The Abdus Salam International Centre for Theoretical Physics

2419-12

Workshop on Large Scale Structure

30 July - 2 August, 2012

The Millennium-XXL project

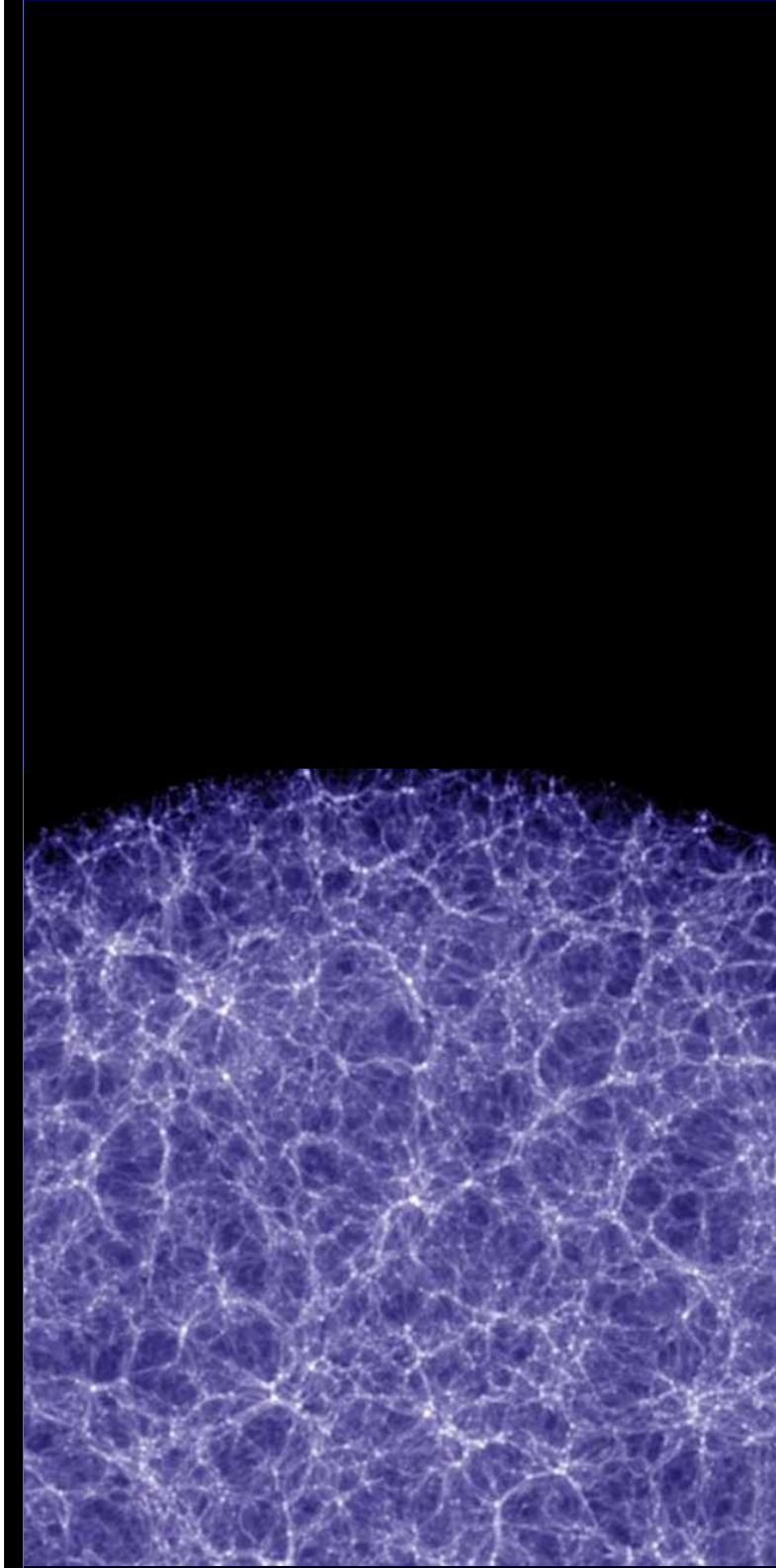
R. Angulo
Max Planck Institute for Astrophysics

The Millennium XXL

Raul Angulo

V. Springel, S. White, A. Jenkins, C. Frenk, C. Baugh

Max-Planck-Institut
für Astrophysik



How can we address the mysteries on the pillars of the Universe?

I) Nature of Gravity

II) Dark Matter

III) Accelerated expansion of the Universe

IV) Properties of the initial fluctuations

How can we address the mysteries on the pillars of the Universe?

I) Nature of Gravity

f(R) or GR?

II) Dark Matter

Cold or Warm? Free streaming scale?

III) Accelerated expansion of the Universe

Cosmological constant or w(z)?

IV) Properties of the initial fluctuations

Any primordial NonGaussianity?

How can we address the mysteries on the pillars of the Universe?

I) Nature of Gravity

f(R) or GR?

Redshift distortions, gravitational redshifts

II) Dark Matter

Cold or Warm? Free streaming scale?

Profiles, rotation curves, substructures, P(k)

III) Accelerated expansion of the Universe

Cosmological constant or w(z)?

BAO, cluster counts, ISW

IV) Properties of the initial fluctuations

Any primordial NonGaussianity?

Large scale bias & clustering, massive clusters

Many tests can be carried out using only the position and redshift of galaxies, specially on large scales

A new era for observational cosmology

Many tests can be carried out using only the position and redshift of galaxies, specially on large scales

A new era for observational cosmology

Requirements:

I) Large volumes

ii) Large number of objects

Planned and current surveys at different redshifts will cover up to hundreds of cubic Gpc

Surveys are gradually approaching the limits of the observable Universe (it is not that big!!).

Many tests can be carried out using only the position and redshift of galaxies, specially on large scales

A new era for observational cosmology

Requirements:

I) Large volumes

ii) Large number of objects

Planned and current surveys at different redshifts will cover up to hundreds of cubic Gpc

Goal: Relate the observed galaxy distribution to the underlying mass concentration

The required percent level accuracy in the interpretation of these mission can only be reached through a much better understanding of several effects

Non-linearities

z-space distortions

Dark Matter haloes

Galaxy formation

Evolution of clustering

Intrinsic

Radial selection function

Discreteness effects

Photometric calibration

Fiber collisions

Contamination by stars

Extrinsic

Direct cosmological simulation are essential for this. They can also tell us how to extract the maximum info about Cosmology and galaxy formation physics.

New generation of observations require new generation of simulations!!

The challenges for numerical simulations:

1) **Simulate volumes as large as future surveys**
2) **Mass resolution adequate to model galaxies; their scale-dependent bias and velocity**

The Millennium XXL

- L-Gadget 3 (65 bytes per particle)
- 6720^3 ~ 303 billion particles
- 3000 Mpc/h box
- 12288 cores: 3072 MPI-tasks/4 threads
- Millennium Cosmology, 2LPT ICs

- 86 trillion force calculations.
- 2.7 million CPU hours
- 15% in SF/FoF, P(k) and other on-the-fly calculations, 14% in I/O, 3% in long-range PM force.
- Peak memory usage: 29 TB (105 bpp)

The computational challenge

Such calculation poses challenges in the raw execution time, memory requirements and data handling.

CPU & load Imbalances; Almost 1000 times more force calculations than in the MS.

RA Memory: 10Tb are needed simply to hold Ids and phase-space variables. L-Gadget 2 (used for the MS) would need 30Tb (without memory imbalances or group finders)

I/O & Data: If we take the same approach as for the MS and MII, then we would need about 700 Tb of disk space.

We require a combination of extremely efficient and scalable algorithms, and a large Supercomputer!

MXXL

- **L = 3000 Mpc/h**
- **N = 6720^3 particles**
- **Є = 10 kpc/h**
- **M = 6.18×10^9 Msun/h**

JuRoPa/JSC

Jülich Research on Petaflop Architectures

Peak Performance: **207 Teraflop/s**

2208 Nodes with 2 processors Intel Xeon X5570 (Nehalem-EP) quad-core, 2.93 GHz .

RAM: **52 Terabyte**

10th largest supercomputer in the world in June 2009.

However….

The MPI library would use 50 Gb of RAM per node; more memory than the available in a node.

Using the whole machine is unfeasible.

Lean-Gadget3

- **Support for hybrid calculations with MPI+Pthreads+OpenMP**
- **Extremely memory-efficient tree algorithm**
  (can run with ~64 bytes/particle in TreePM mode if needed)
- **Improved communication scheme for the PM calculation**
  (works more efficiently and uses less peak memory)
- **We have thread-parallelized FFTW-2 ourselves**
- **New parallel distributed sort that scales much better than before**

CPU load & Imbalances

Data Handling & Analyses

New memory efficient and fast inlined FOF and SUBFIND code (can be used on-the-fly or stand-alone)

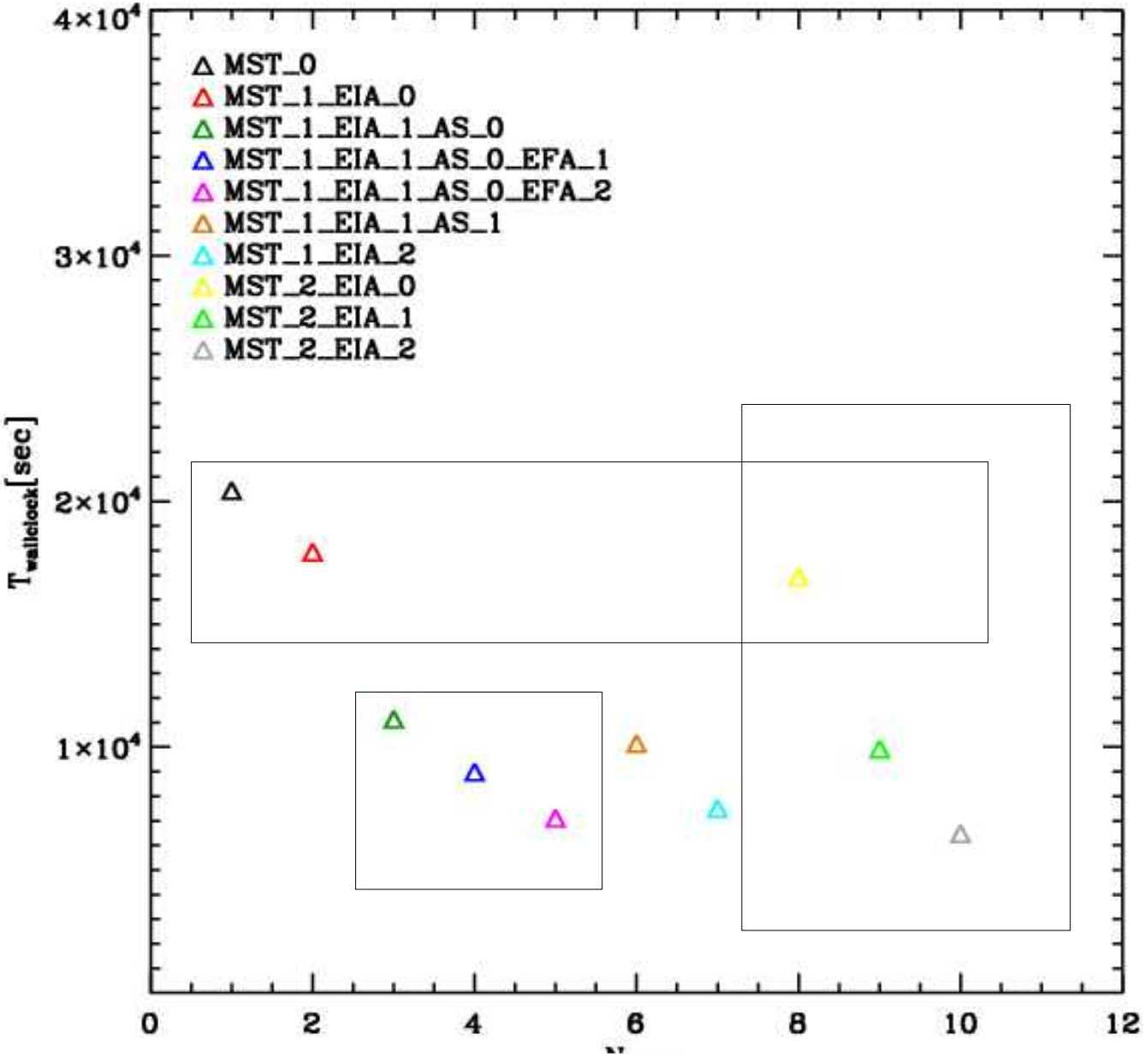
Improved output structure that simplifies later postprocessing.

Many more quantities calculated on the fly.

