution is discretizing the distribution function

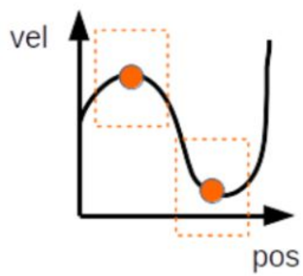
$$\dot{\mathbf{x}}_c = \frac{\mathbf{v}_c}{a^2}$$
$$\dot{\mathbf{v}}_c = -\nabla_{\mathbf{x}}\phi|_{\mathbf{x}_c}$$

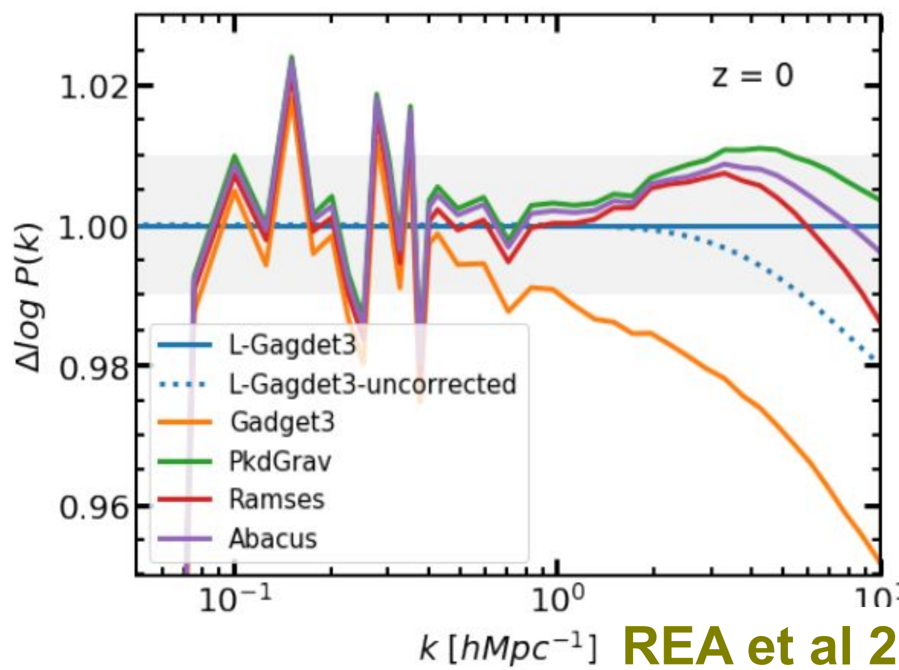
The solution yields the equation of motions of the Hamiltonian of classical mechanics

This is the correct solution as N goes to infinity



N-body simulation particle



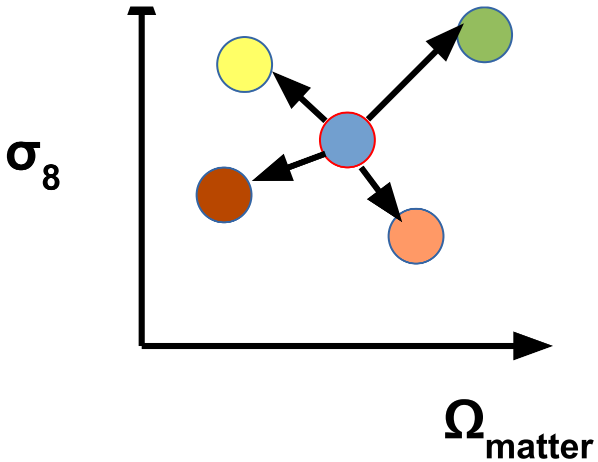


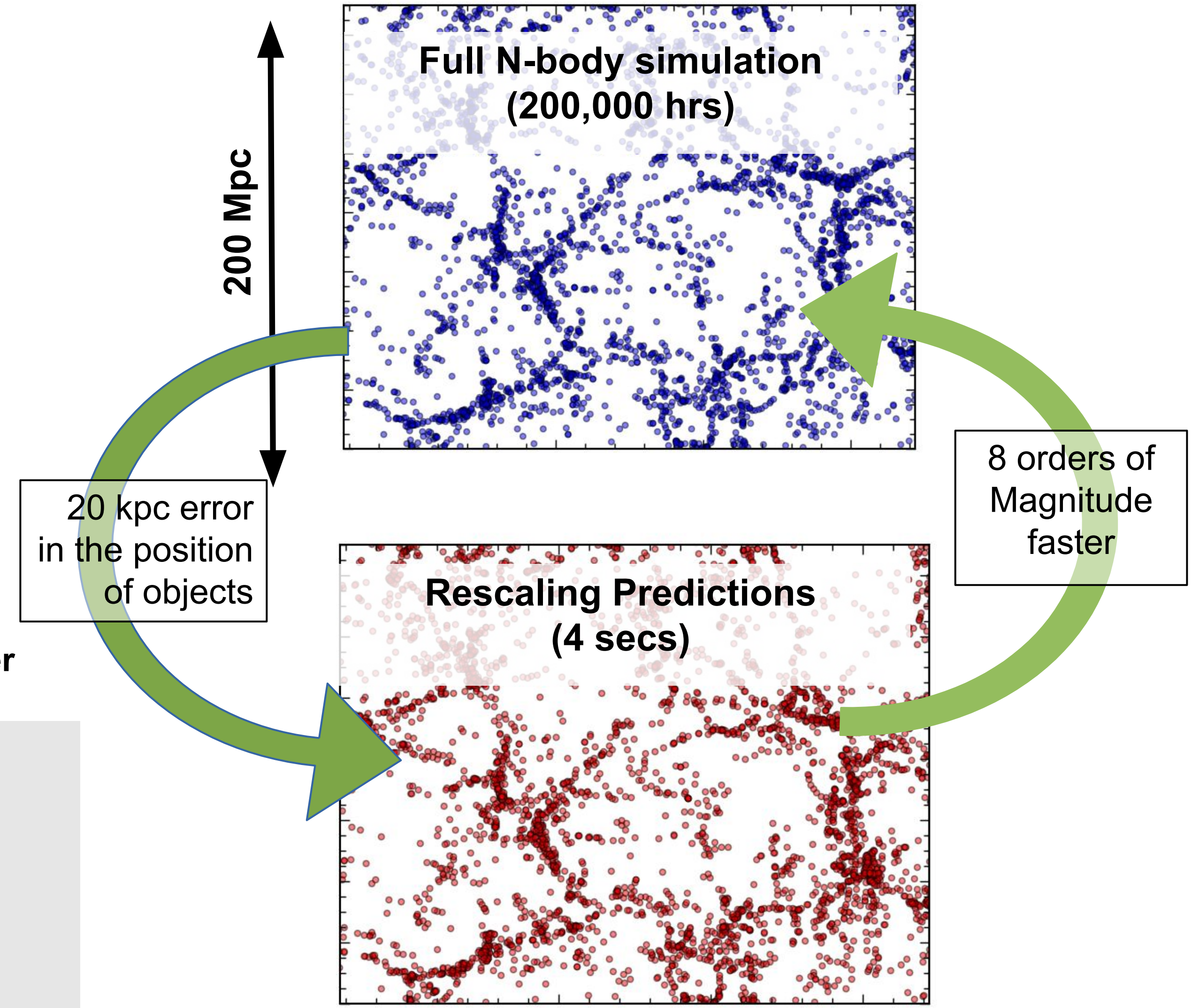
REA et al 2020

How to sample the cosmo-parameter space?

Rescaling modifies the outputs of simulation to mimic other cosmologies

Angulo & White (2010)
Angulo & Hilbert (2015)
Ruiz et al (2011)
Renneby, Hilbert, Angulo (2018)

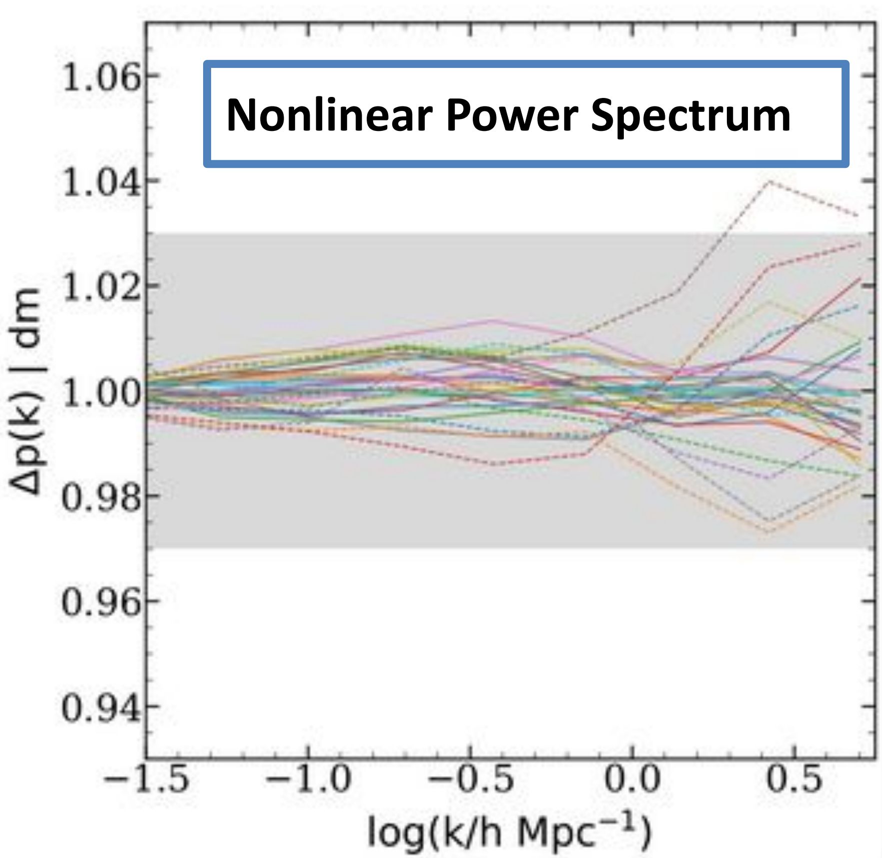




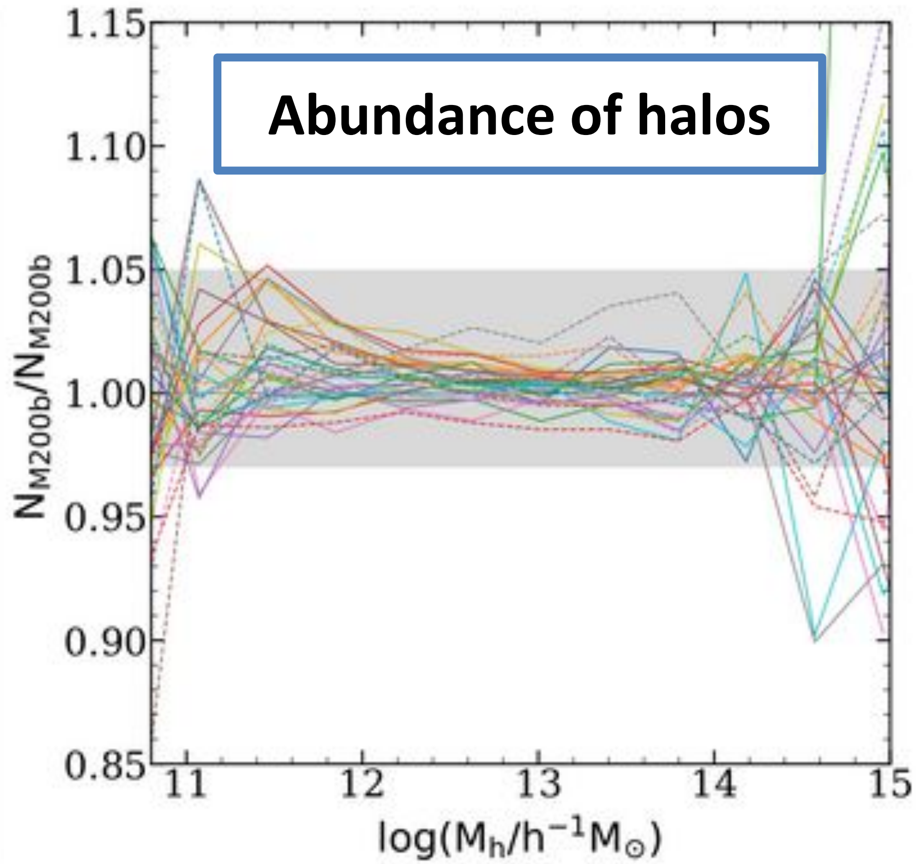
Validated for:

- real/redshift-space
- correlation function/power spectra
- 3-point correlation functions
- (sub)halo mass function
- abundance of voids
- different redshifts/cosmologies

Accuracy in predicting the nonlinear distribution of matter and collapsed objects