Compression of data output

Number of particles is not precision

Force and time integration parameters can change execution times by factors of a few



MST: Maximum allowed timestep

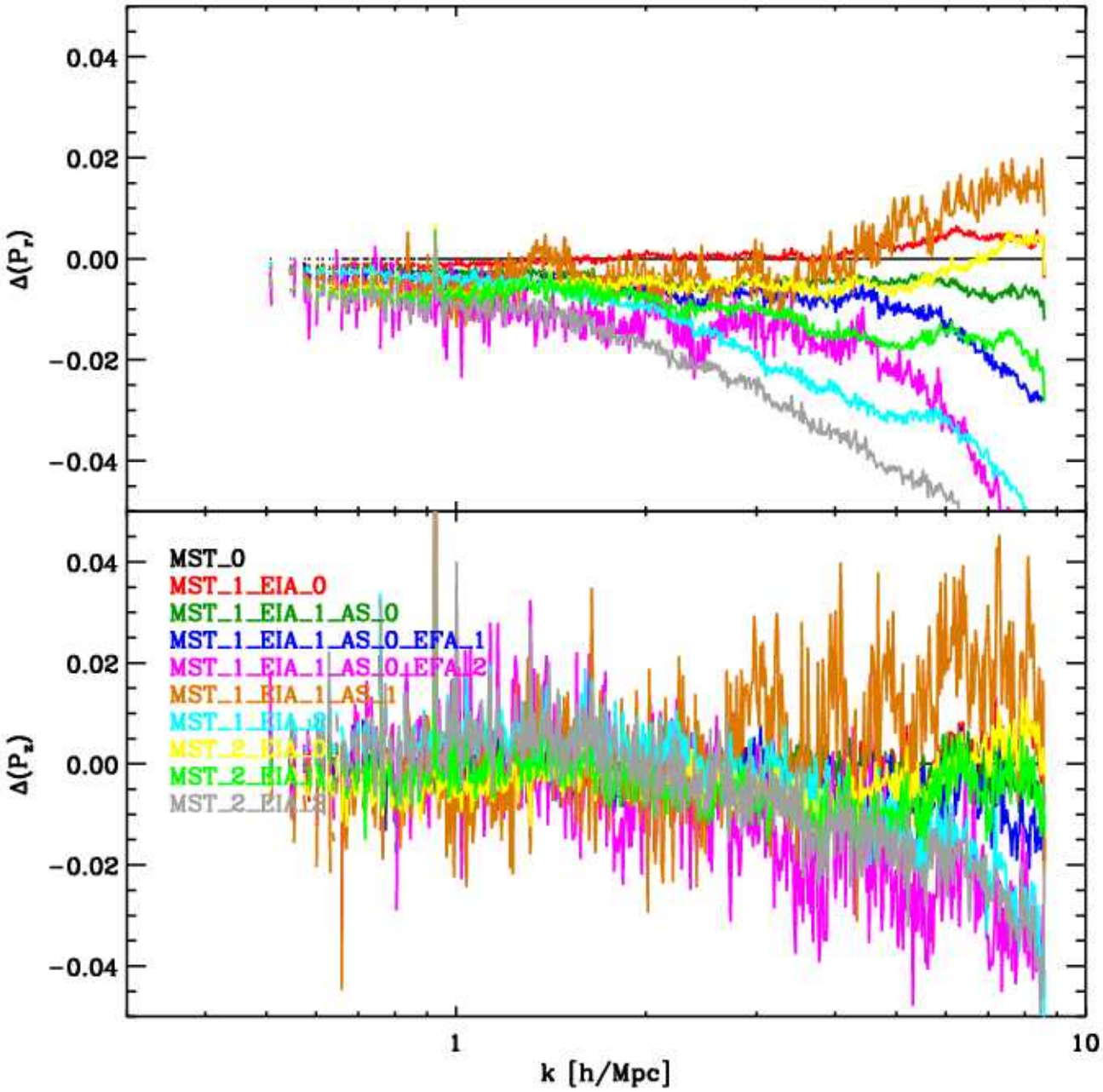
EIA: Error in time Integration

AS: Smoothing applied to mesh force

EFA: Error in Force Calculation

Number of particles is not precision

Errors on the power spectra induced by numerical errors



MST: Maximum allowed timestep

EIA: Error in time Integration

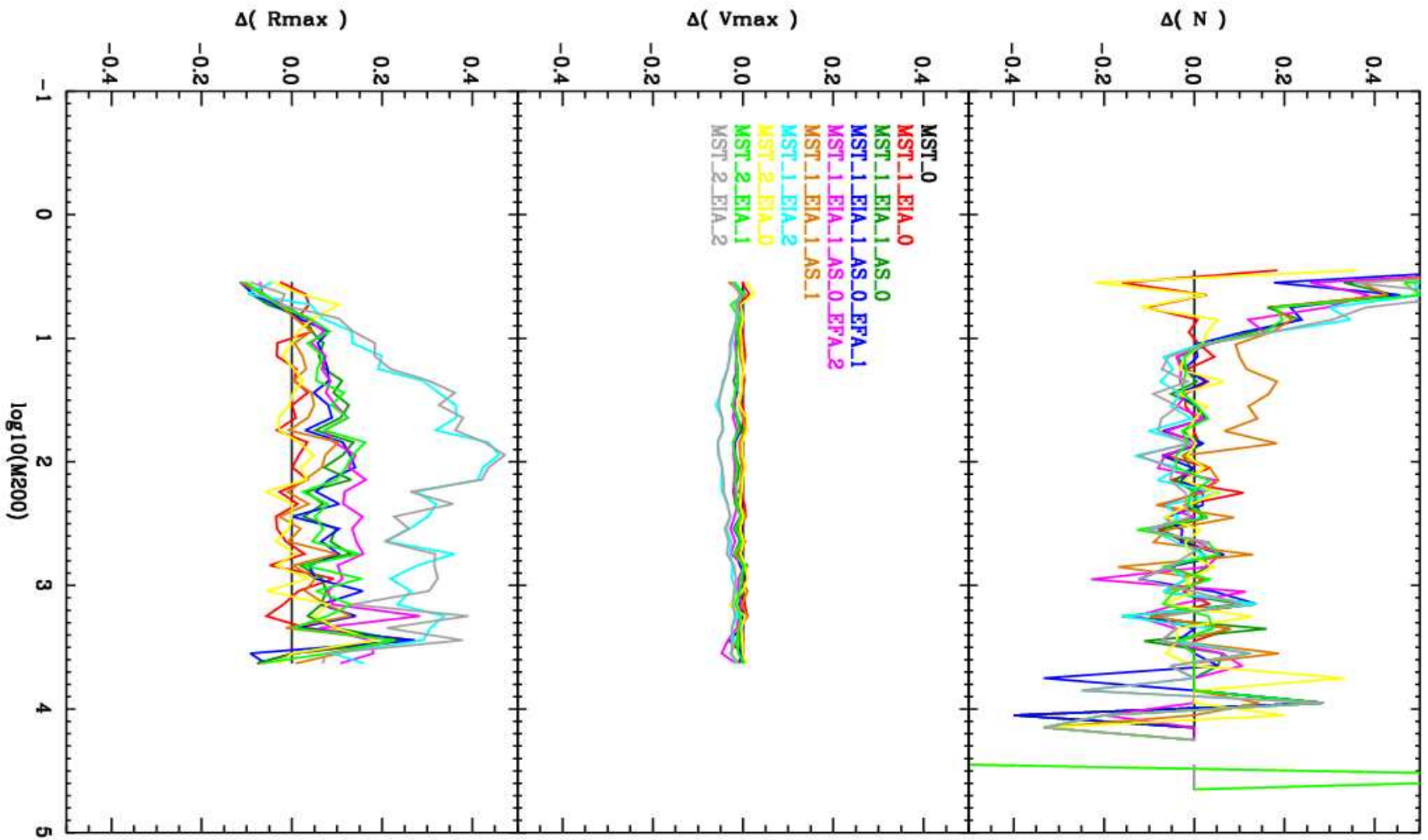
AS: Smoothing applied to mesh force

EFA: Error in Force Calculation

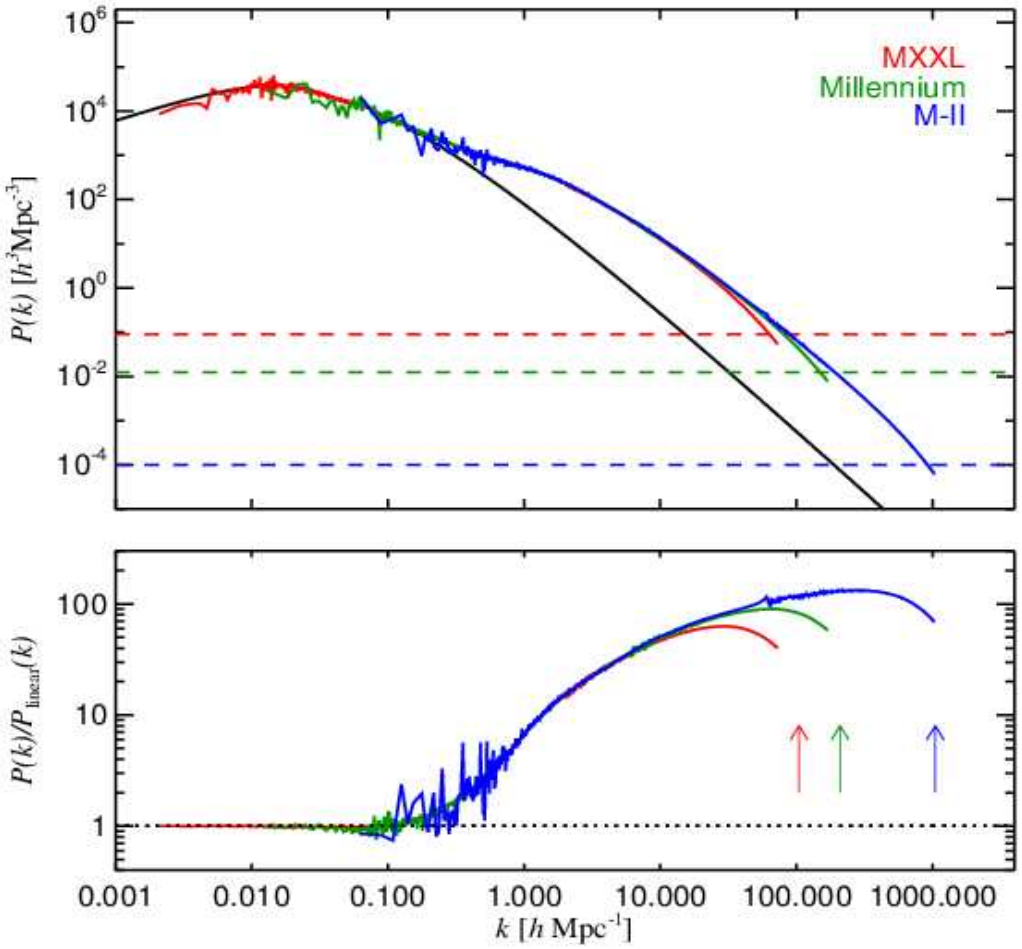
Not including:
1) Starting redshift
2) Transients from the ICs
3) Softening Length

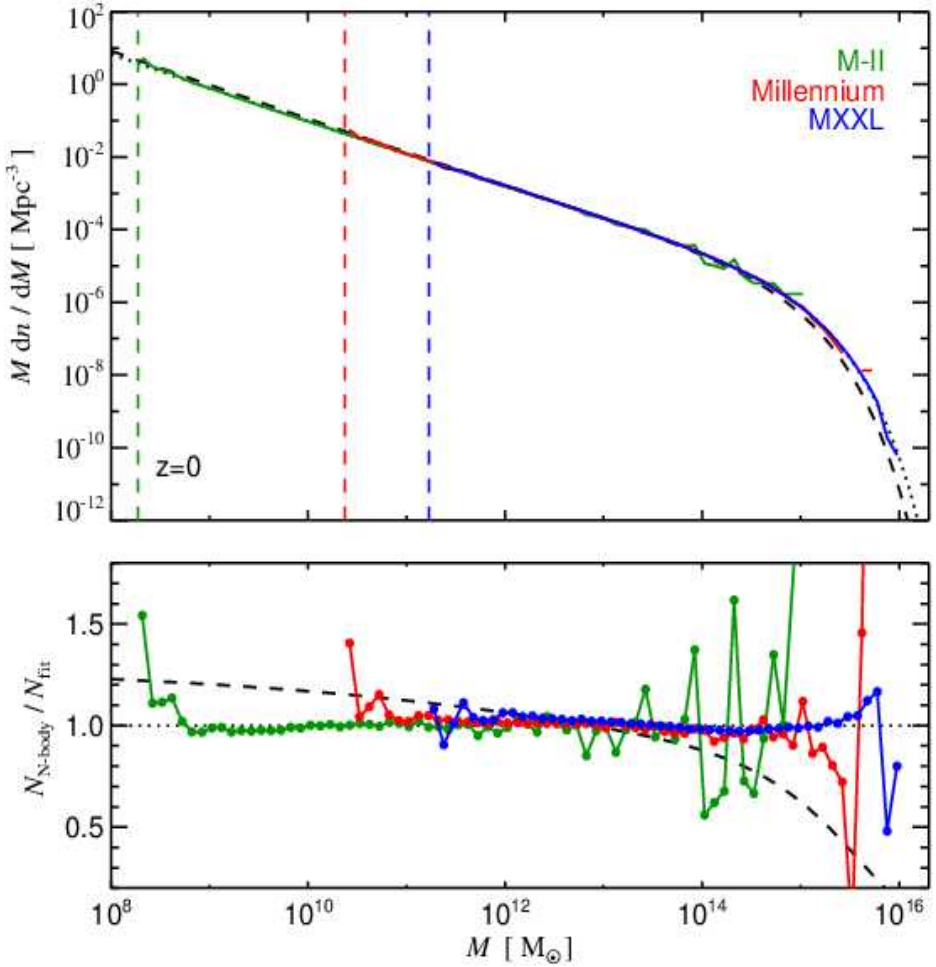
Number of particles is not precision

Force and time integration parameters do matter for the internal properties of DM haloes