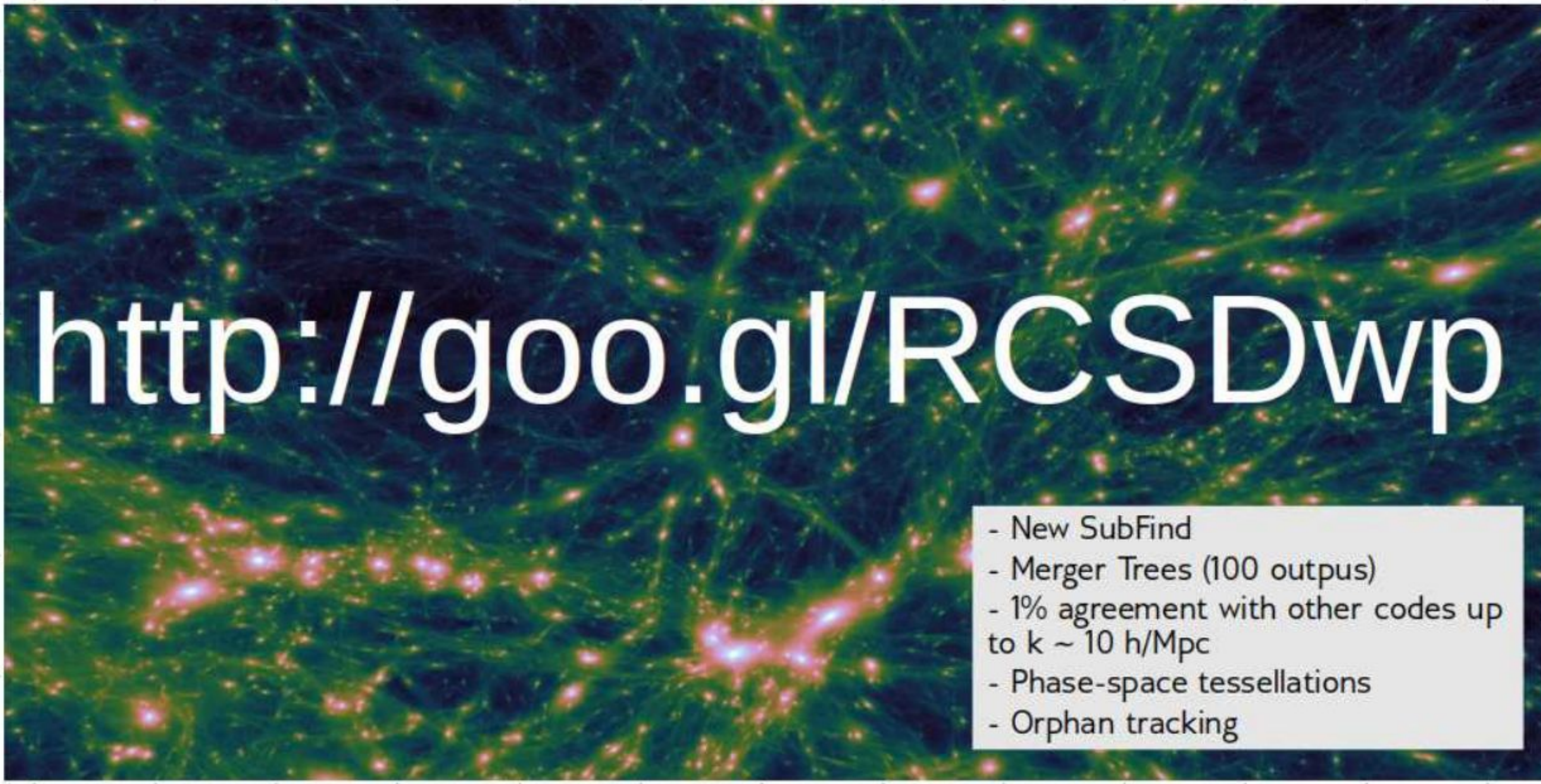
Contreras, Angulo, et al (2020)



Ondaro, Angulo, et al (2021)

Region of ~10-sigma around Planck's best fit values, including massive neutrinos and dynamical dark energy

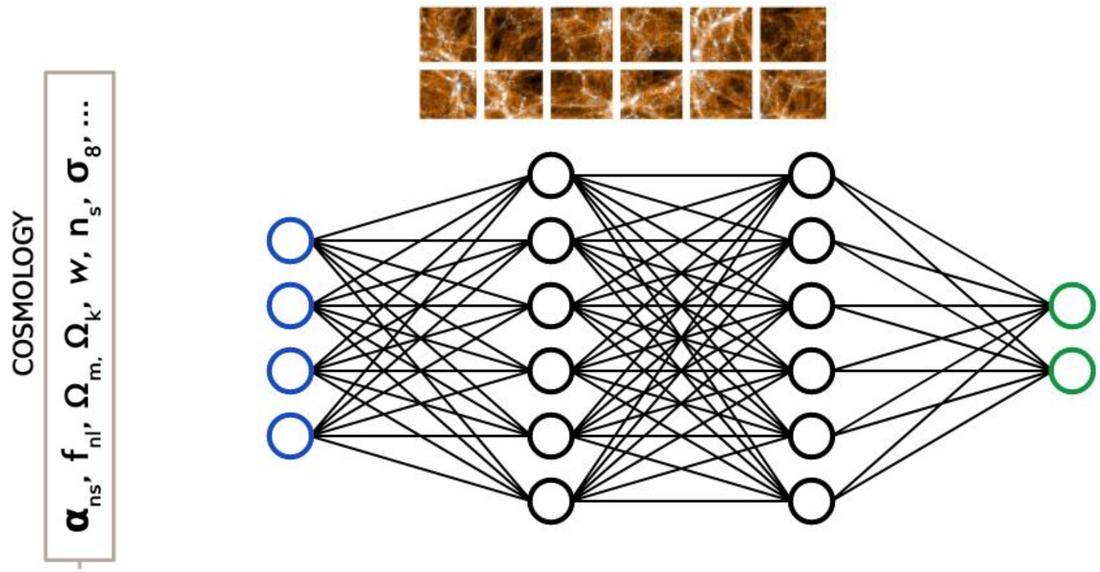
Bacco simulations: The challenge of predicting the nonlinear mass field at the 1%

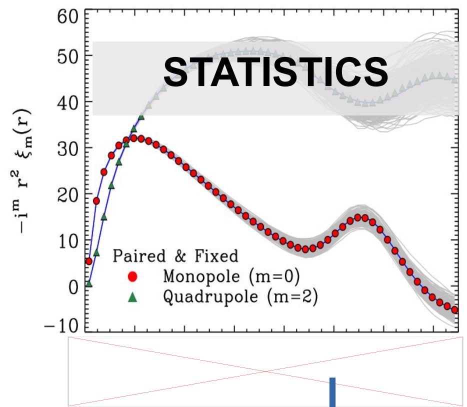


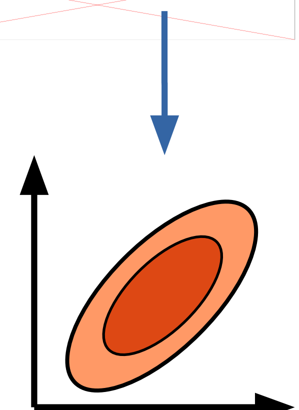
6 simulations: 2000 Mpc with 4320^3 particles; Paired-Fixed 2LPT ICs; 100 outputs; halos ~ 10^{10} Msun/h

Constraining Cosmology and Astrophysics

Feed-forward neural networks trained on 10,000 combinations of cosmology and astrophysics

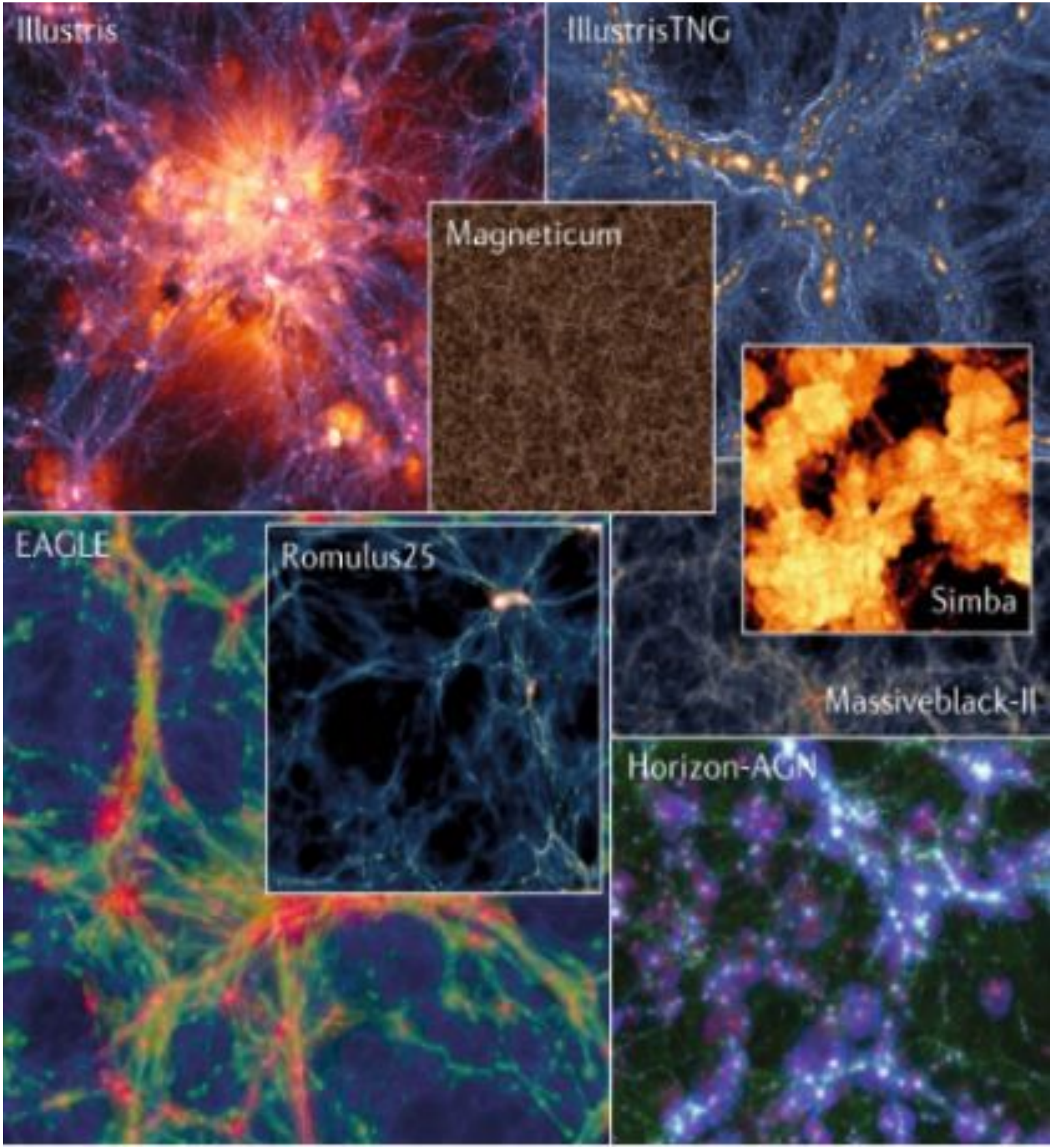




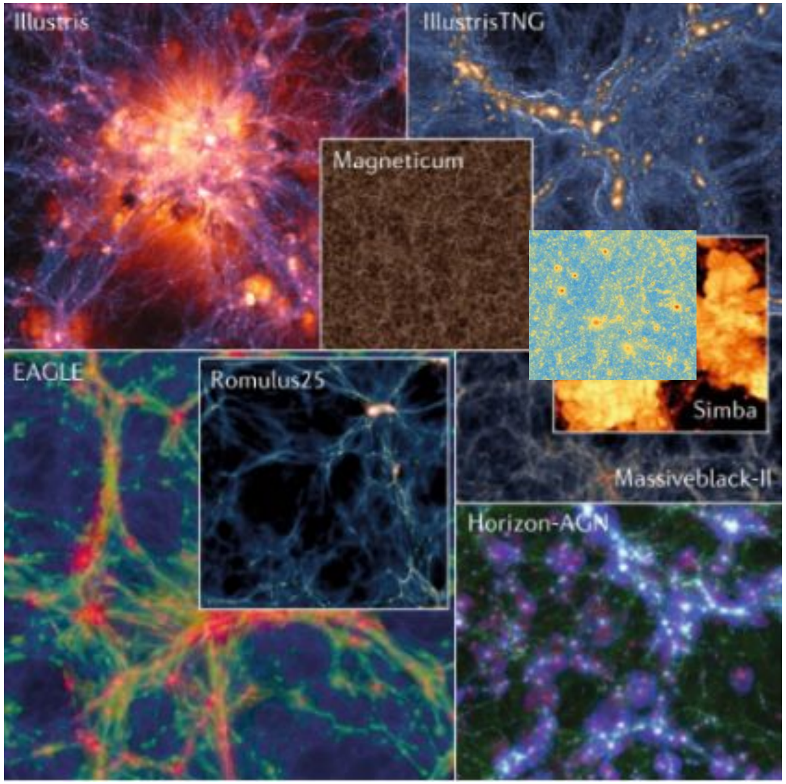


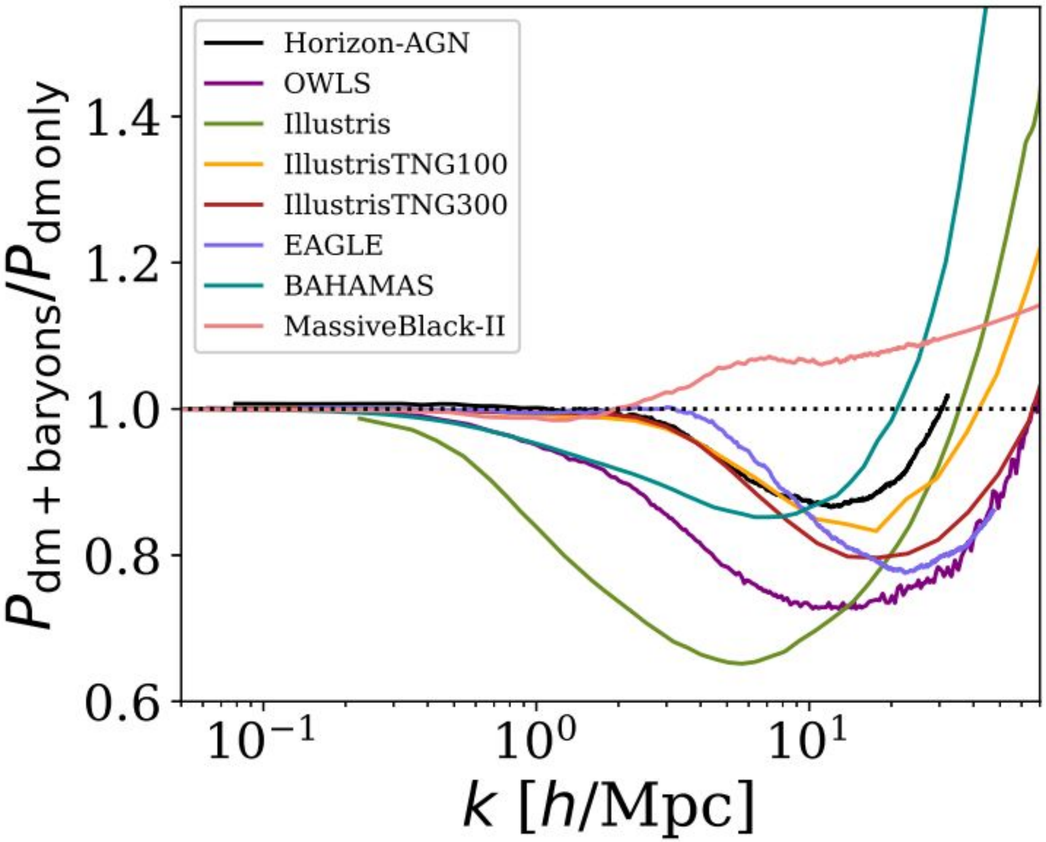
- (Non) linear matter power spectrum
- Baryonic effects
- Galaxy correlation functions
- Halo Mass function
- Perturbative bias in real & redshift space
- Linear theory

Gas and hydro forces



Hydrodynamical forces and galaxy formation is important to predict the large-scale structure of a given cosmology

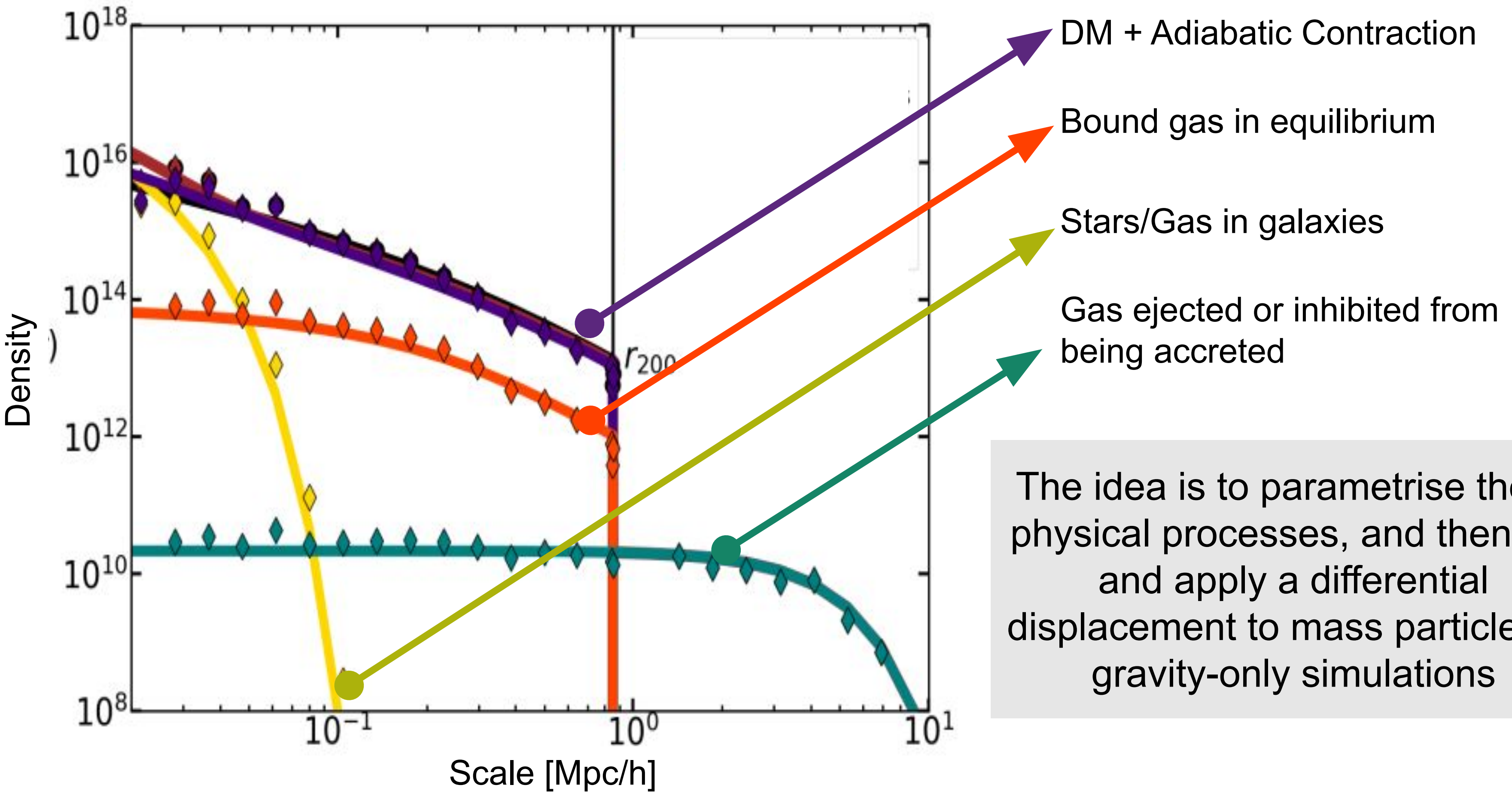




Chisari et al (2019)

Modelling the baryonic effects on the matter clustering

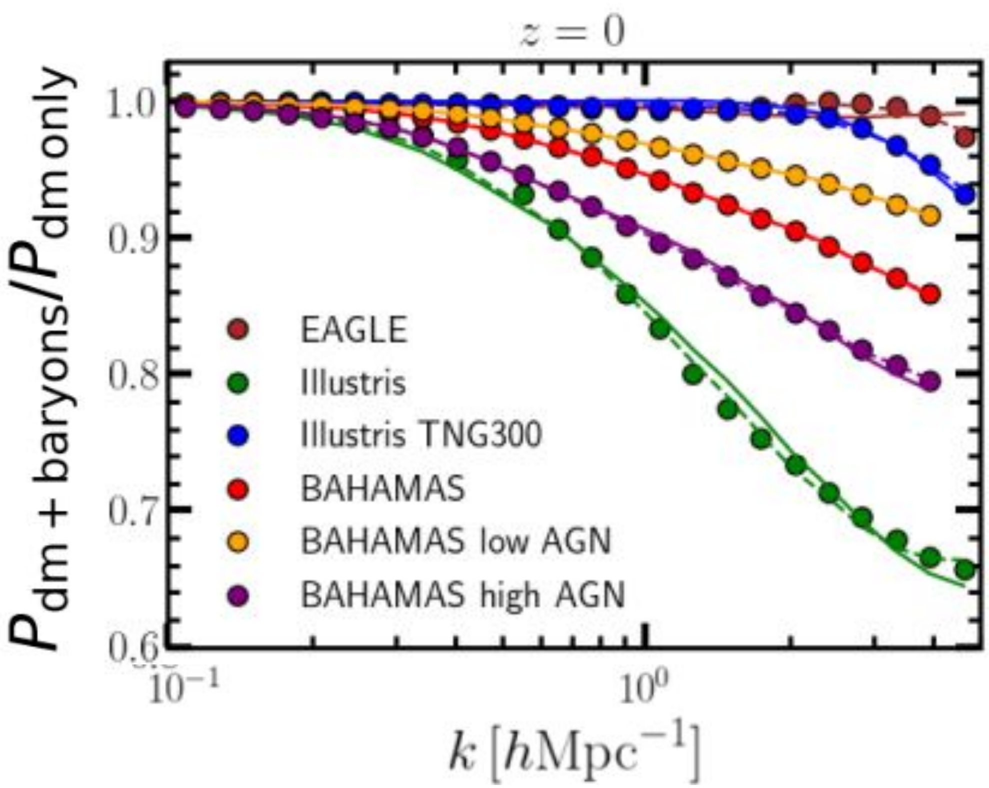
Arico, REA+ (2020,2021); Schneider+(2017, 2019)



The idea is to parametrise these physical processes, and then tag and apply a differential displacement to mass particles in gravity-only simulations

Modelling the baryonic effects on the matter clustering

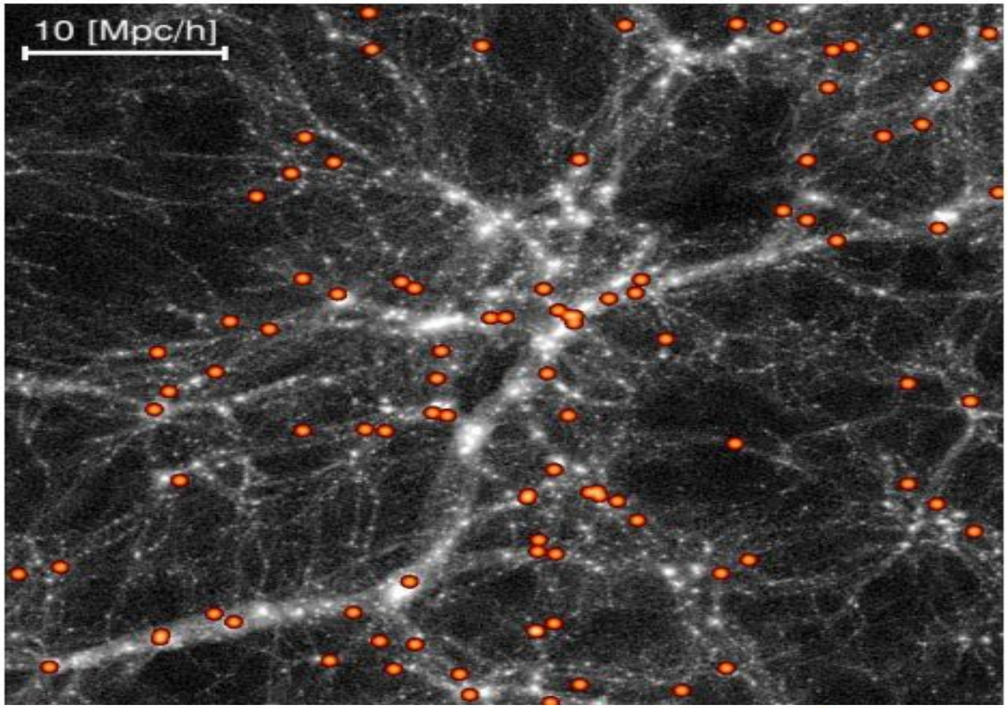
The impact of baryons can be reproduced to ~1% accuracy on 7 different state-of-the-art hydrodynamical simulations

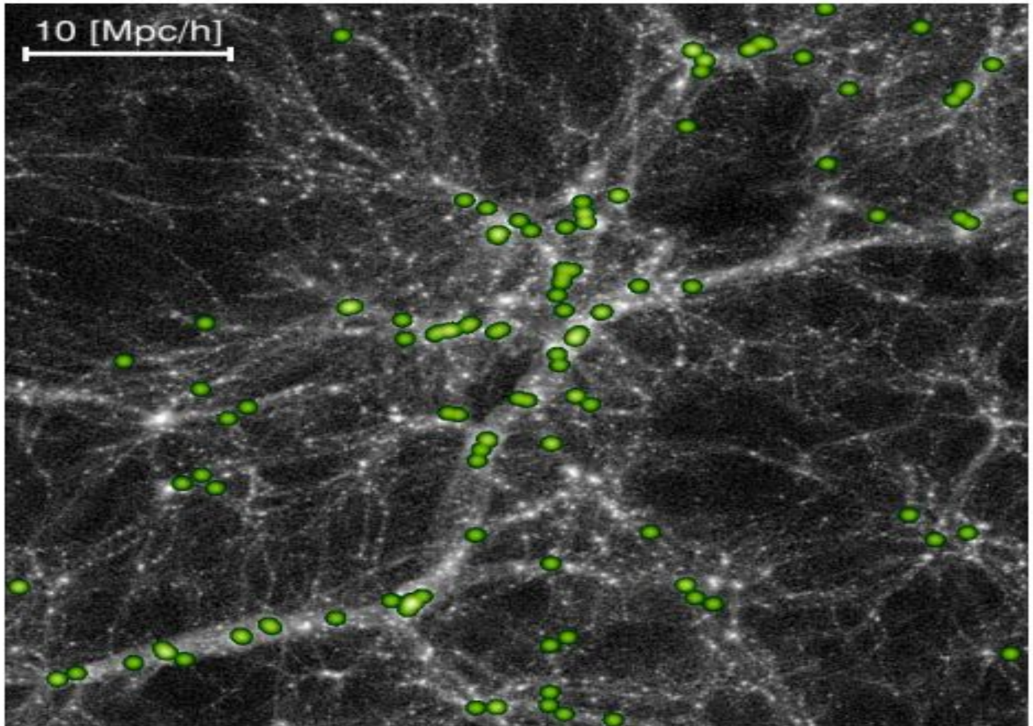


1. Fast & flexible 3D predictions
2. It can be calibrated with observations
3. It does not depend on any specific simulation
4. Also validated for WL peaks and bispectrum and at high redshift