Comparison with previous MS



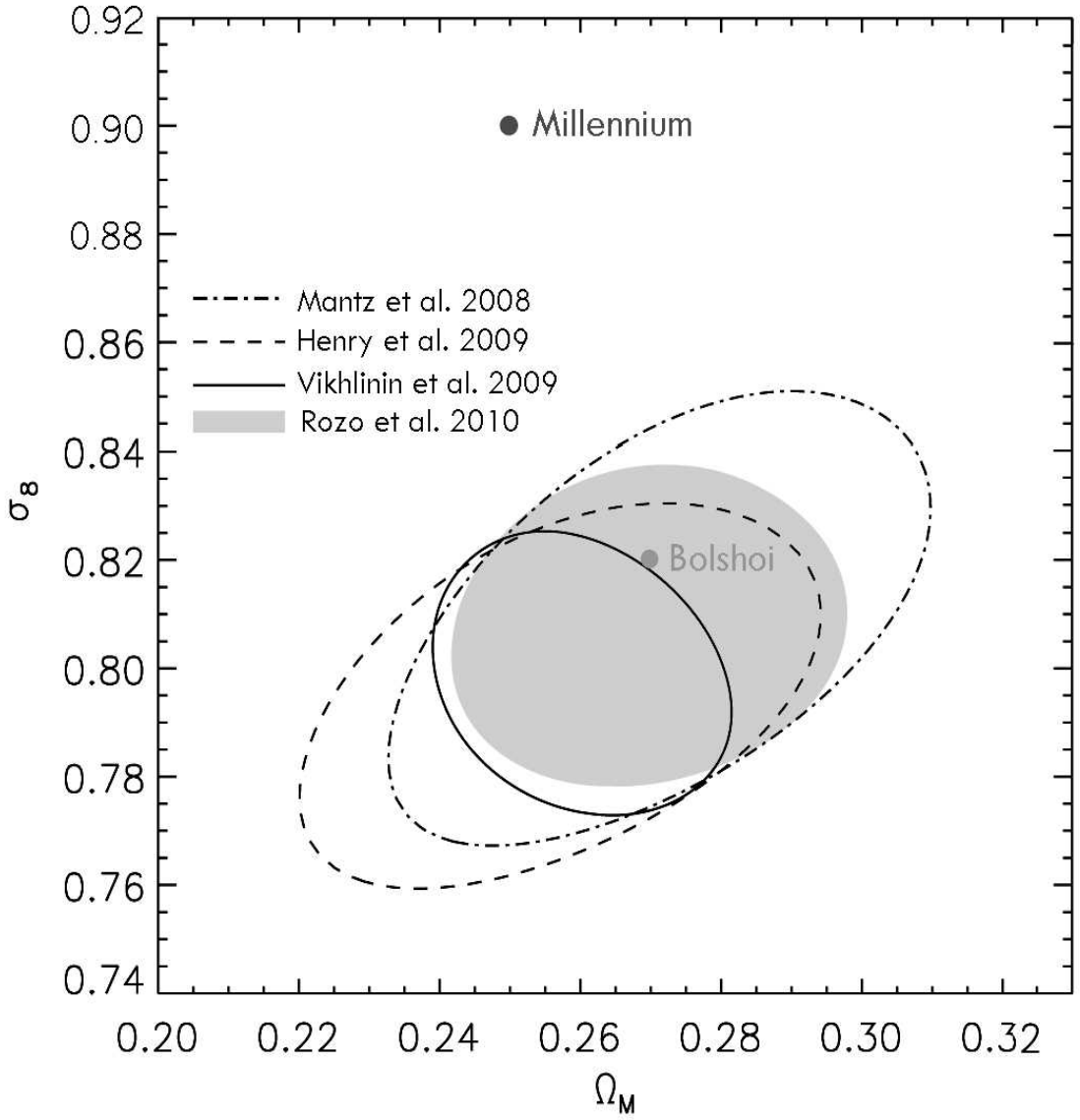


Combining these simulations we have predictions for the LCDM paradigm;

… from a 7kpc up to 4 Gpc.

… from 10^8 up to 10^{16} Mo haloes.

Hey, but you've got the wrong cosmology!



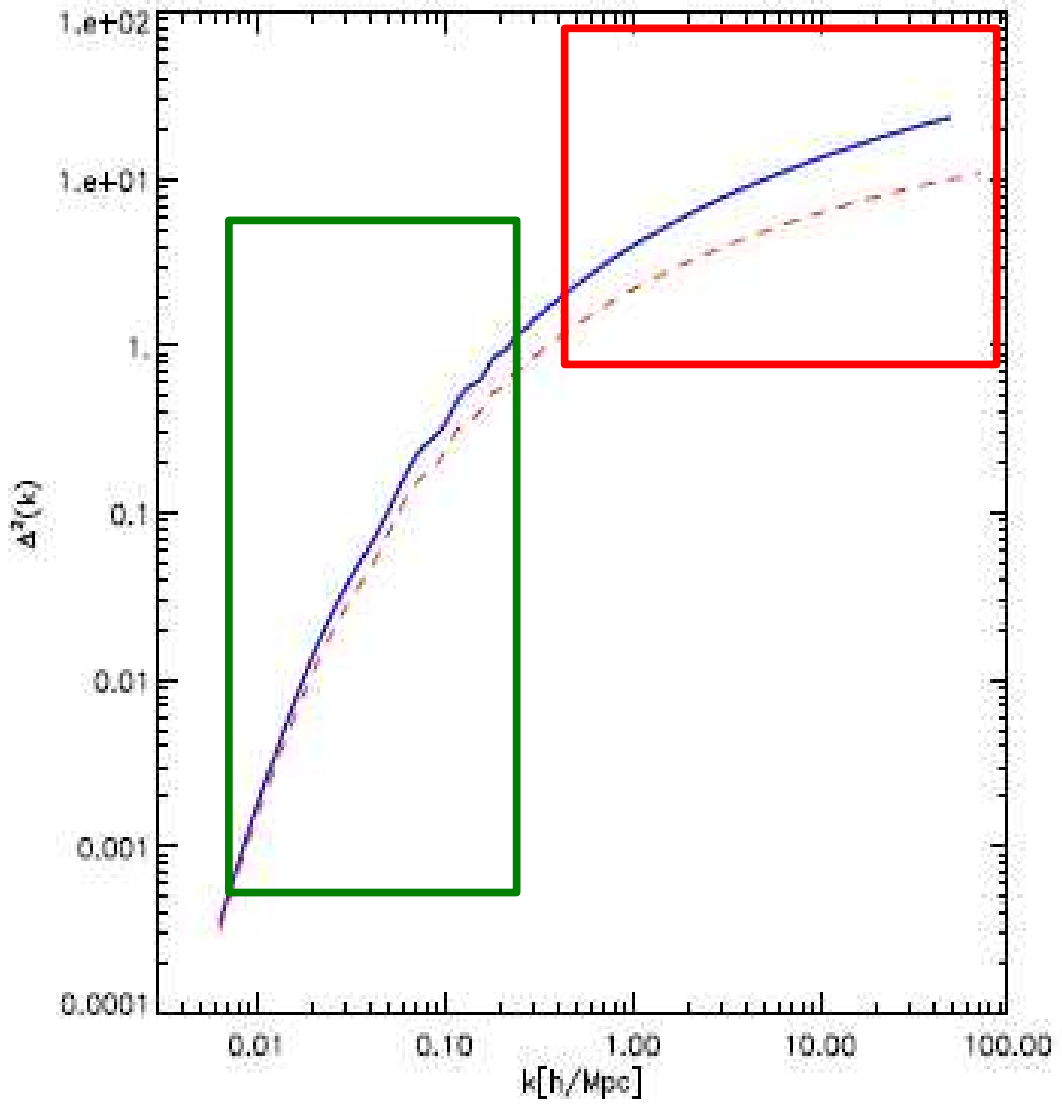
Klypin et al 2011

Most of the tension is due to a somewhat high value for sigma_8

How does cosmology affect structure formation?

Responsible for nonlinear structures

Information about early physics and cosmology



Can we mimic structure formation on different cosmologies?

The variance of the linear field as a function of scale determines all the properties of nonlinear objects

$$\sigma^2(R) = \langle \delta^2(\vec{x}; R) \rangle = \int \mathrm{d}\ln k\ \Delta^2(k) |W(k; R)|^2$$

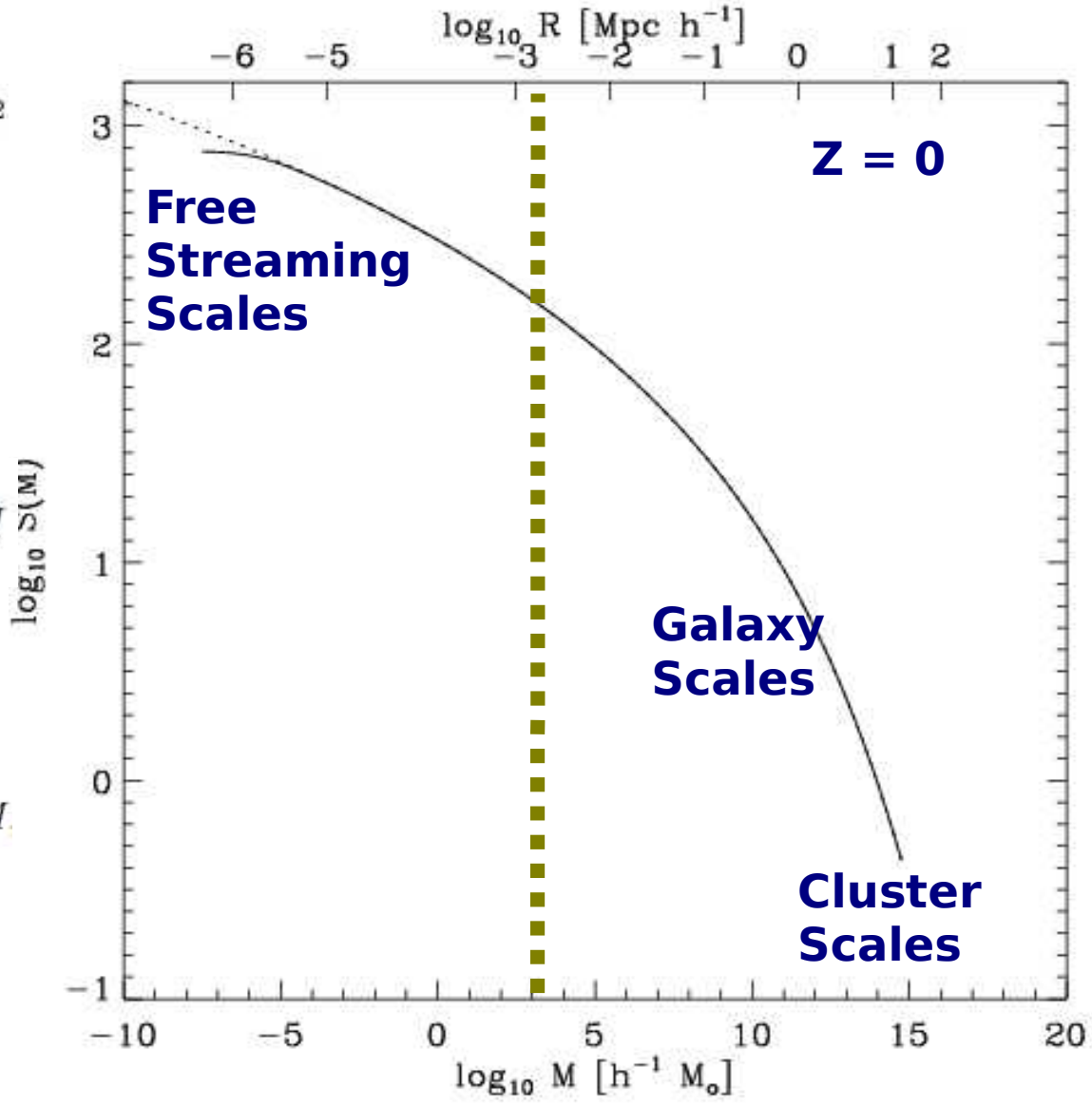
$$P(\delta; R)\mathrm{d}\delta = \frac{1}{\sqrt{2\pi\sigma^2(R)}} \exp[-\delta^2/2\sigma^2(R)]\mathrm{d}\delta$$

$$\frac{\mathrm{d}n}{\mathrm{d}M}\,\mathrm{d}M = \sqrt{\frac{2}{\pi}} \frac{\rho_\mathrm{M}}{M^2} \frac{\delta_\mathrm{c}}{\sigma} \left| \frac{\mathrm{d}\ln\sigma}{\mathrm{d}\ln M} \right| \exp\left(-\frac{\delta_\mathrm{c}^2}{2\sigma^2} \right) \mathrm{d}M$$

$$f(M|\delta_0, S_0) \left| \frac{\mathrm{d}S}{\mathrm{d}M} \right| \mathrm{d}M \equiv$$

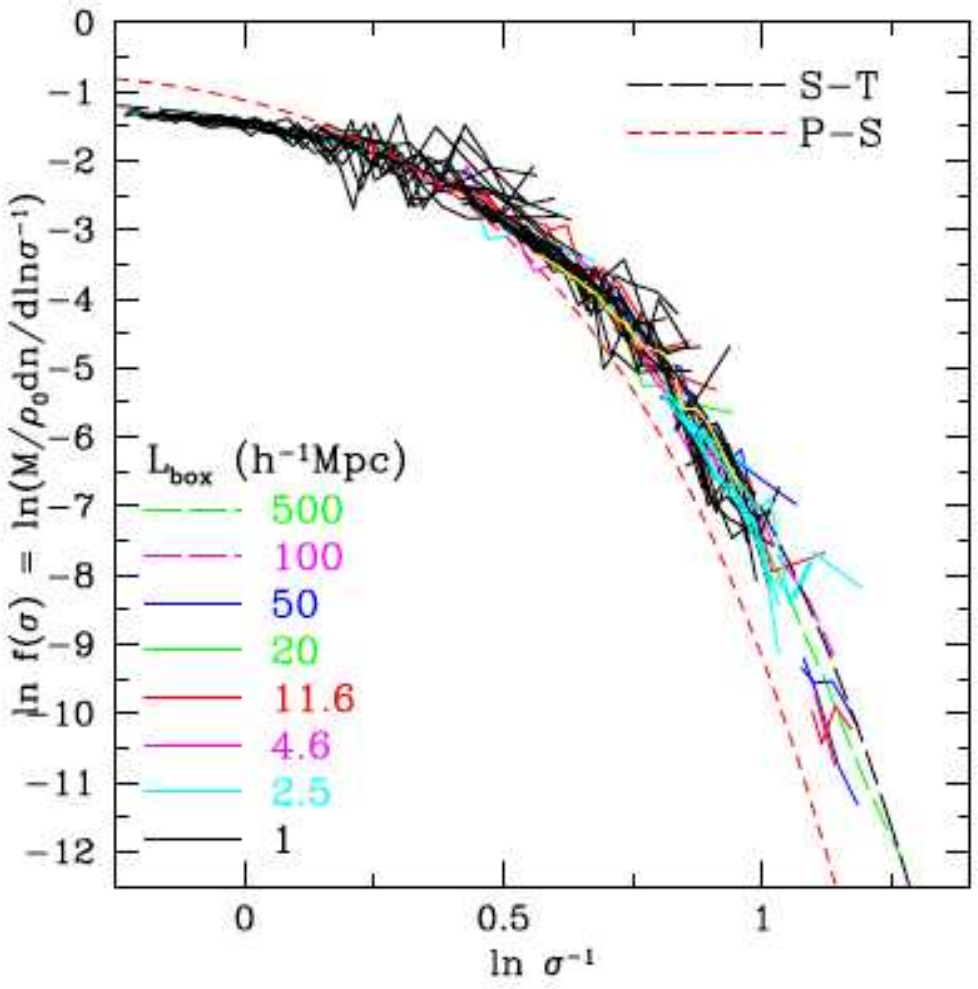
$$= \frac{1}{\sqrt{2\pi}} \frac{\delta_\mathrm{c} - \delta_0}{\Delta S^{3/2}} \left| \frac{\mathrm{d}S}{\mathrm{d}M} \right| \exp\left[-\frac{(\delta_\mathrm{c} - \delta_0)^2}{2\Delta S} \right] \mathrm{d}M$$

$$\delta_\mathrm{halo}^\mathrm{L} = \frac{\nu^2 - 1}{\delta_\mathrm{c}} \delta_0.$$



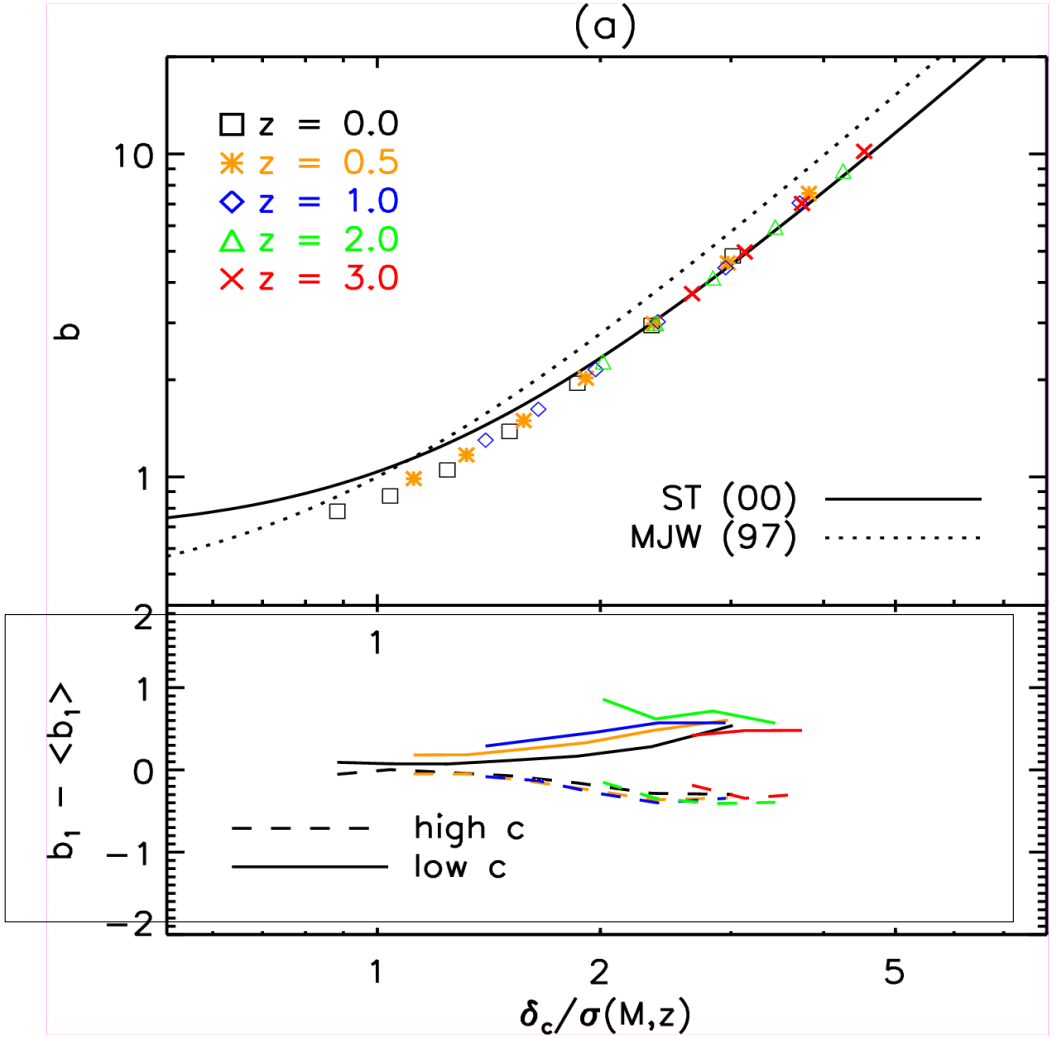
This is also true in simulations. Main halo properties depend mainly on the associated peak height.

Unconditional Mass function



Reed et al 2007

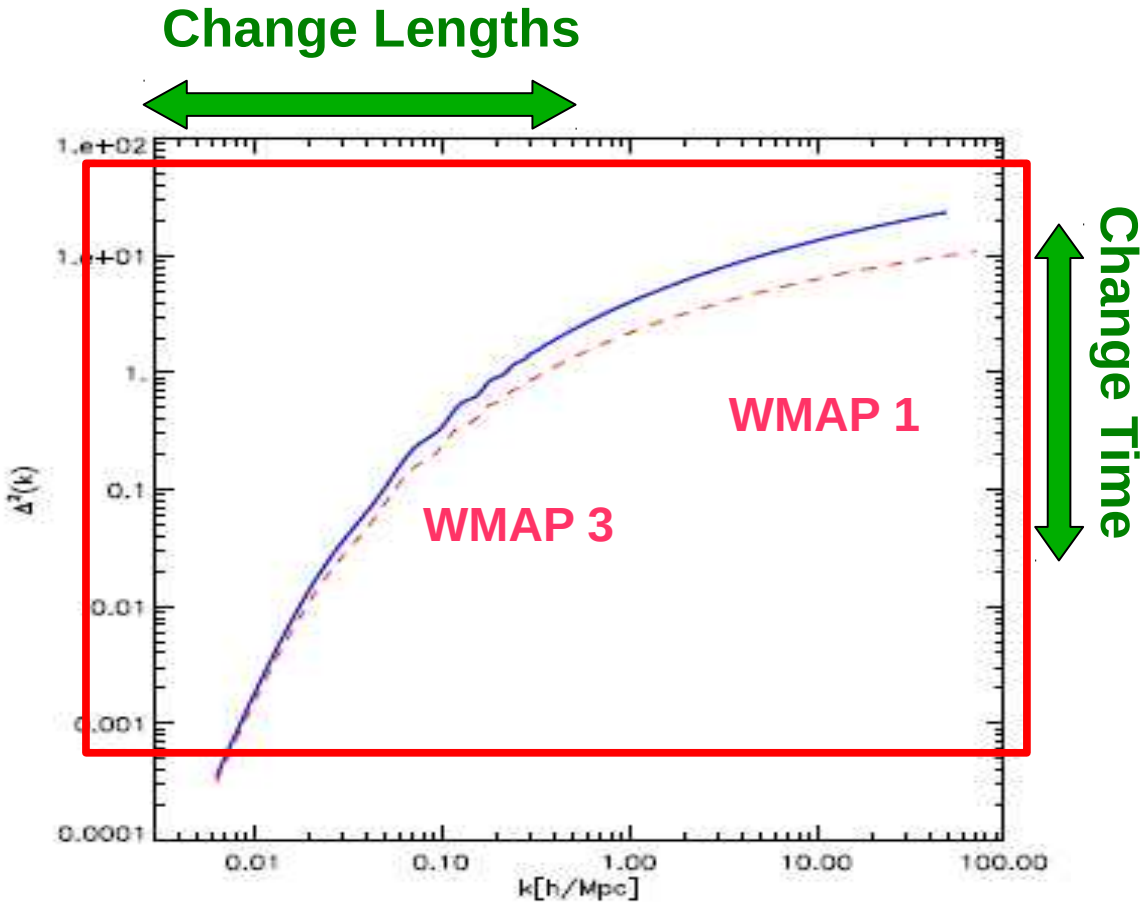
Bias parameters (b1,b2,b3 & b4)



Angulo et al 2008

One Simulation to Fit them All

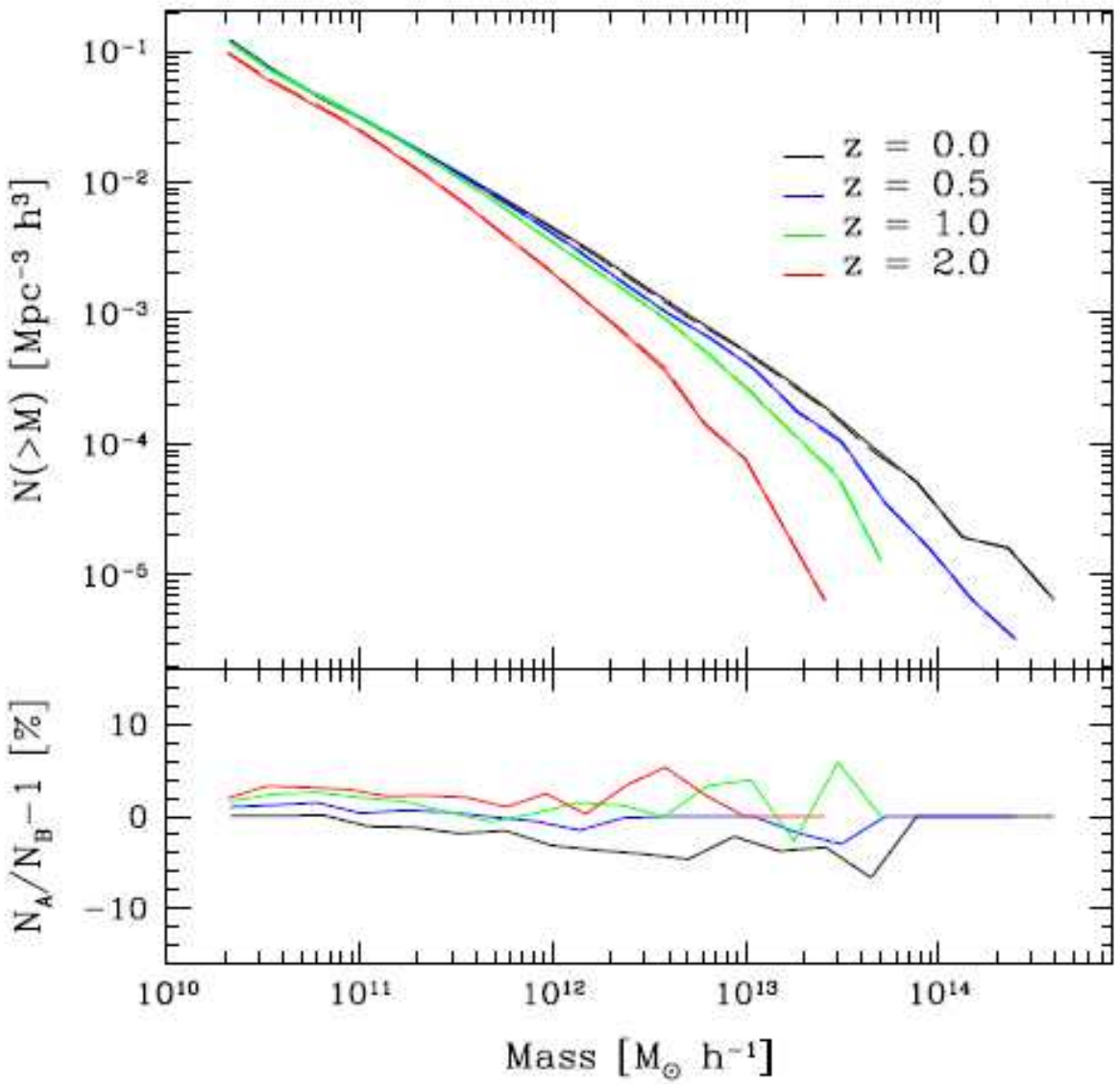
(Angulo & White 2010)



Linear and quasi linear scales can be modelled in the ZA

How well can we scale the mass function?