Arico, Angulo, et al (2019, 2020, 2021);

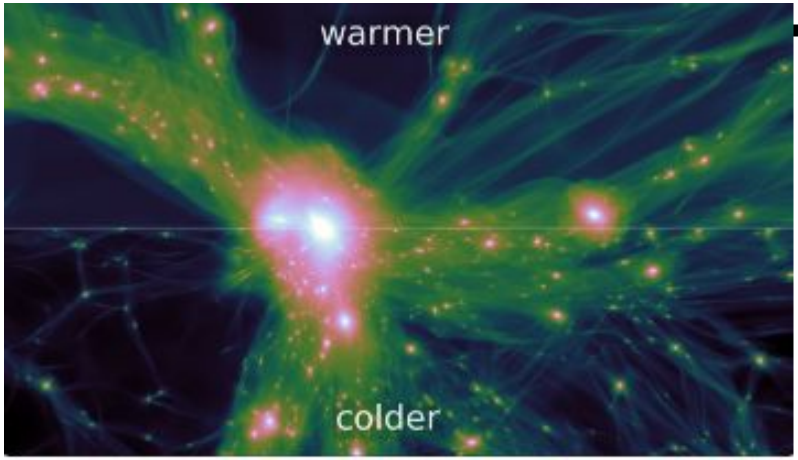
GALAXIES

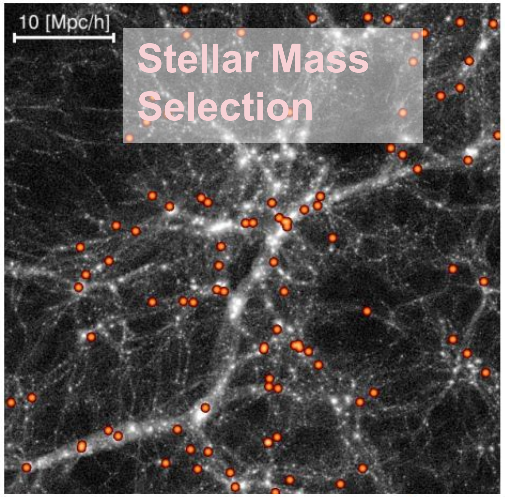


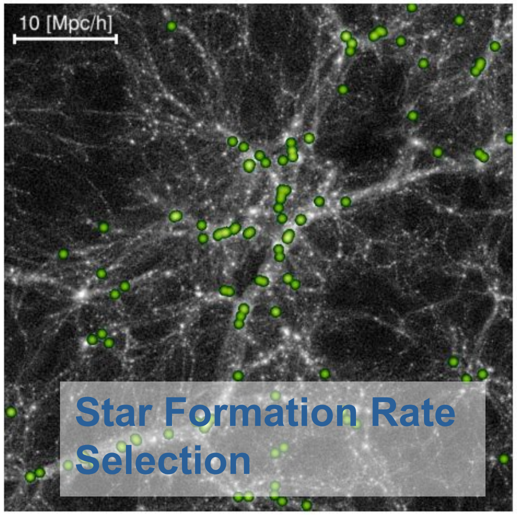


How to efficiently model galaxies in DM sims?

- Halo Occupation Distribution Models
- **(Extended) Subhalo Abundance Matching**
- Empirical Modelling
- Semi-Analytic Galaxy Formation



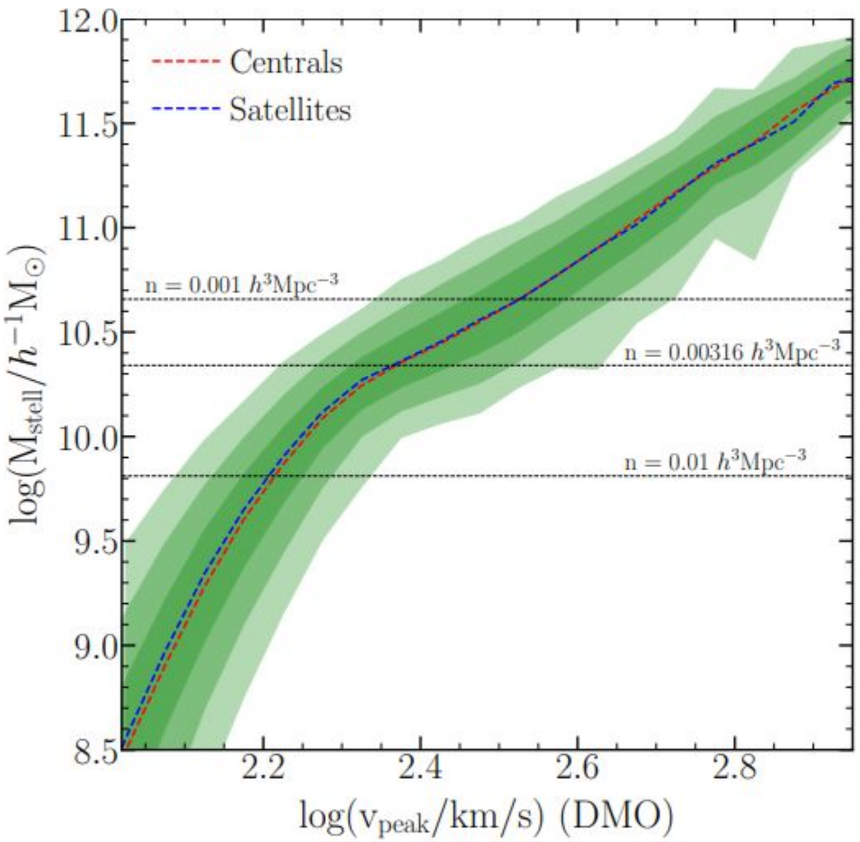




Contreras, REA, Zennaro (2020)
Contreras, REA, et al (2021)
Contreras, REA, MTNG (2022)

SHAMe: Subhalo Abundance Matching Extended

Goal: Flexible and efficient model to fit all available sims



Subhalo Abundance Matching

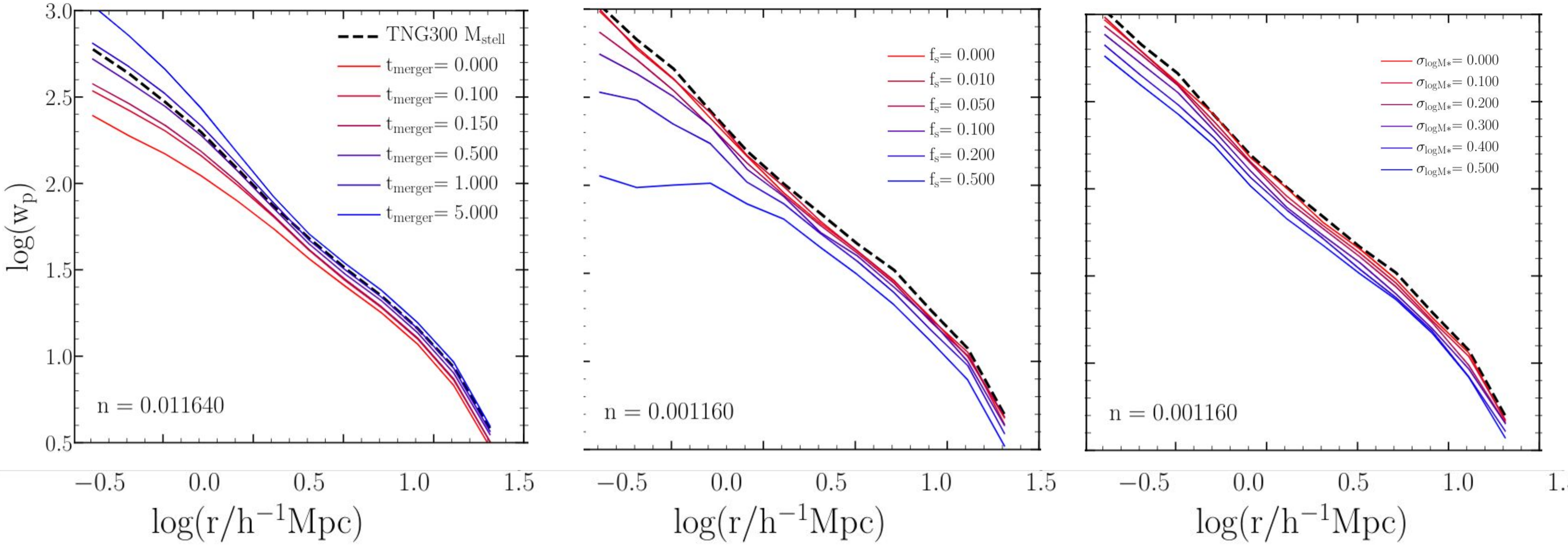
1. Describes the Mhalo-Mstar and clustering of galaxies in observational and simulated data
2. Fails in redshift space
3. Incorrect environmental dependence
4. Incorrect stripping and disruption

extended Subhalo Abundance Matching (alla Emerge/UniverseMachine)

1. Model for the stripping of stars
2. Predictions for Star Formation Rates
3. Model for dynamical friction (orphan tracking)
4. Flexible degree of assembly bias

The impact of astrophysics in the spatial distribution of galaxies

Merger Timescale | Tidal stripping of stars | Stellar-DM Mass scatter



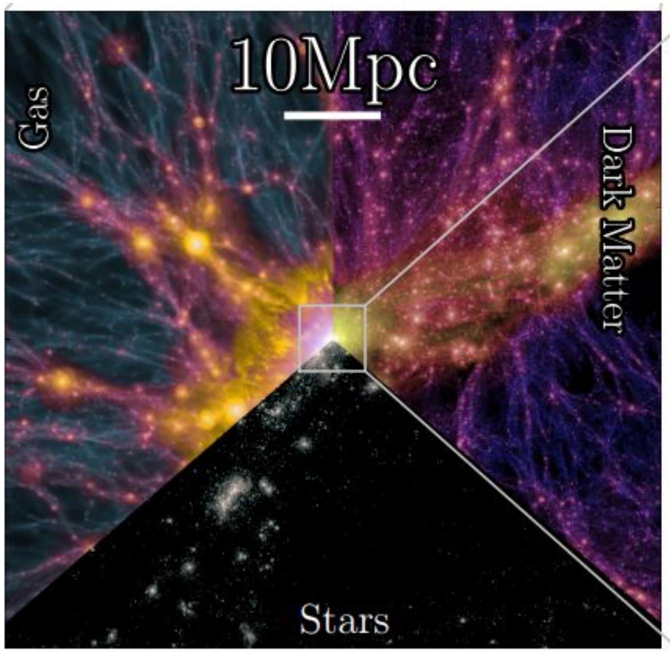
Validated for:

- redshift space multipoles
- 3-point correlation functions
- Amount of assembly bias
- k-Nearest neighbours
- Stellar Mass and SFR galaxy selections

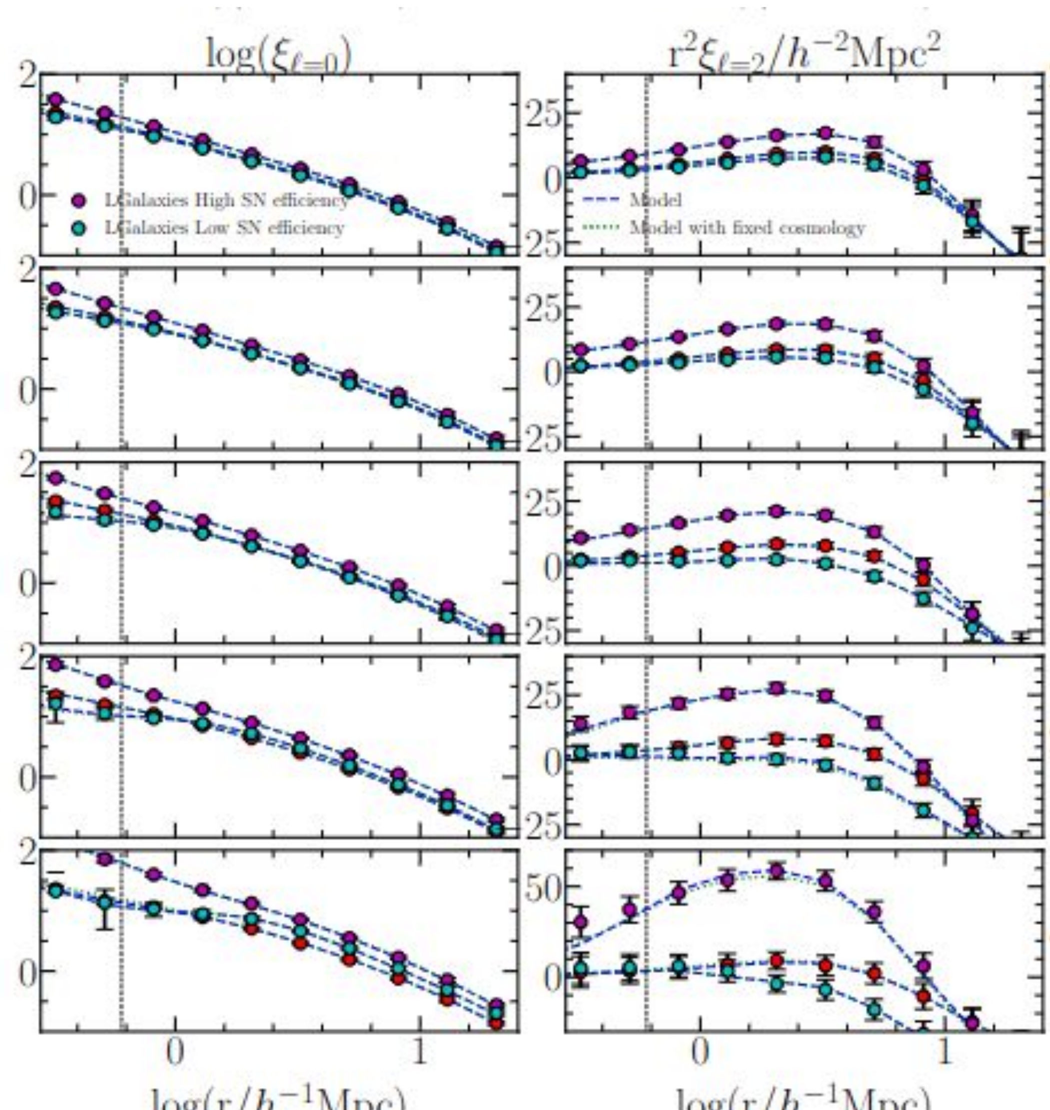
Contreras, Angulo, et al (2021)

SHAMe meets Millennium-TNG hydro simulation

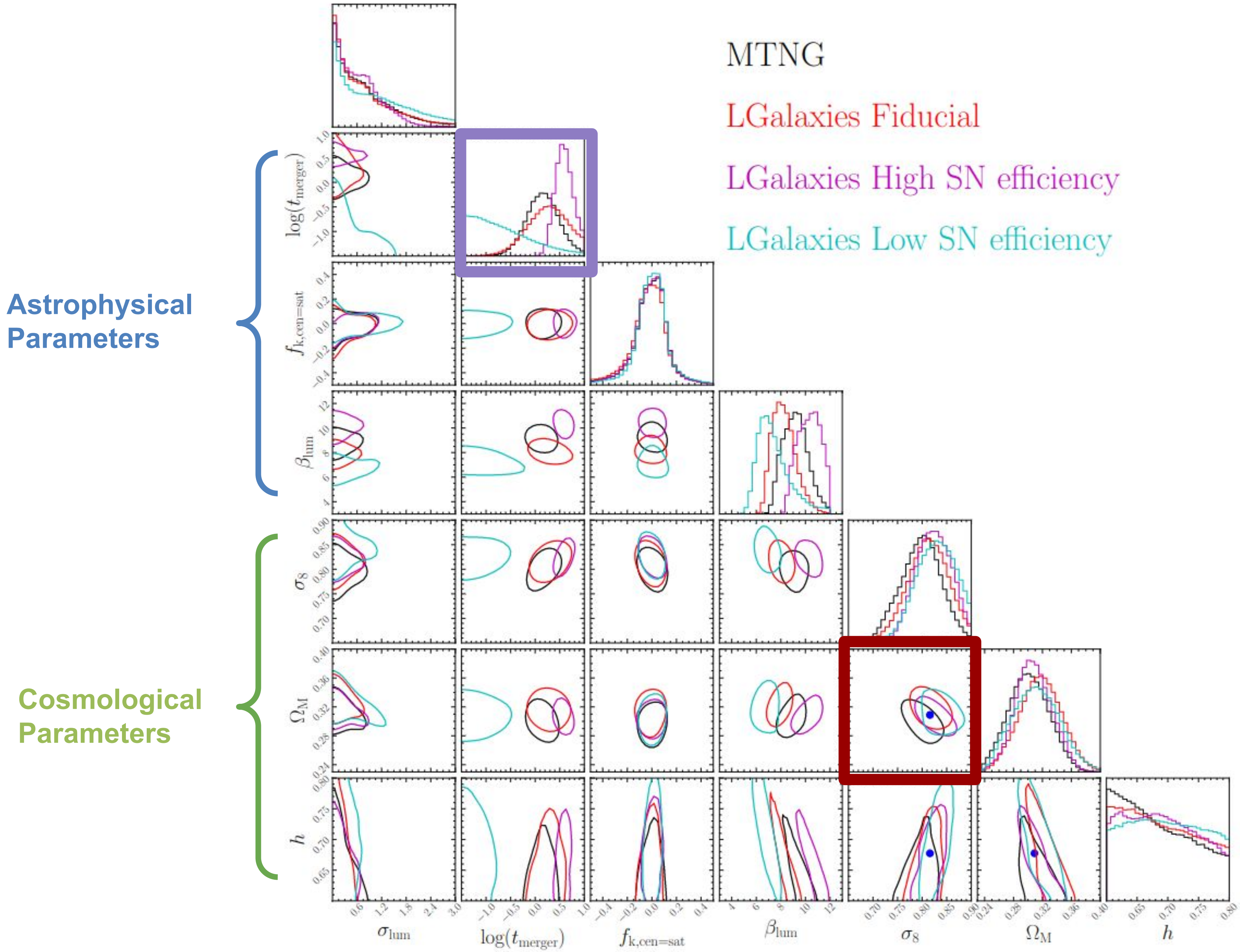
SHAMe is able to describe the redshift-space clustering of MTNG galaxies



L = 740 Mpc
(similar volume to SDSS main)
Same resolution as TNG300
4 additional SAM catalogues
4 number densities

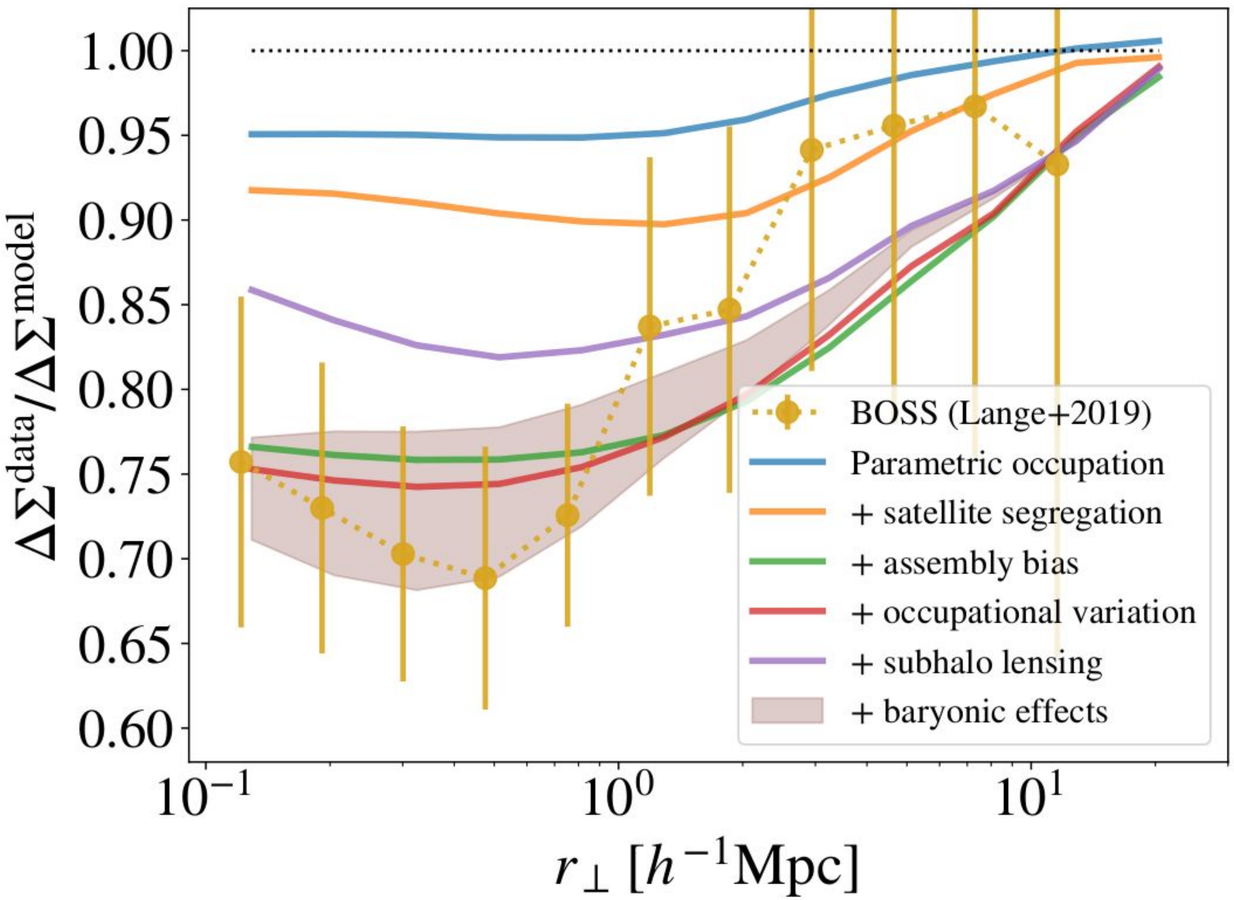


Contreras, REA+ (2022)


MTNG
LGalaxies Fiducial
LGalaxies High SN efficiency
LGalaxies Low SN efficiency
Astrophysical Parameters
Cosmological Parameters
$\sigma_{\rm lum}$
$\log(t_{\rm merger})$
$f_{\rm k,cen=sat}$
$\beta_{\rm lum}$
σ_8
$\Omega_{\rm M}$
h

“Lensing is low” tension

HOD modelling lead to artificial tensions in clustering and gg-lensing



Chavez-Montero, REA, et al (2023)

BACCO's Hybrid bias expansion

A rigorous model that is agnostic to selection effects and gas-physics

The galaxy overdensity field, can be written as a function of matter (and velocity) statistics

$$\delta_g(\boldsymbol{x}, \tau) = \mathcal{F}[\Phi, \Phi_v]$$

Which can then be written perturbatively up to a given order. Specifically, in Lagrangian coordinates:

$$1 + \delta_g(\boldsymbol{q}) = 1 + b_1^L \delta(\boldsymbol{q}) + b_2^L \delta^2(\boldsymbol{q}) + b_{s^2}^L s^2(\boldsymbol{q}) + b_{\nabla^2\delta}^L \nabla_{\boldsymbol{q}}^2 \delta(\boldsymbol{q}) + \cdots ,$$

$$1 + \delta_g(\boldsymbol{x}) = \int d^3\boldsymbol{q}\, \delta^D(\boldsymbol{x} - \boldsymbol{q} - \boldsymbol{\psi}(\boldsymbol{q}))(1 + \delta_g(\boldsymbol{q}))$$

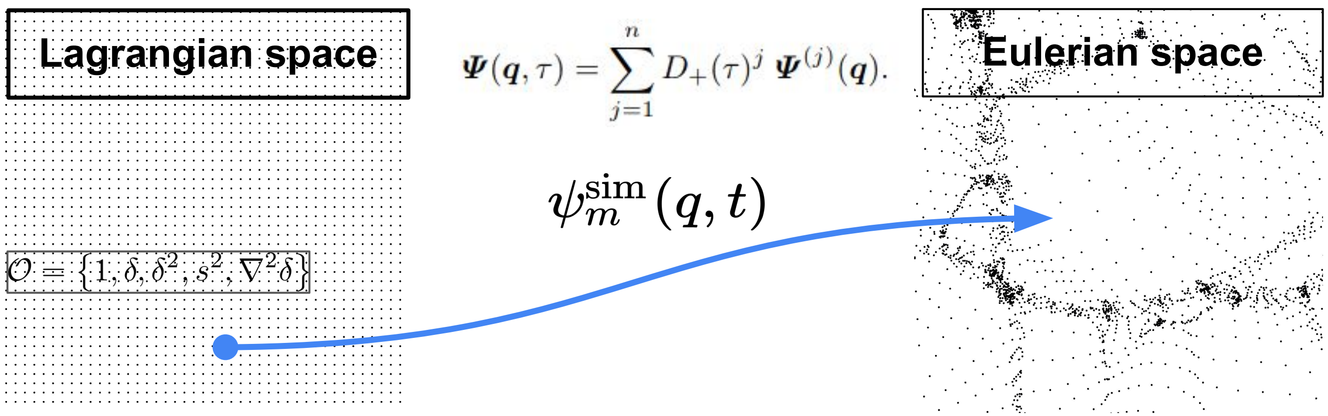
Additionally, the displacement field can be computed perturbatively

$$\boldsymbol{\Psi}(\boldsymbol{q}, \tau) = \sum_{j=1}^{n} D_+(\tau)^j \, \boldsymbol{\Psi}^{(j)}(\boldsymbol{q}).$$