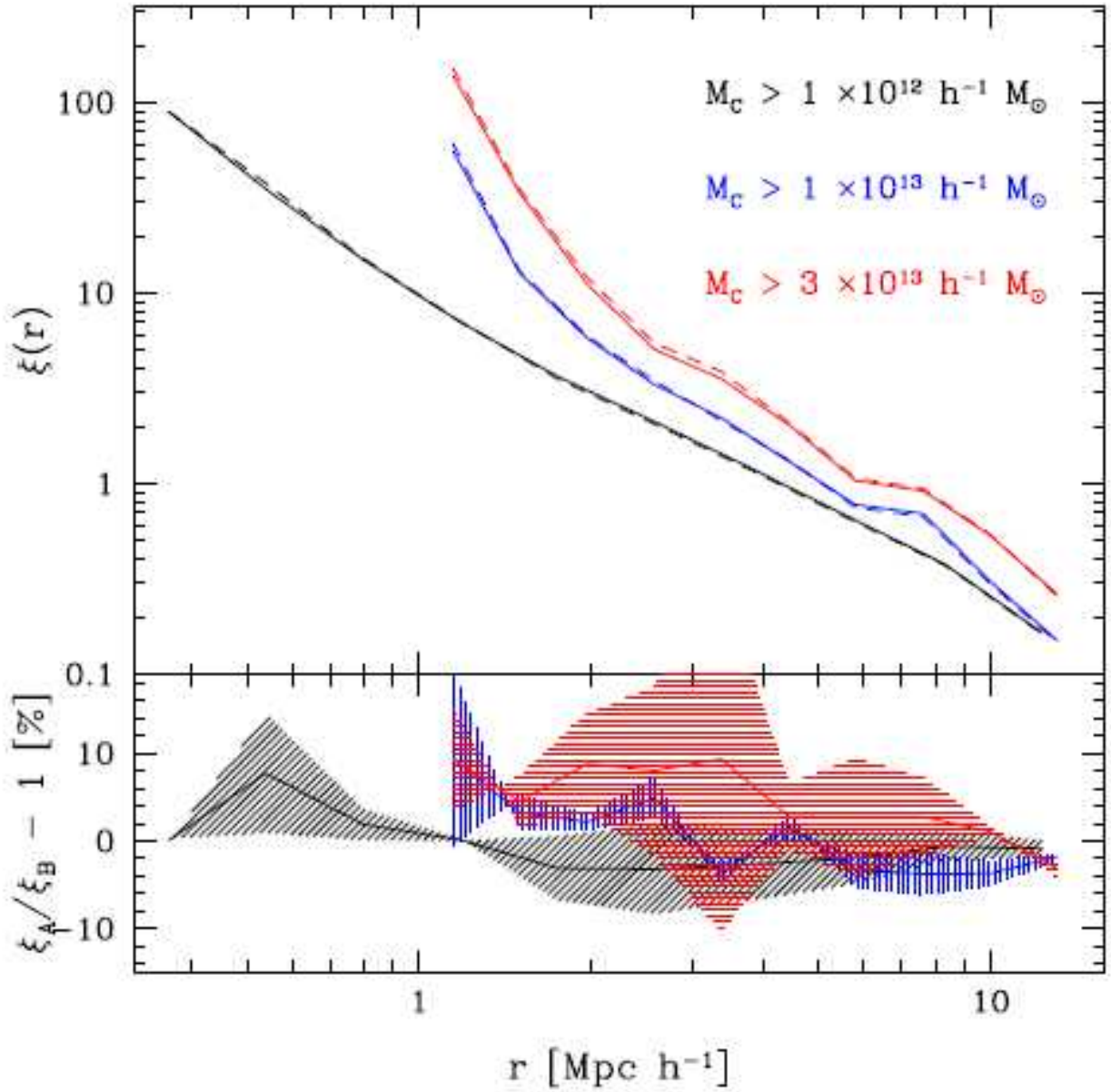
(Ruiz et al 2011)



Original: WMAP 1
Target : WMAP 5

How well can we scale the clustering?

(Ruiz et al 2011)

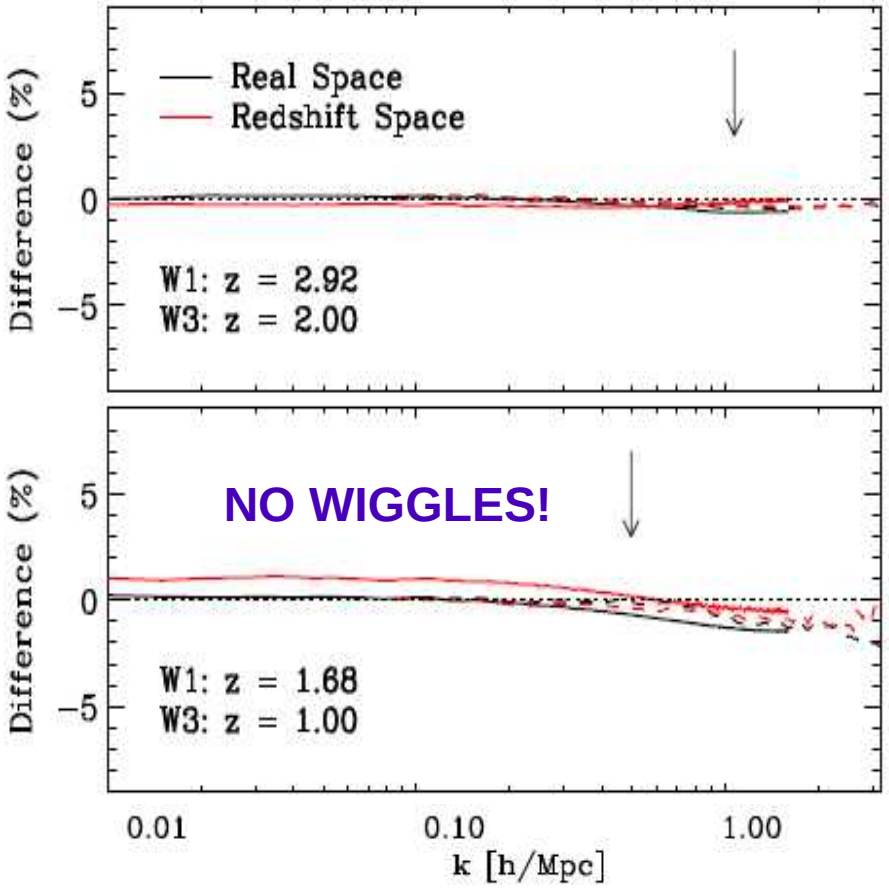


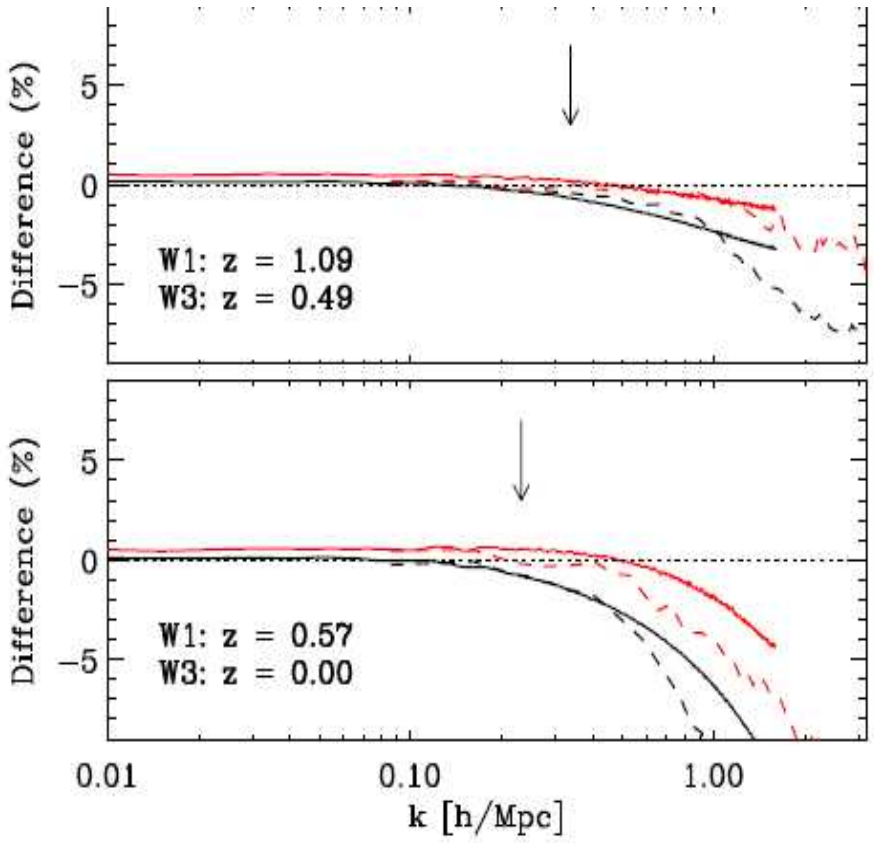
Original: WMAP 1
Target : WMAP 5

One Simulation to fit them all!

(Angulo & White 2010)

Dark Matter power spectra in the WMAP 1 & 3 cosmologies



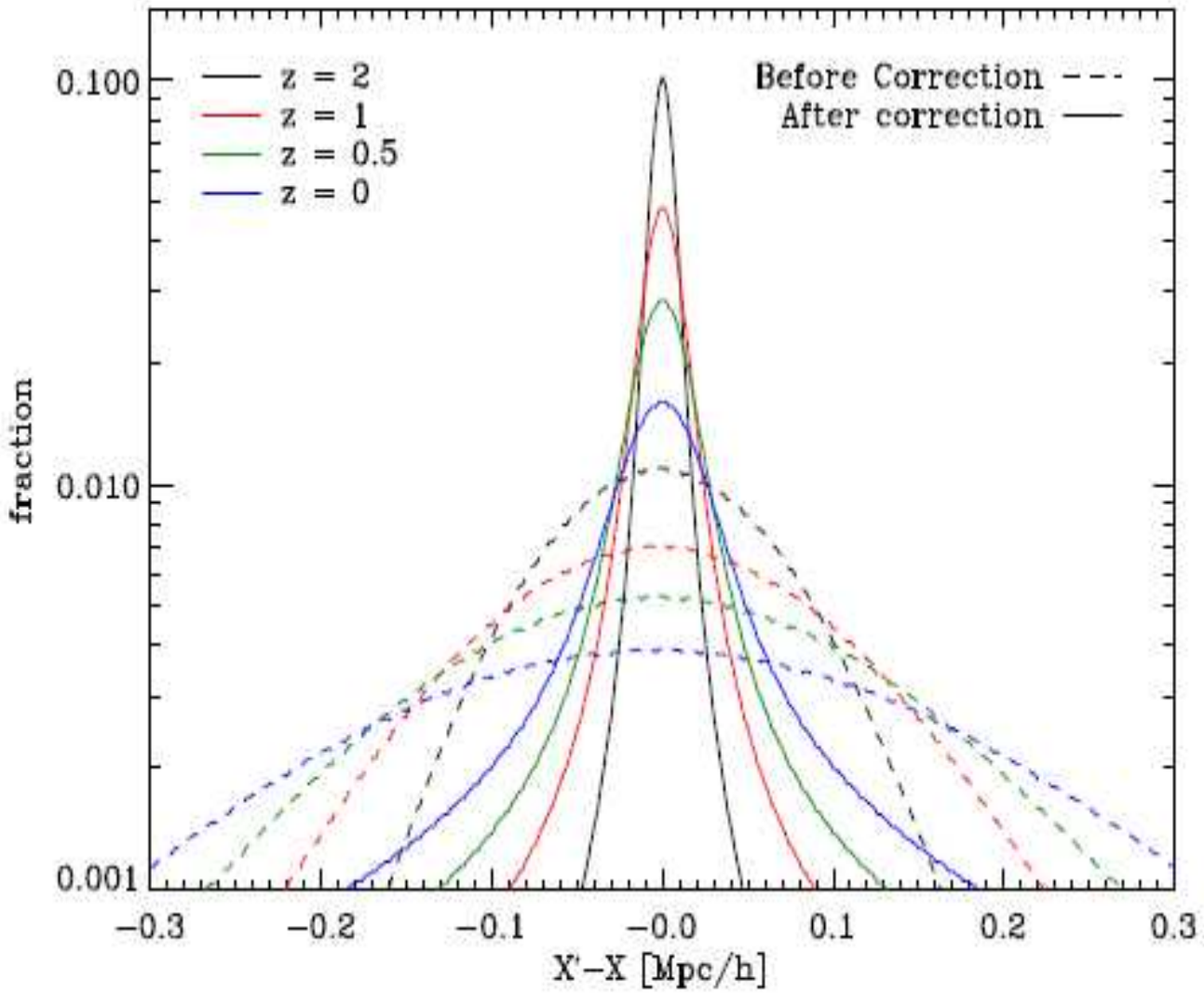


Velocities, subhaloes, M200, Concentrations, spin, merger histories are also scaled.

One Simulation to fit them all!