However, perturbative approaches break at ~0.1 h/Mpc – potentially neglecting 100x modes

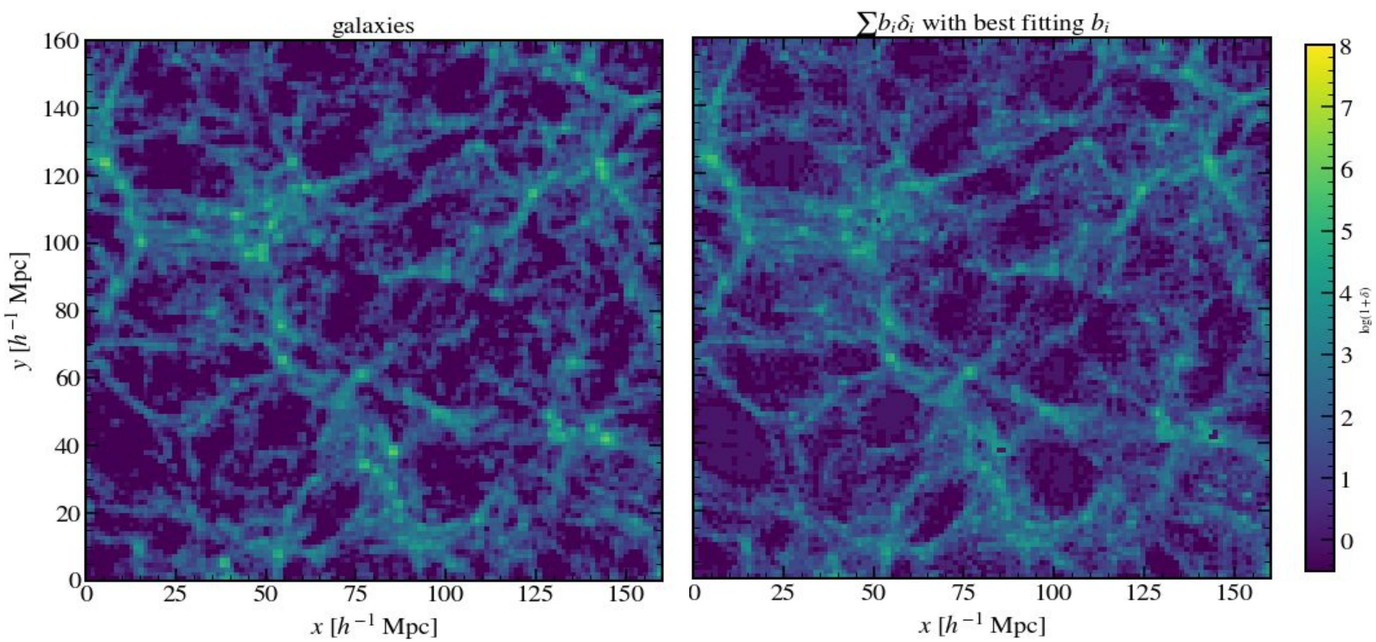
Alternative approach: Hybrid bias expansion

Combining numerical and perturbative worlds (M. Pellejero & M. Zennaro)



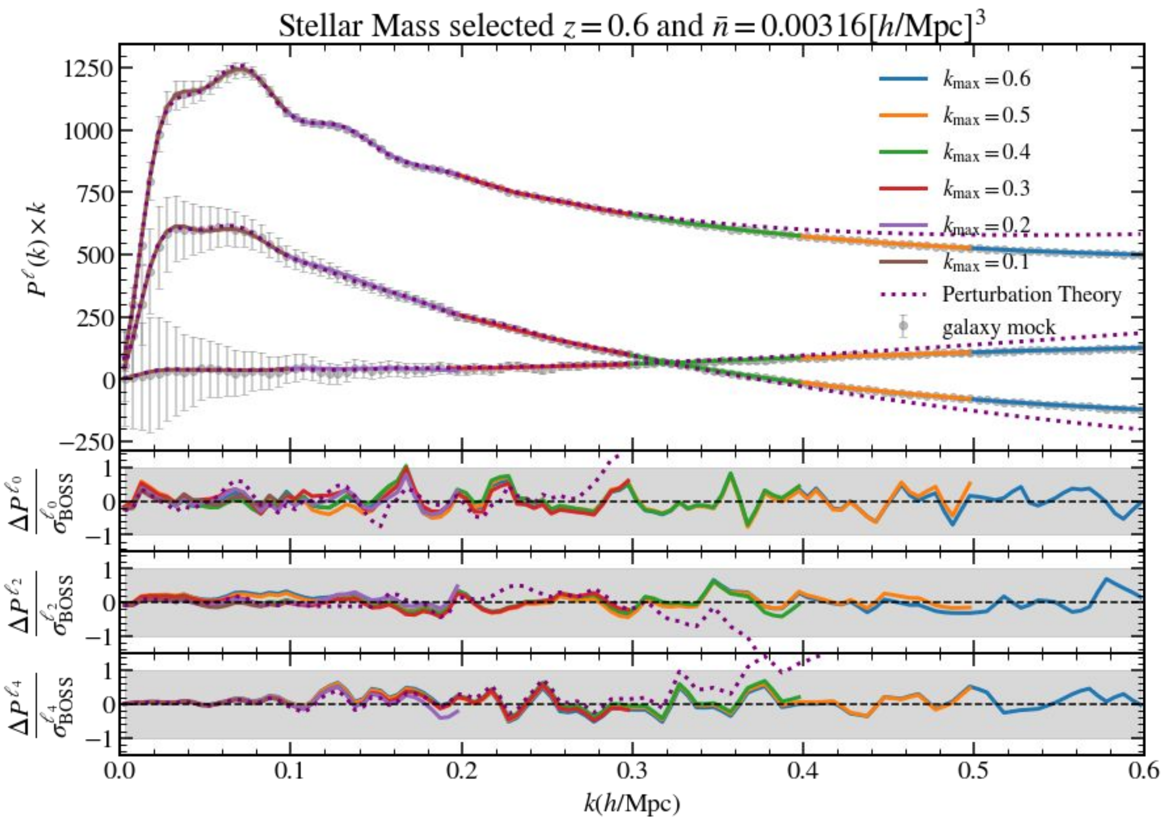
The unc… order

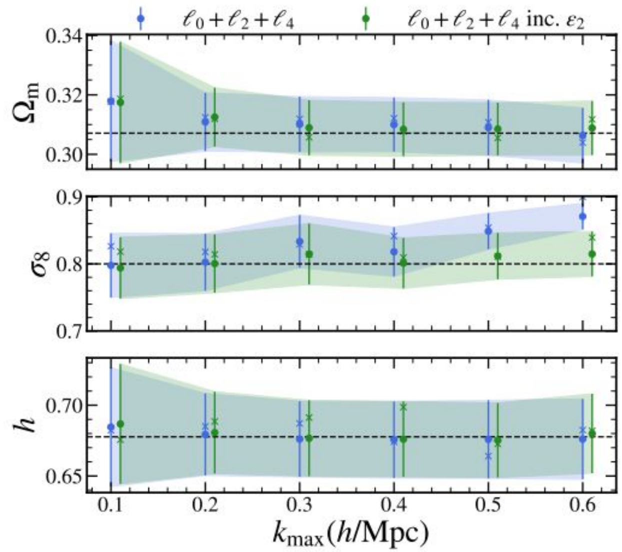
The tran… ft space), can be p…

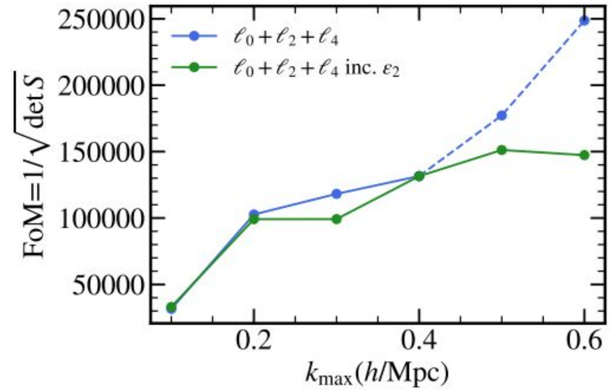


BACCO's Hybrid bias expansion in z-space

Fits to the 300 official BOSS mock catalogues



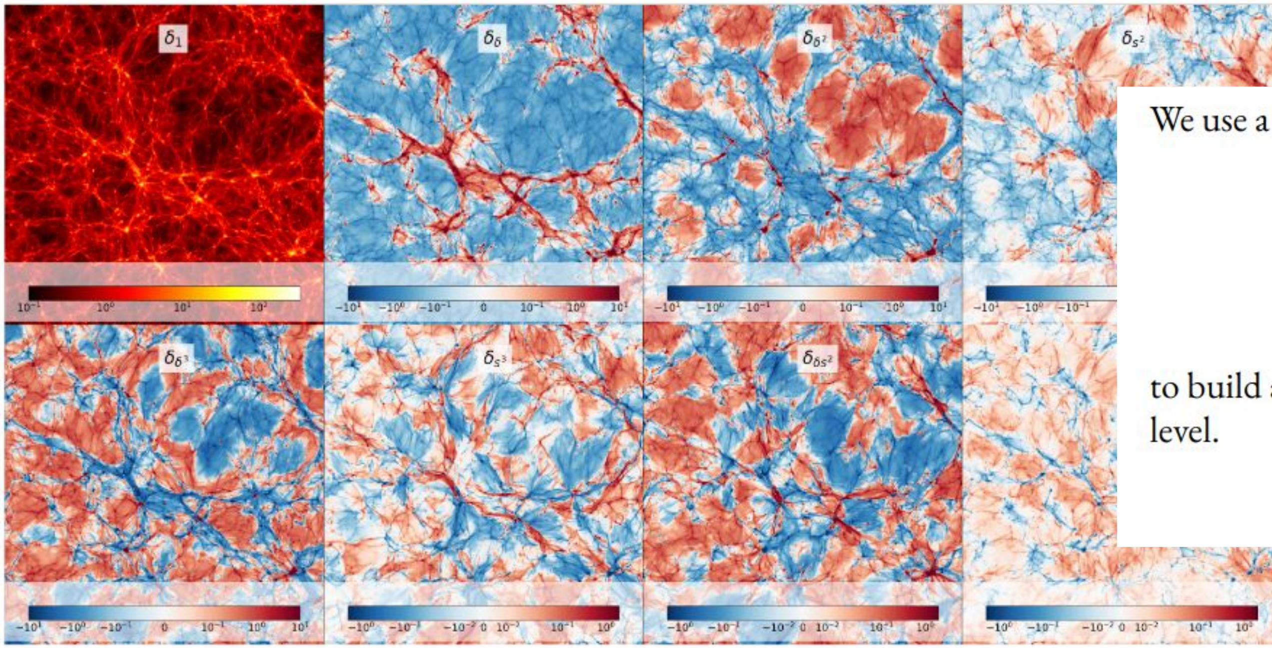




Zennaro, REA,+ (2021,2022)
Pellejero-Ibañez, RA+ (2021,2022)

Describing galaxies at the “field level”

Galaxy field is built by combining all 8 terms that appear at 3rd order



We use a third order bias model

$$\begin{cases} \text{1st Order} & \delta \\ \text{2nd Order} & \delta^2, s^2 \\ \text{3rd Order} & \delta^3, s^3, s^2\delta \\ \text{Non-Local} & \nabla^2\delta \end{cases}$$

to build a model for biased tracers at the field level.

$$\delta_{h,det} = \sum_i b_i \mathcal{O}_i$$

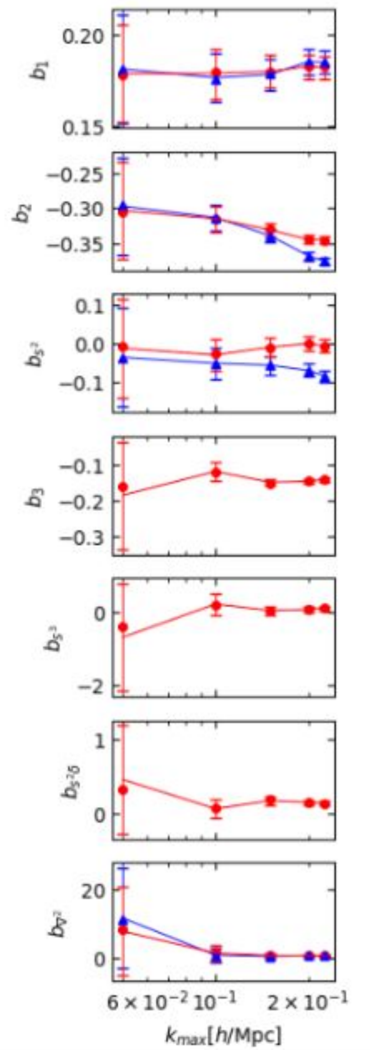
$$-2\ln\mathcal{P}\left(\delta_g|\delta_{g,\mathrm{det}}\right) = \int_{|\boldsymbol{k}|<\Lambda} \frac{\mathrm{d}^3\boldsymbol{k}}{(2\pi)^3}\left[\frac{\left|\delta_g(\boldsymbol{k}) - \delta_{g,\mathrm{det}}(\boldsymbol{k})\right|^2}{P_\epsilon(k)} + \ln\left(2\pi P_\epsilon(k)\right)\right].$$


Francisco Maion

Describing galaxies at the "field level"

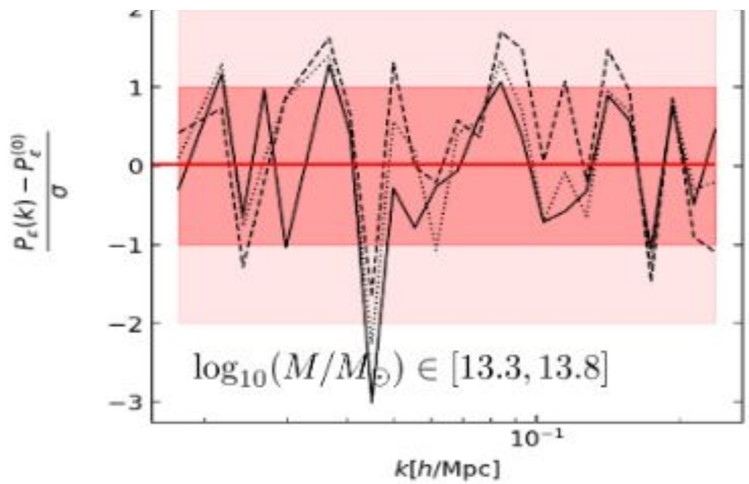
In the most extreme test – fixed phases and cosmology – every single pixel is compatible with stochastic noise up to k=0.25 h/Mpc

Bias parameters measured from galaxy mock created using SHAMe method (Contreras et. al (2021))



2nd order bias
3rd order bias

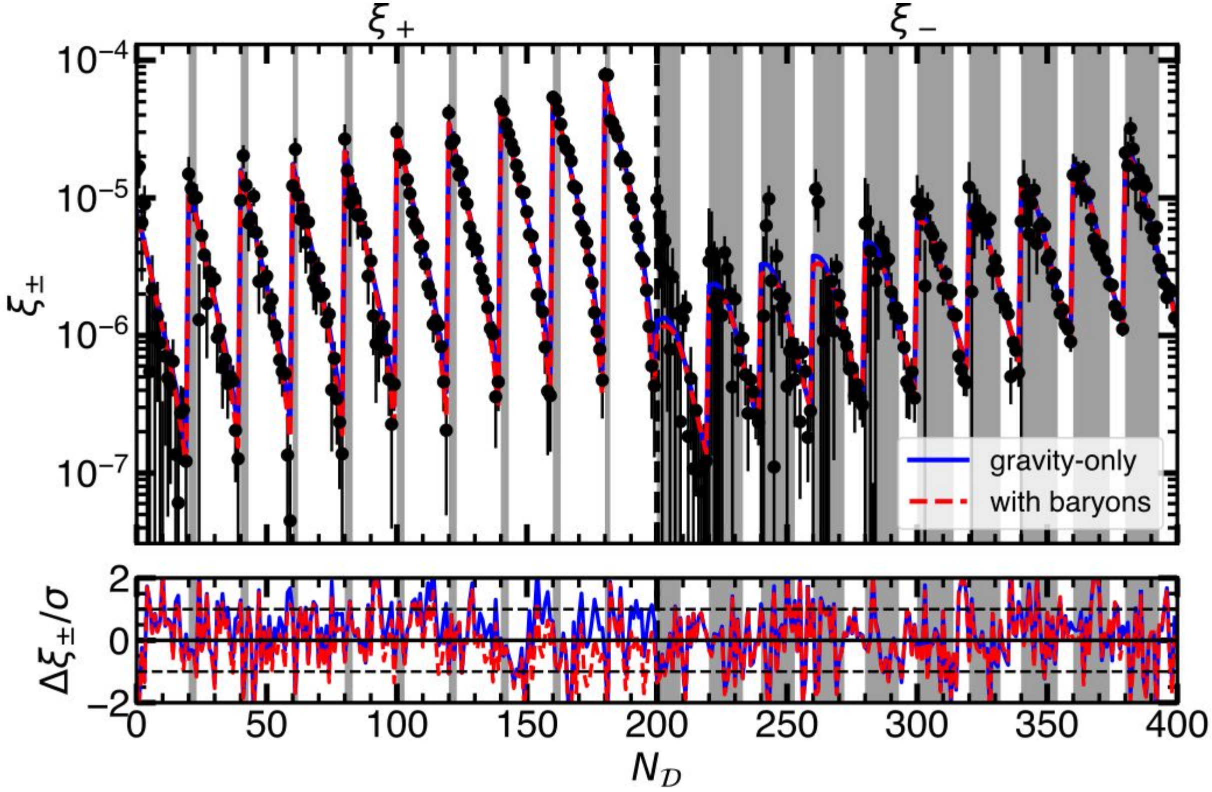
$$P_\epsilon(k) = |\delta_h - \delta_{h,det}|^2$$



1. Extension to redshift space
2. Quantification in constraining power
3. Combination with ML techniques

Exploring the S8 tension

Constraints on cosmology and baryons from small scales



DES Y3 Cosmic shear correlation functions $\xi_\pm$, all angular scales $\theta \in [2.5, 250]$ arcmins.

DARK ENERGY SURVE

Arico, REA+ (2023)

Exploring the S8 tension

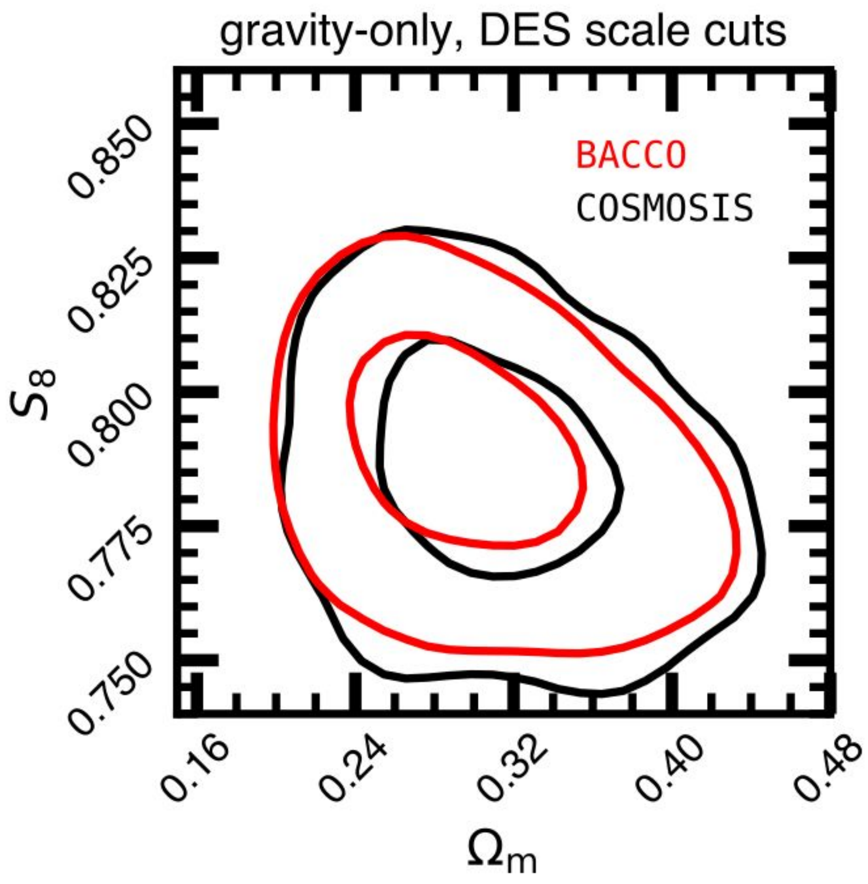
Constraints on cosmology and baryons from small scales

- Independent pipeline written from scratch (0.4s in a laptop)

-More accurate Pnl – BACCOemu

- Explicit baryon model via baryonification

- All scales in shear
- No shear ratios
- NLA intrinsic alignment



DARK ENERGY SURVEY

Exploring the S8 tension

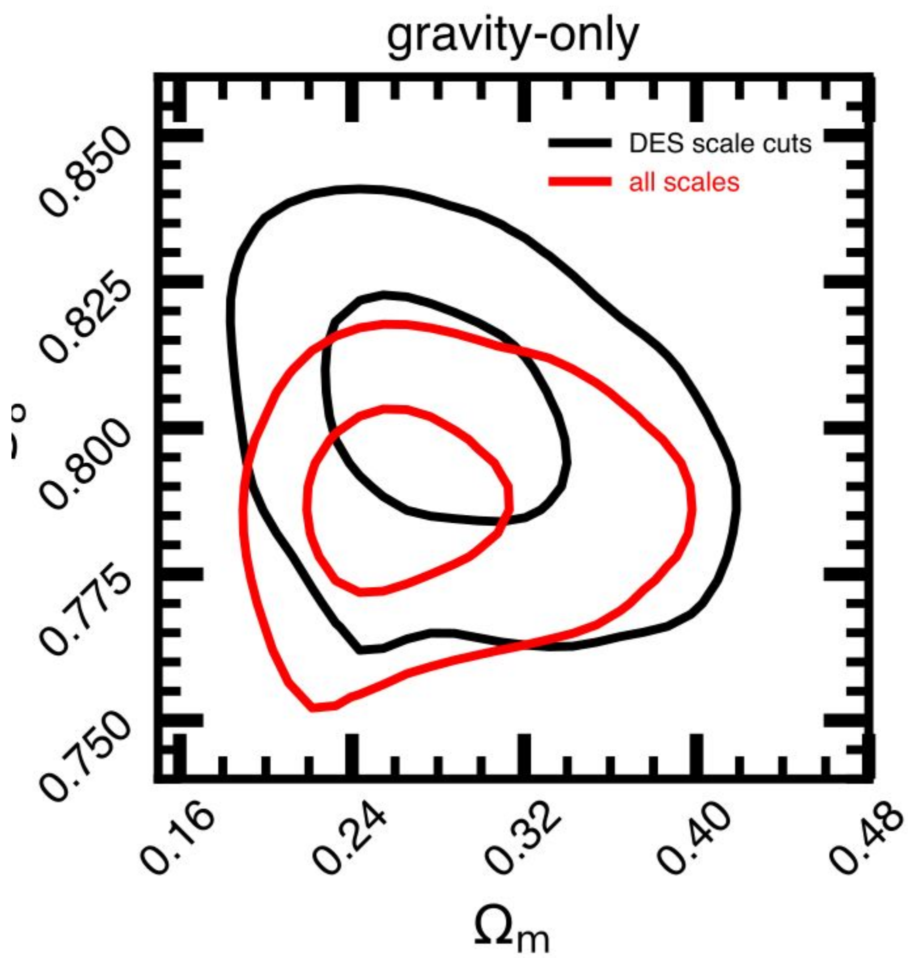
Constraints on cosmology and baryons from small scales

- Independent pipeline written from scratch (0.4s in a laptop)

-More accurate Pnl – BACCOemu

- Explicit baryon model via baryonification

- All scales in shear
- No shear ratios
- NLA intrinsic alignment



DARK ENERGY SURVEY

Exploring the S8 tension

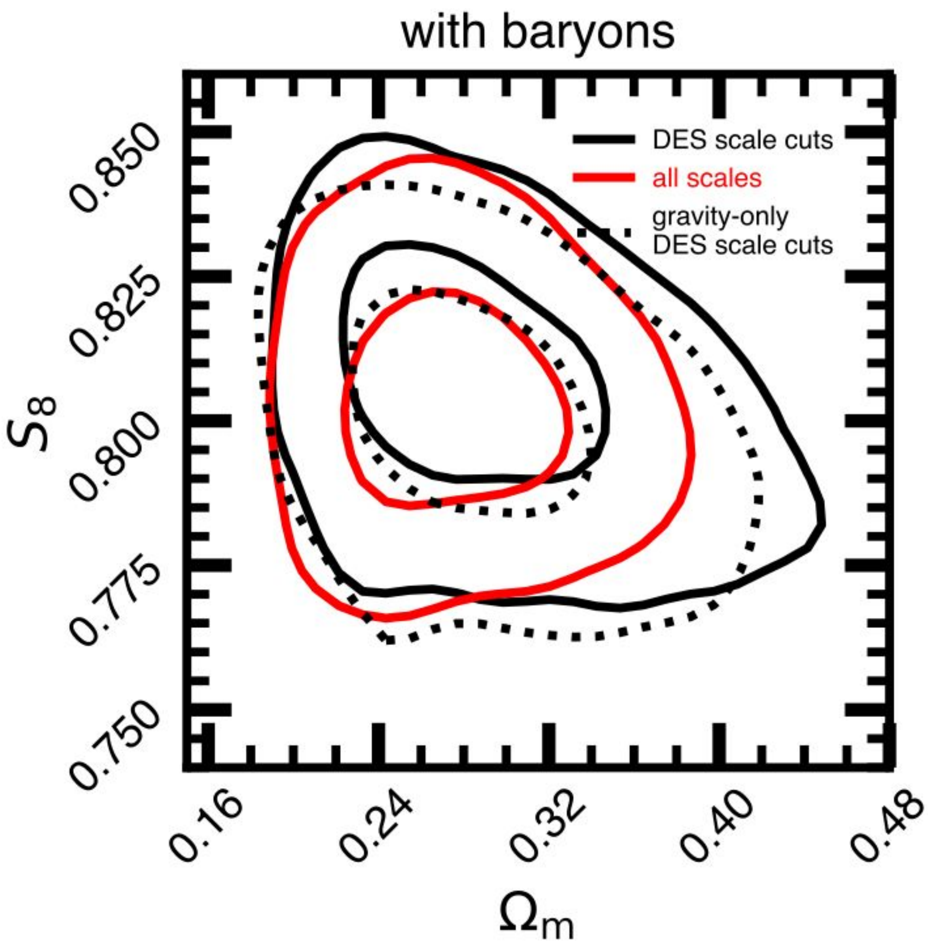
Constraints on cosmology and baryons from small scales