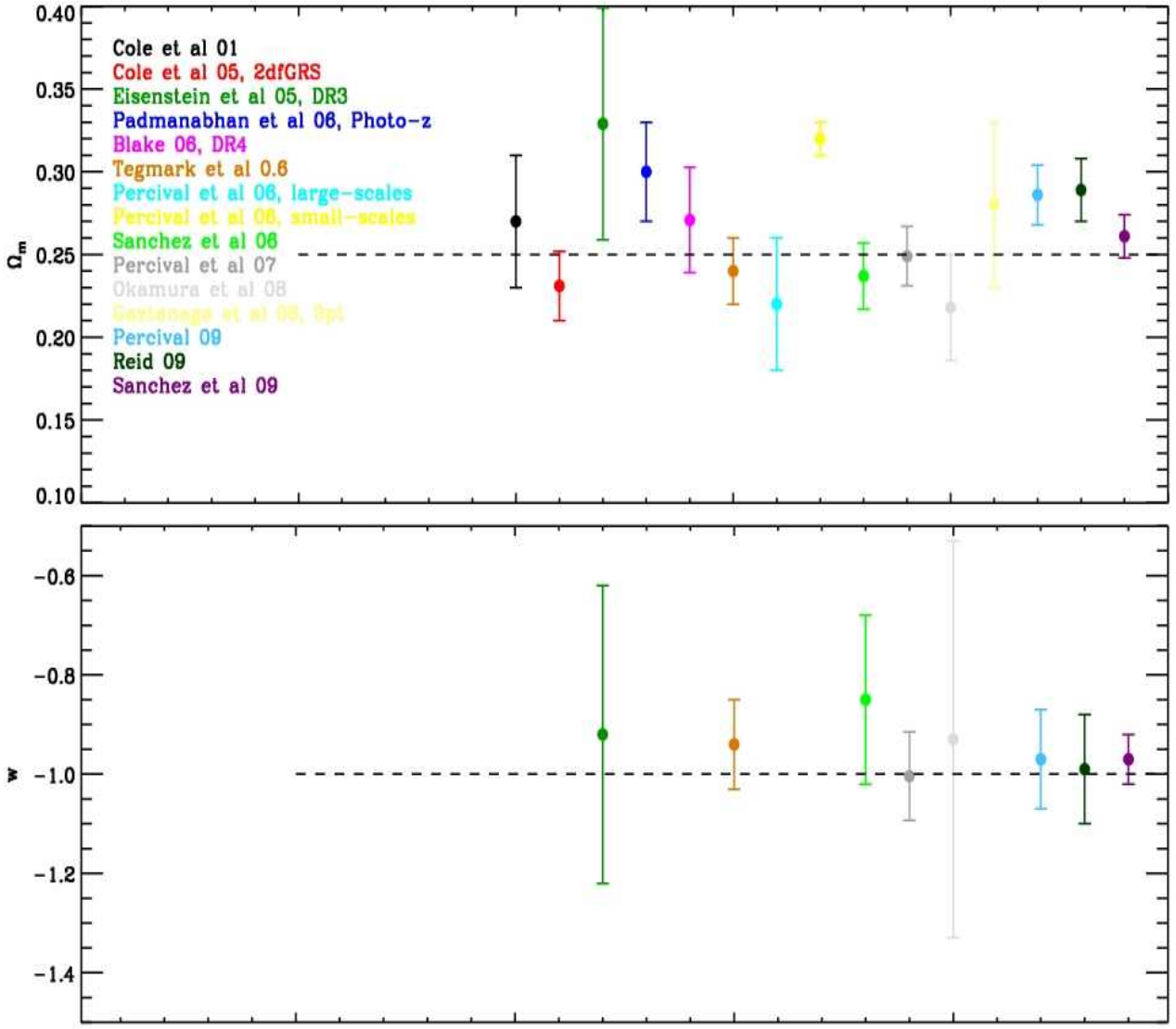
(Angulo & White 2010)

Distribution of distances of every particle in two simulations



But, is it useful?

Cosmological parameters are far from set



Angulo et al 2012b
Qi et al 2012

Goal: Relate the observed galaxy distribution to the underlying mass concentration

So far the baryonic effects on the LSS have been neglected (or coarsely estimated), but they need to be understood and taken into account in order to achieve precise constraints from missions such as EUCLID.

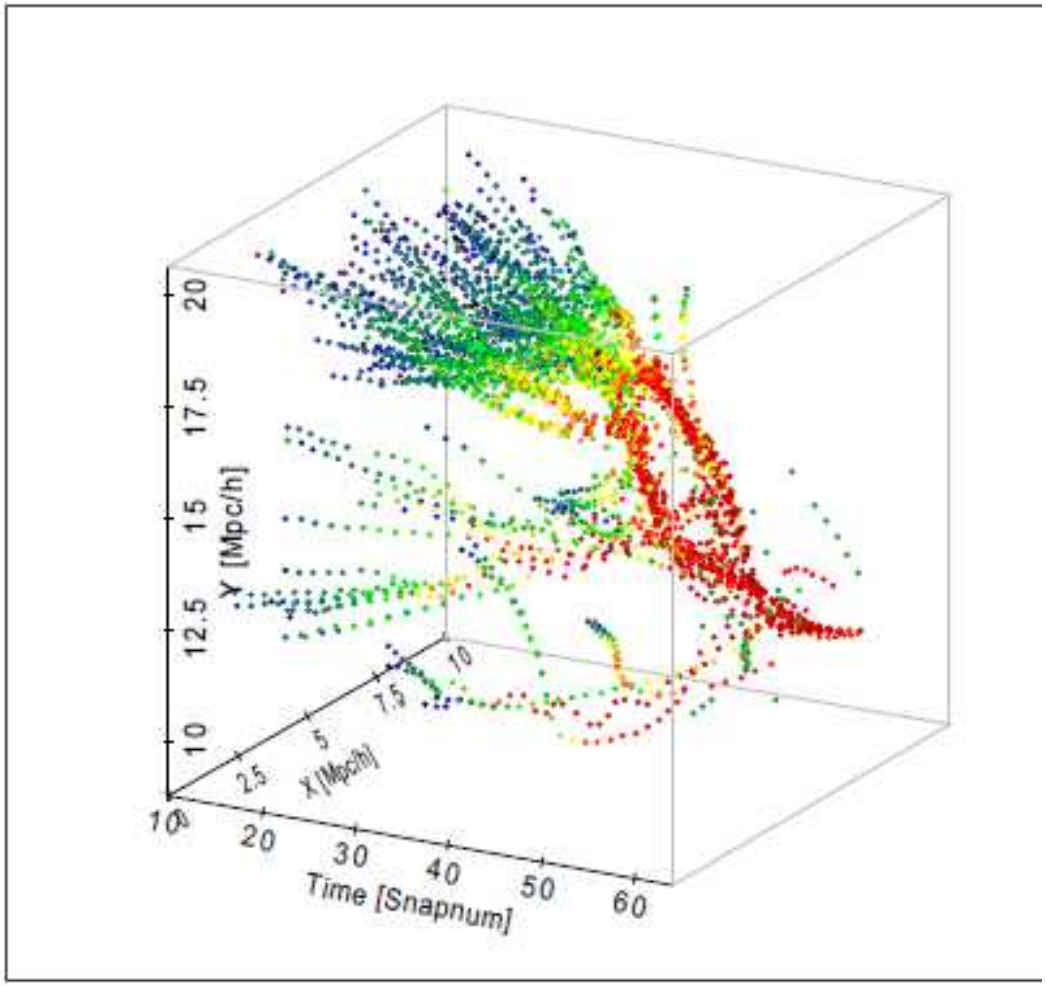
So far the baryonic effects on the LSS have been neglected (or coarsely estimated), but they need to be understood and taken into account in order to achieve precise constraints from missions such as EUCLID.

Hydro-sims: Computationally unfeasible, immature and still miss to reproduce low star formation efficiencies.

SAMs: Although implementing very drastic approximations, it is a much more successful description of galaxy formation.

Semi-analytic galaxies on the MXXL

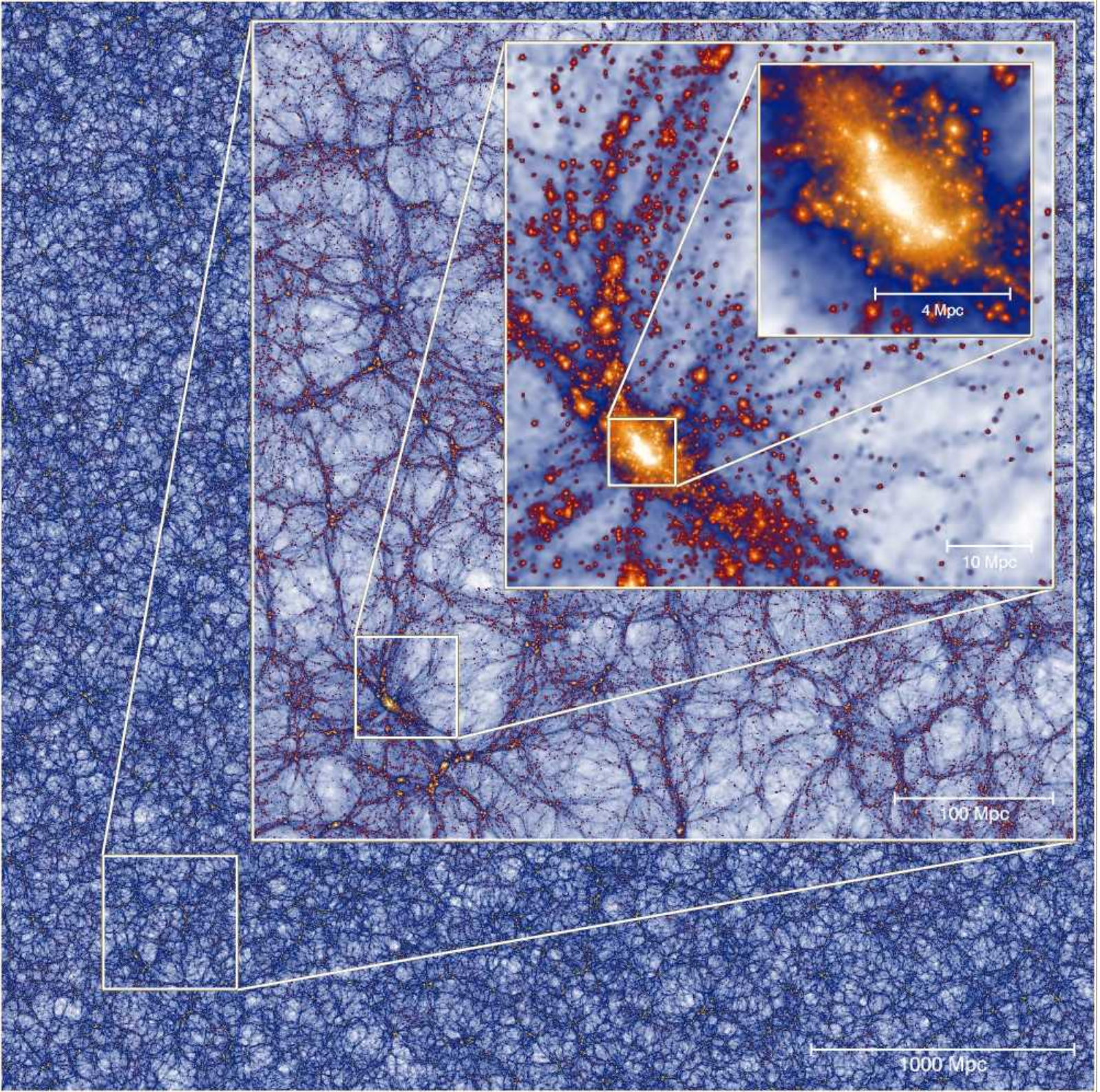
A physically motivated model for the properties of galaxies and their spatial distribution

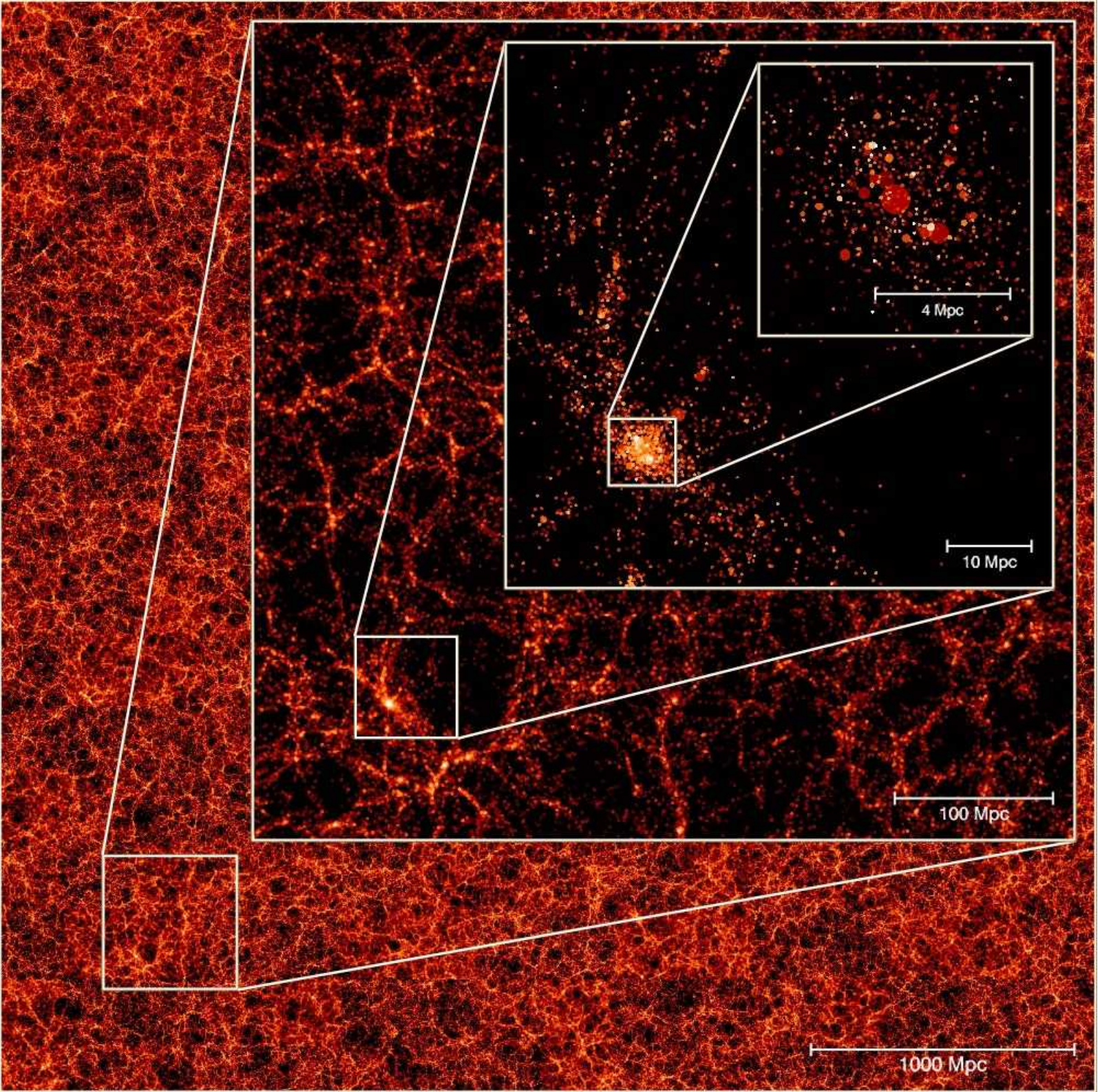


Treat baryonic physics within the evolving population of DM, including:

- ***-Gas cooling***
- ***-Evolution and formation of stars***
- ***-Central black hole growth***
- ***-Stellar and AGN feedback***
- ***-Mergers***
- ***-Metals***

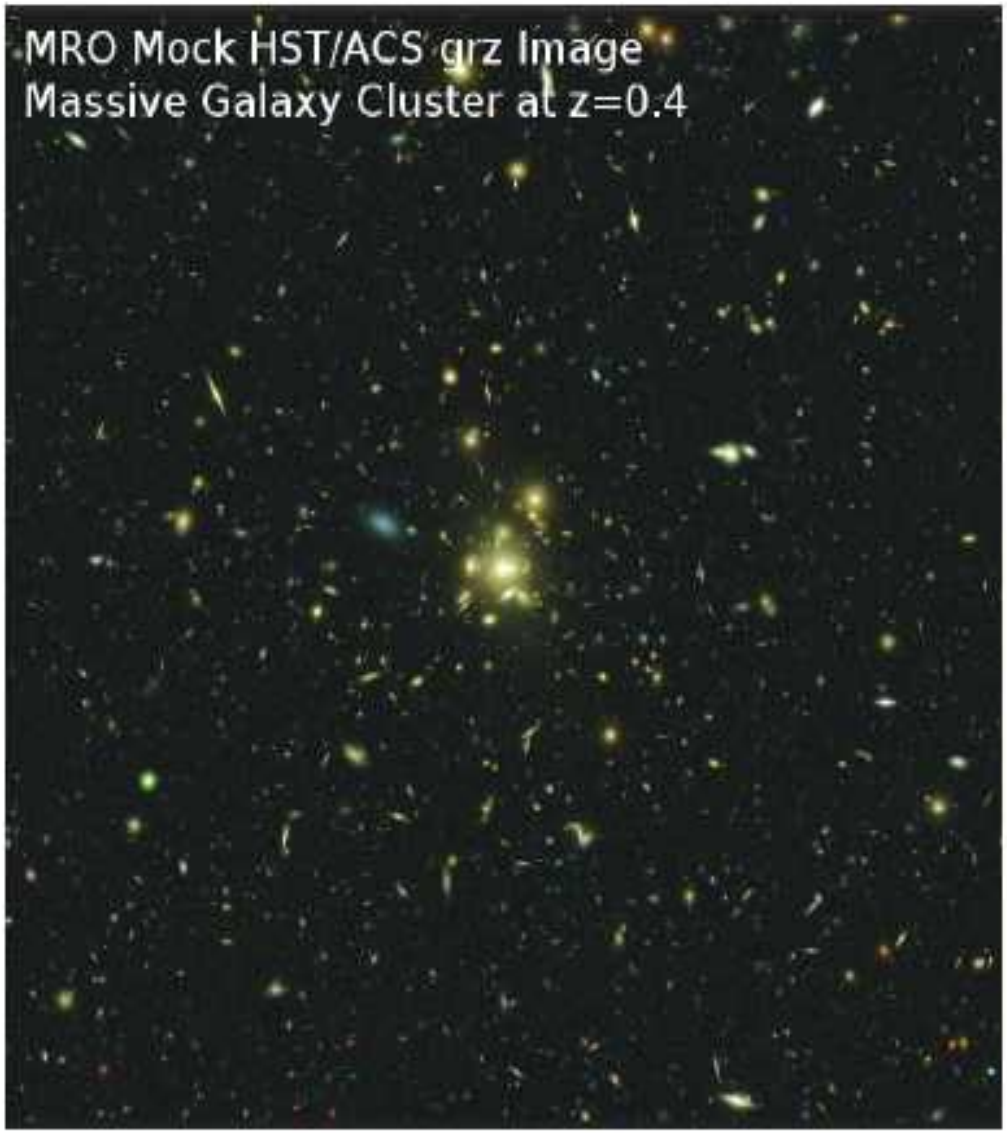
Outcome:
Time-evolution for observable properties of galaxies, including their scale-dependent bias


4 Mpc
10 Mpc
100 Mpc
1000 Mpc


4 Mpc
10 Mpc
100 Mpc
1000 Mpc

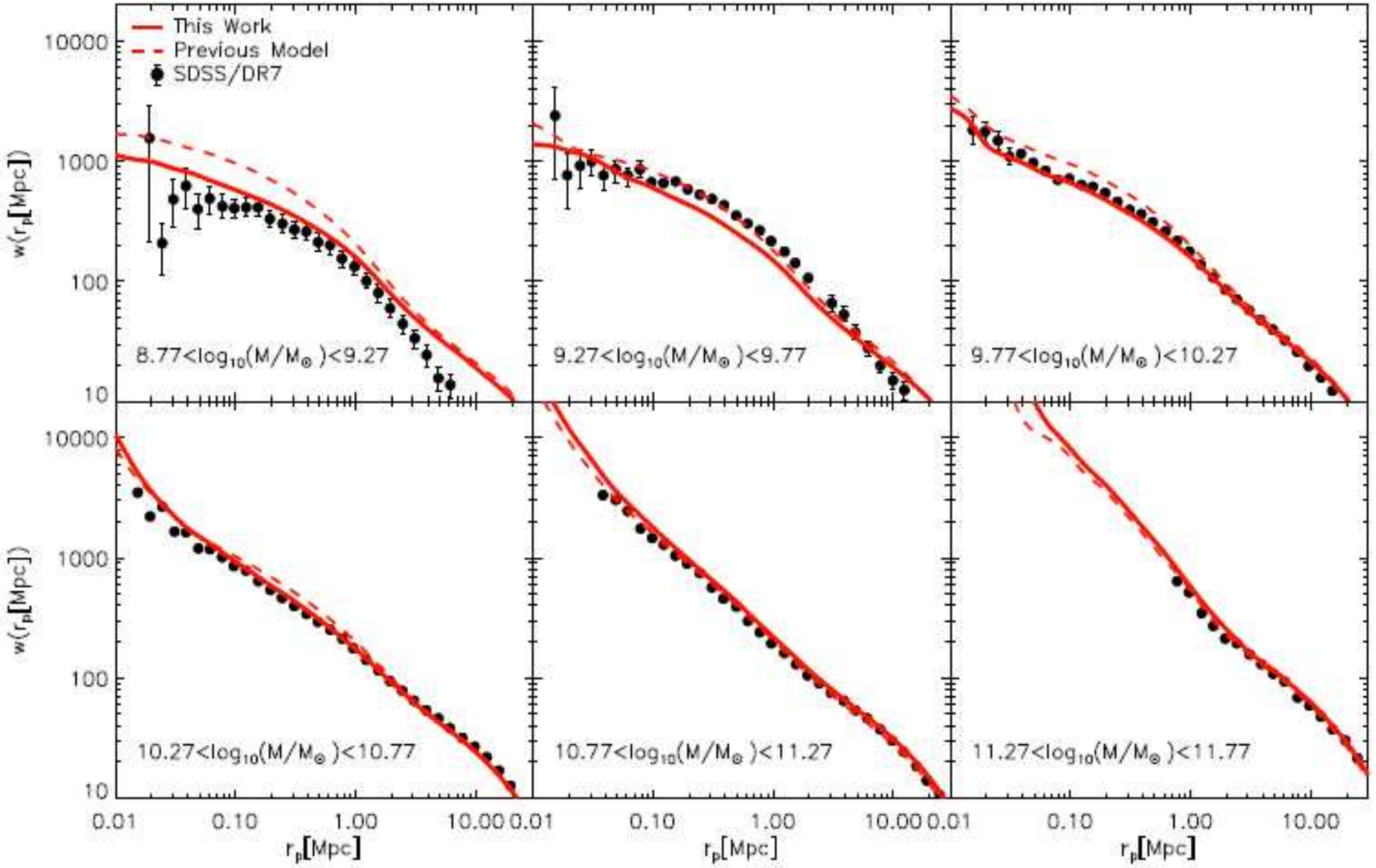
The Millennium run Observatory

(Overzier, Lemson, RA, Henriquez, Marleau & White, 2012)



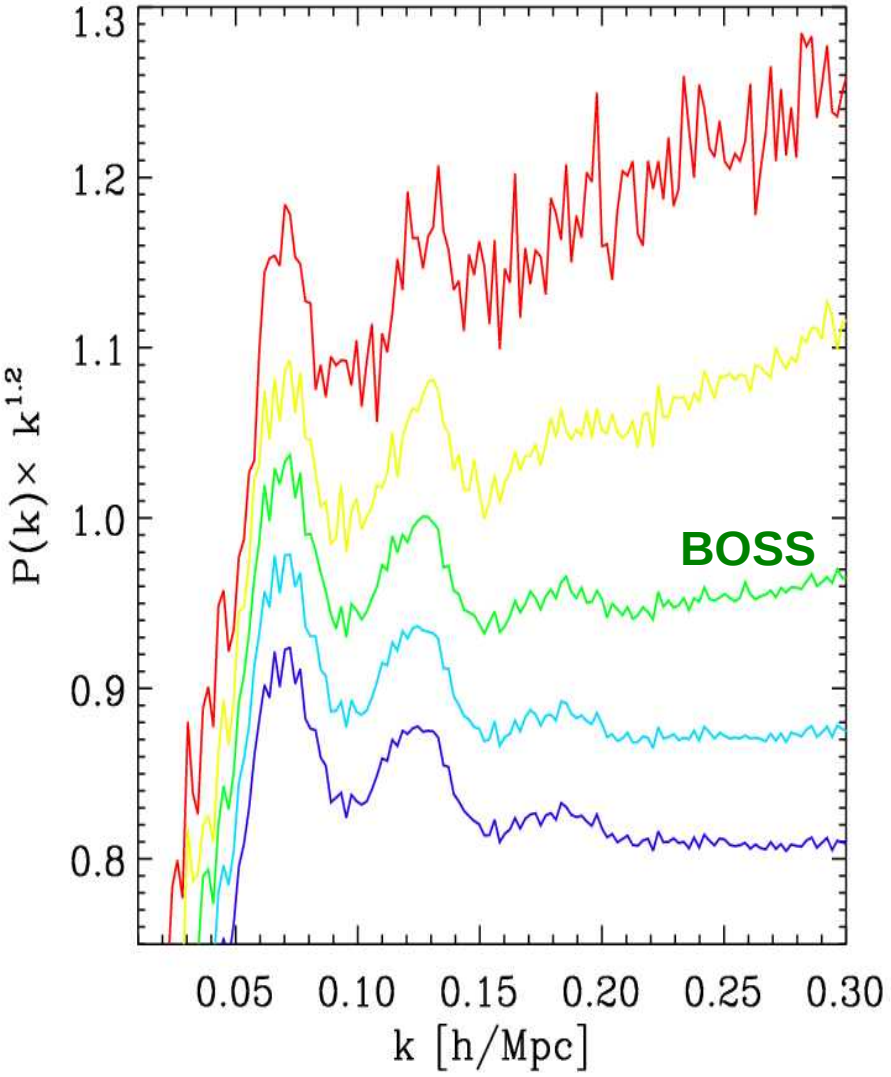
COMPARISON BETWEEN THE PROJECTED CORRELATION FUNCTION IN SDSS AND IN THE MILLENNIUM SIMULATION + SAM

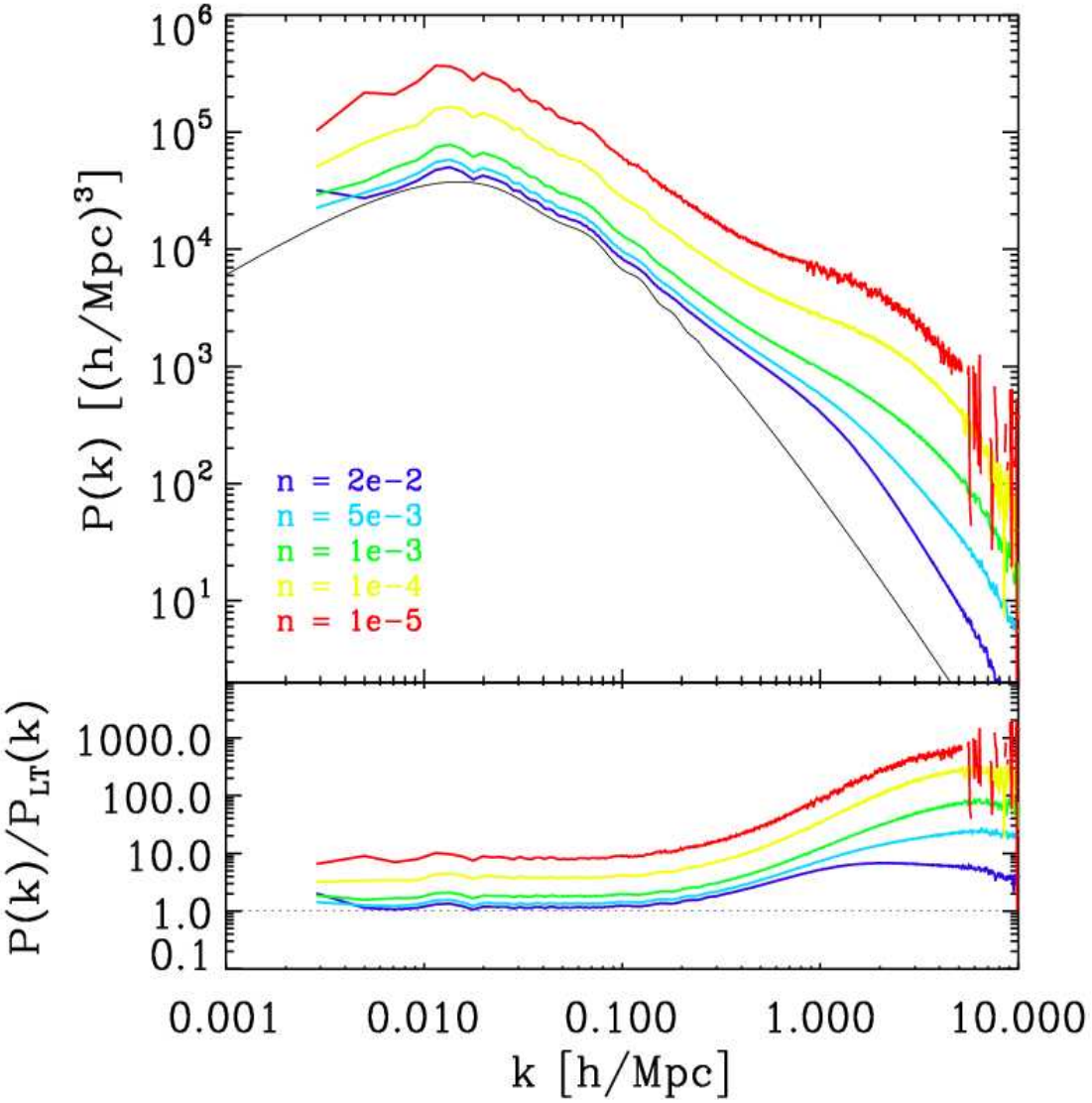
Henriques, White, Thomas, Lemson, Guo & Angulo 2012, in prep)



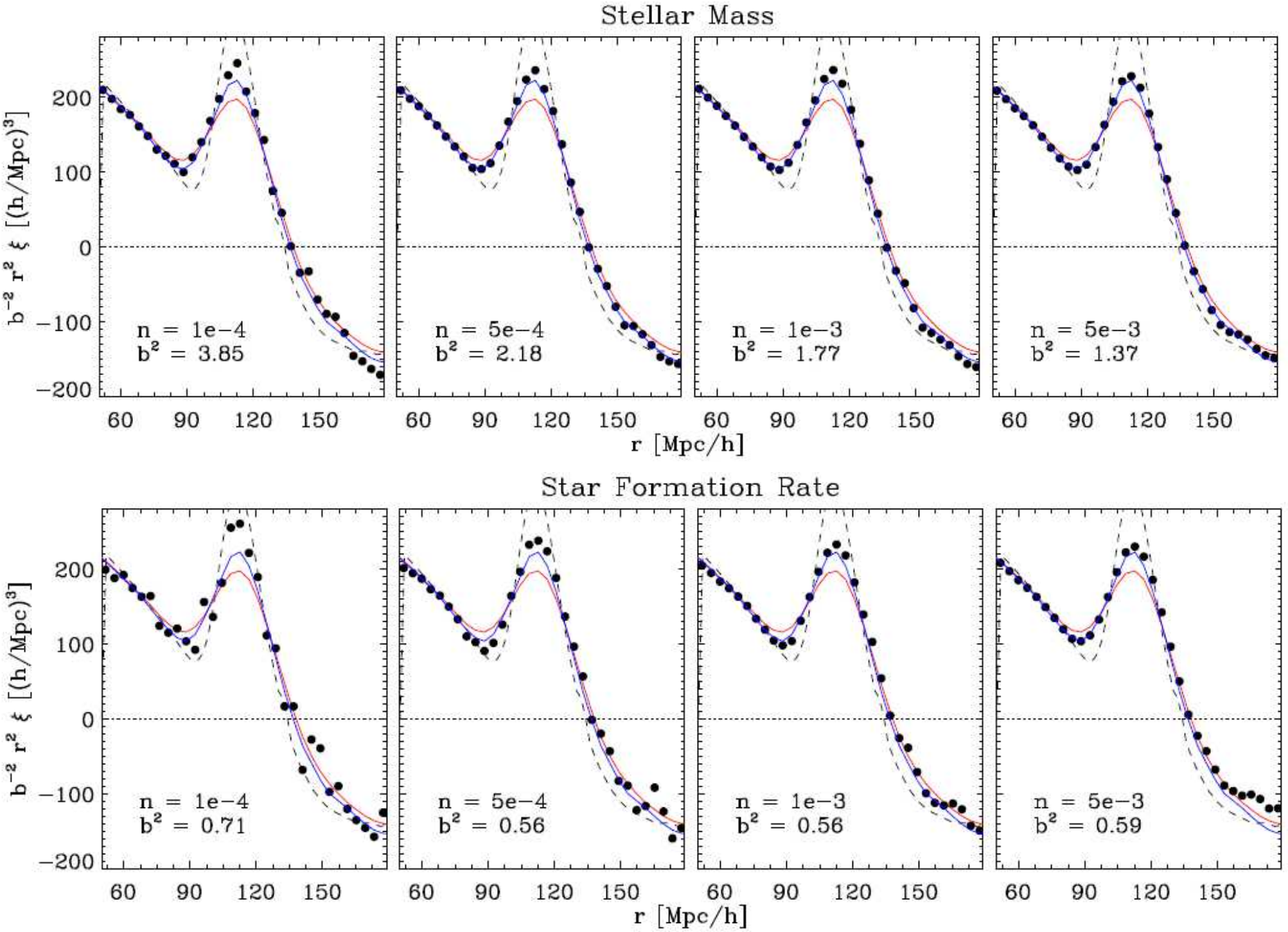
Different galaxy catalogues in the MXXL simulation trace the BAO features with a scale-dependent bias

POWER SPECTRA OF THE GALAXY DISTRIBUTION AT Z=0 FOR DIFFERENT SPACE DENSITIES

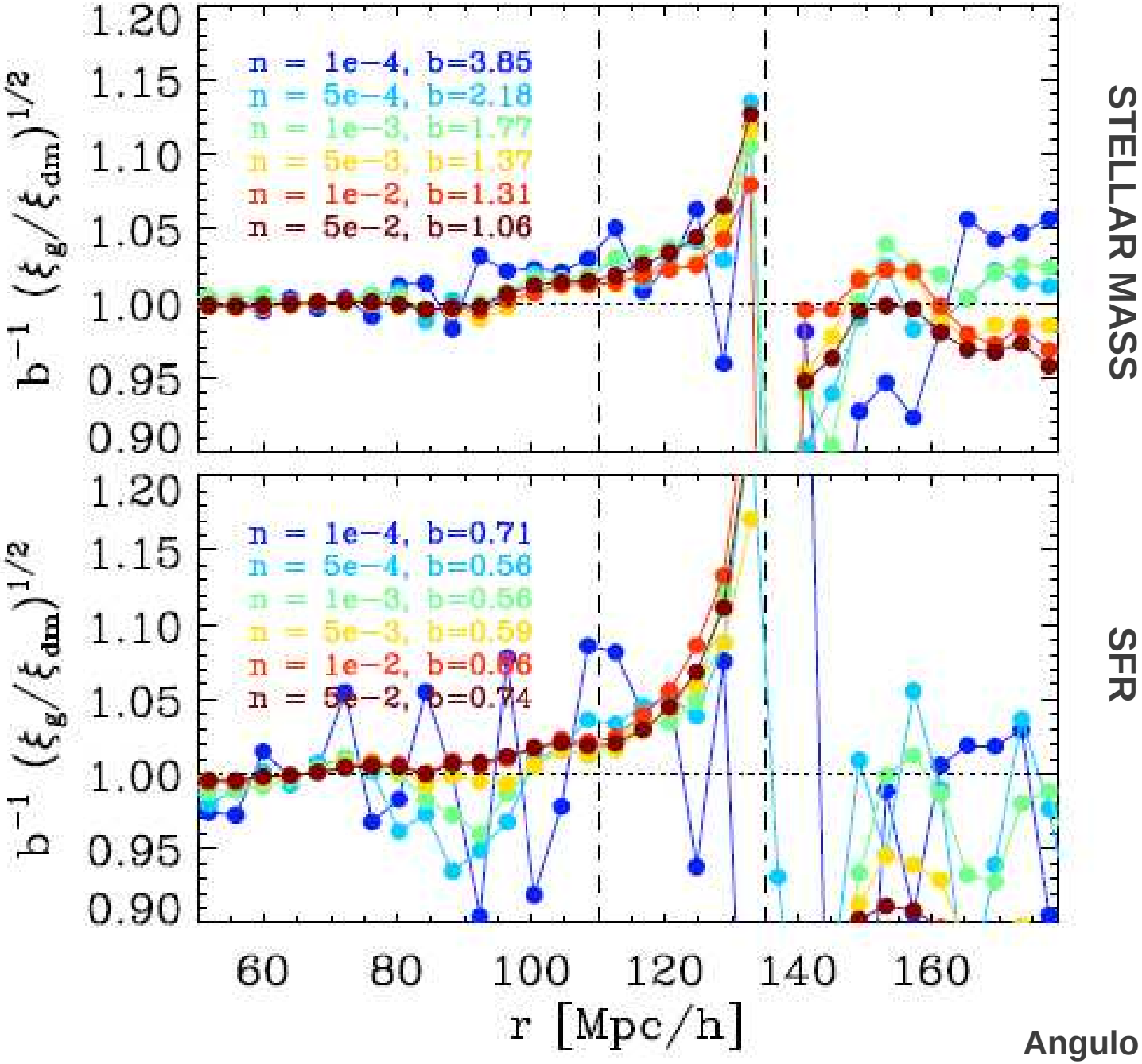




Angulo 2012, in prep



Galaxies show an excess clustering at the BAO position of 2%



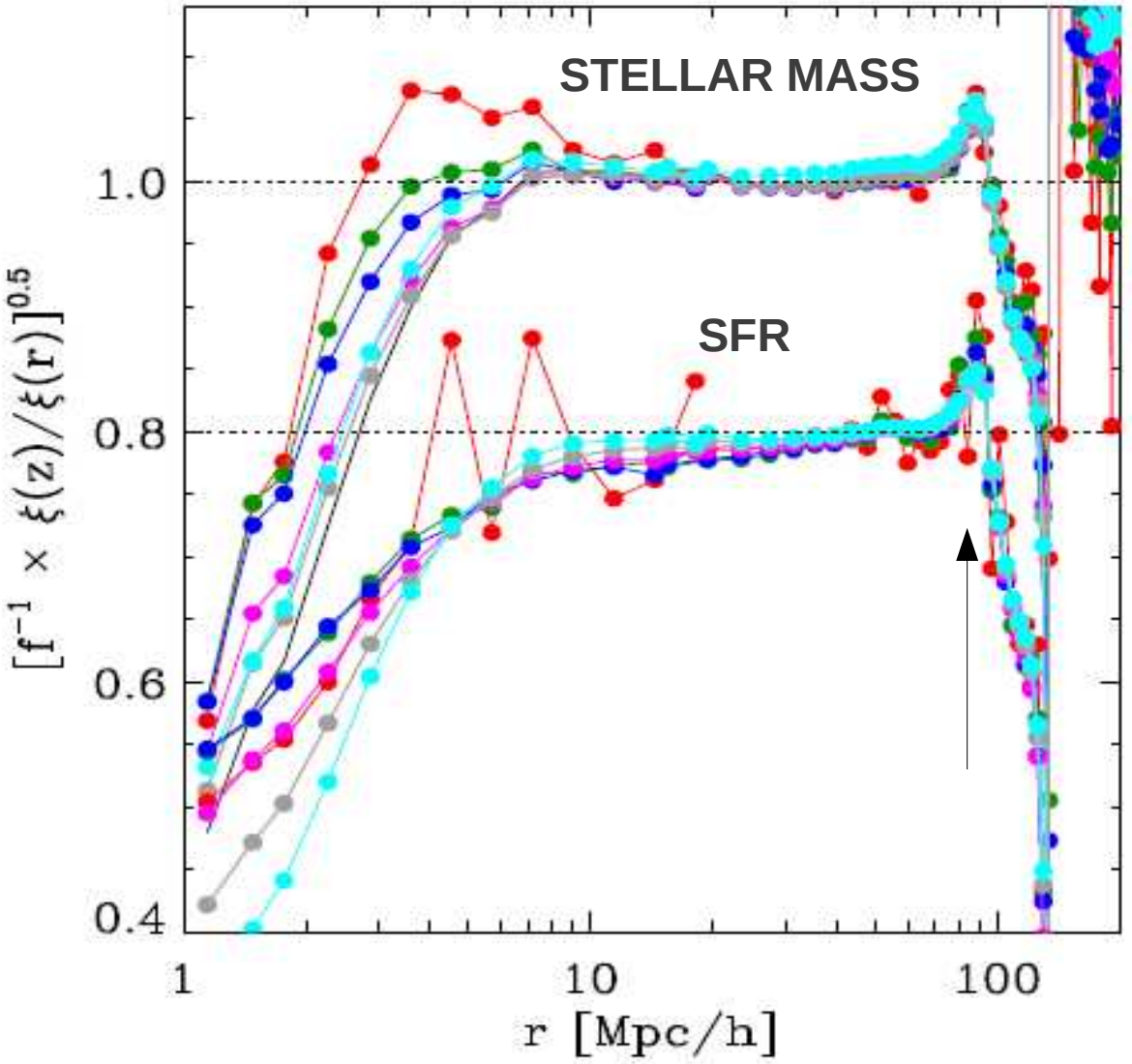
Angulo 2012, in prep

Different galaxy catalogues show departures from the Kaiser formula in different manners and scales.

THE RATIO OF THE REDSHIFT- TO THE REAL-SPACE CORELATION FUNCTION FOR DIFFERENT SAMPLES.

Angulo 2012, in prep

Small and intermediate-scale clustering properties depend not only on gravity and cosmological parameters, but also on the baryonic physics.



Final message

The appearance of the LSS is slightly different for different galaxy samples.

New generation of simulations will be essential to interpret future high-precision measurements and for an optimal data exploitation.

Thus, they will be crucial to understand more about the main ingredients of our Universe

The Millennium XXL simulation is the most realistic prediction of the statistical properties and spatial distribution of the galaxy population on the largest scales.