- Independent pipeline written from scratch (0.4s in a laptop)

-More accurate Pnl – BACCOemu

- Explicit baryon model via baryonification

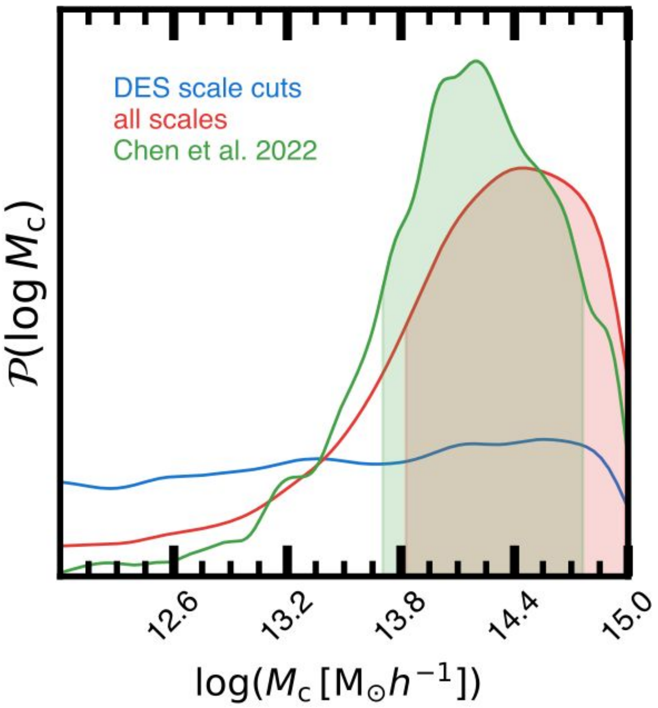
- All scales in shear
- No shear ratios
- NLA intrinsic alignment

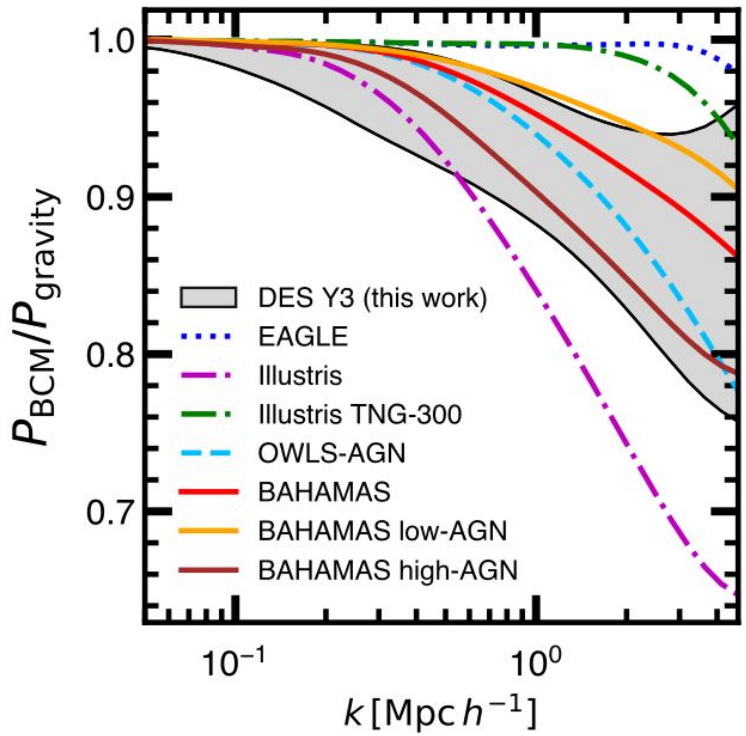


DARK ENERGY SURVEY

Exploring the S8 tension

Results on baryons. Halo mass for which half og the gas is lost due to feedback



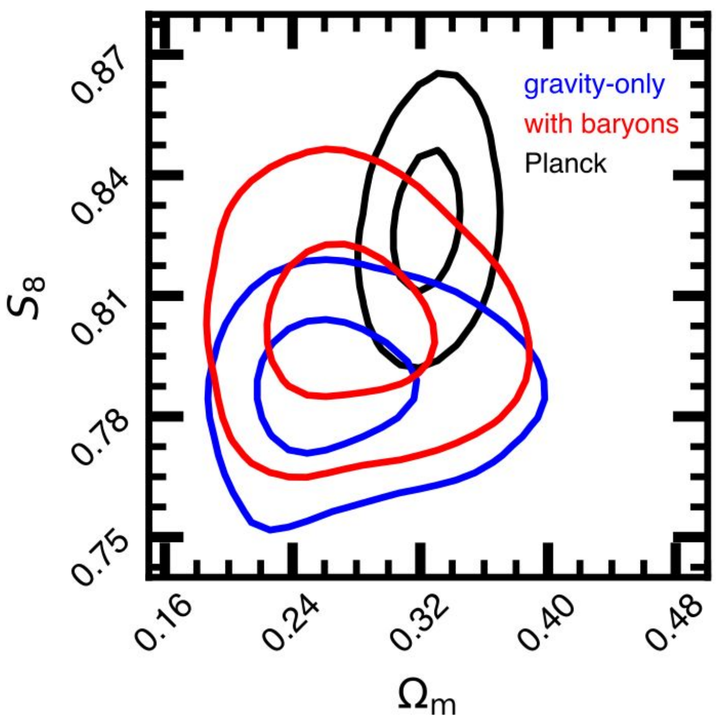


Arico, REA + (2023)

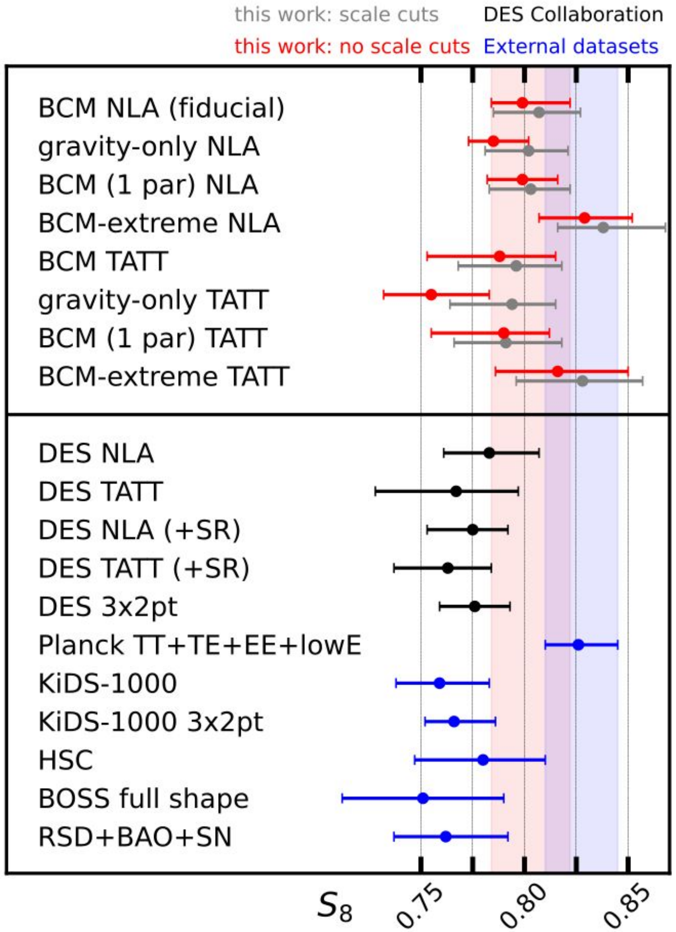
$$\log M_c = 14.38^{+0.60}_{-0.56} \log(h^{-1}M_\odot)$$

Exploring the S8 tension

0.9 sigma tension with Planck for our fiducial case

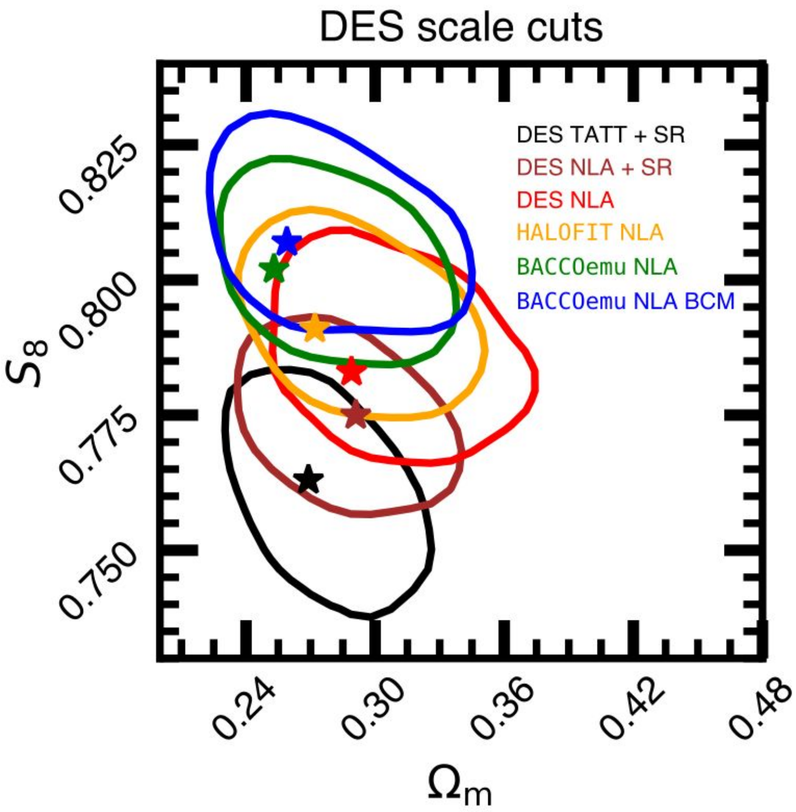


0.9σ tension with Planck in our fiducial case



Exploring the S8 tension

Differences with respect to the DES collaboration results



Disagreement in S_8

Our fiducial - DES fiducial 1.4σ

TATT-NLA $0.4 - 0.7\sigma$

Shear ratios $0.1 - 0.3\sigma$

Pipeline/priors $\approx 0.2\sigma$

Non-linearities $0.4 - 0.5\sigma$

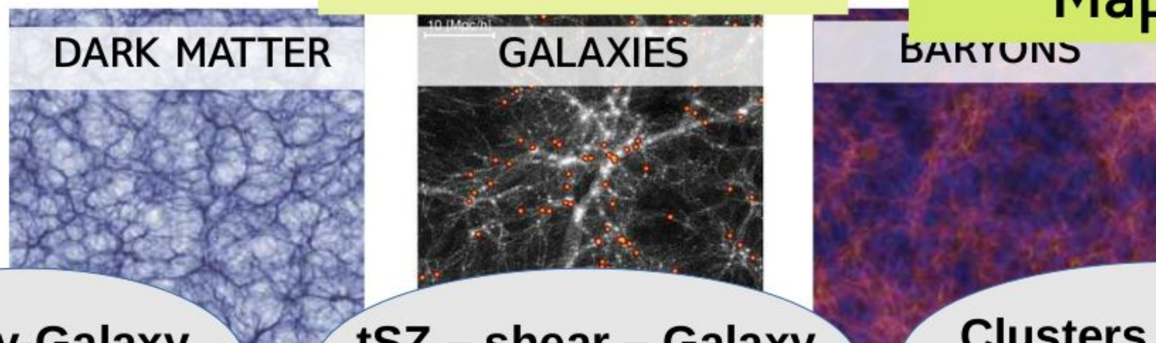
Baryons $\approx 0.2\sigma$

Simultaneous modelling of LSS data

Lensing

Clustering
Clusters

SZ
Intensity
Mapping



Galaxy-Galaxy Lensing

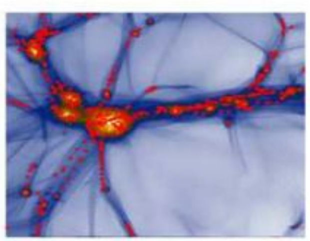
tSZ – shear – Galaxy Cross Correlations

Clusters in X-rays, SZ, optical, lensing

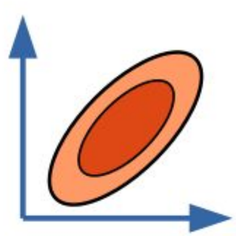
Improvements feedback into a more accurate modelling and better cosmological inferences

Numerical simulations for interpreting large scale structure data

Upcoming large galaxy surveys pose multiple challenges in data analyses and theory



Simulation-based approaches and new summary statistics are providing new options



All these together could offer a robust and accurate path to discovering new physics

http://bacco.dipc.org/