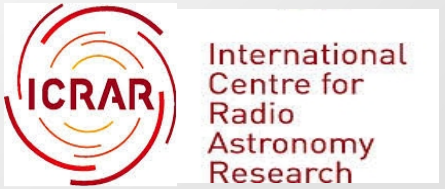
HALOGEN

A tool for fast generation of mock catalogs

by Santiago Ávila and Steven Murray
with A. Knebe, J. García-Bellido,
C. Power & A. Robotham



Introduction

- Large Scale Structure (LSS) is one of the fundamental pillars of LCDM
- And a very hot source of observations: BOSS, DES, PAU, DESi, Euclid, …
- Whereas CMB can tell us more about primordial universe, we need LSS to learn more about late effects: Dark Energy
- BAO: a very clean probe.

Introduction

- We need to compare observations to theory, but analytic approaches in LSS are very limited → we need numerical simulations.
- To constrain models we need ~1000 mock catalogs to control systematic errors and compute covariance matrices.
- To produce a mock catalog:
  - (1) N-Body Simulation → DM-field
  - (2) Halo finder → Halos
  - (3) HOD/SAM/etc → Galaxies
- N-Body Simulations take really long. Not *feasible* to run hundreds of them.
- (1) is very precise, but
  (2) has ~10% scatter in mass function (Knebe et al. 2013)
  (3) can present a huge scatter in 2-point function (50% in Benson et al. 2001)

Introduction

- Do we need that much precision in the Dark Matter field?

Introduction

- Do we need that much precision in the Dark Matter field?
  Maybe Not:
  - PThalos (Scoccimarro & Seth, 2001; Manera et al. 2013 & 2014)
  - PINNOCHIO (Monaco, Theuns & Taffoni, 2002; Monaco & Theuns, 2013; ...)
  - COLA (Tassev, Zaldarriaga & Eisenstein 2013; Koda, Kazin & Blake, 2013)
  - QPM (Martin, Tinker & McBride, 2014)
  - PATCHY (Kitaura, Yepes & Prada, 2014, Kitaura et al. 2014)
  - EZAmocks (Chuang, in prep)
  - HALOgen (Avila et al, in prep)

  All of them are alternative methods, compared in *Chuang et al. in prep.*

Introduction

- Do we need that much precision in the Dark Matter field?
  Maybe Not:

  - PThalos (Scoccimarro & Seth, 2001; Manera et al. 2013 & 2014)
  - PINNOCHIO (Monaco, Theuns & Taffoni, 2002; Monaco & Theuns, 2013; ...)
  - COLA (Tassev, Zaldarriaga & Eisenstein 2013; Koda, Kazin & Blake, 2013)
  - QPM (Martin, Tinker & McBride, 2014)
  - PATCHY (Kitaura, Yepes & Prada, 2014, Kitaura et al. 2014)
  - EZAmocks (Chuang, in prep)
  - **HALOgen (Avila et al, in prep)**

  All of them are alternative methods, compared in *Chuang et al. in prep.* (http://popia.ft.uam.es/nIFTyCosmology)

HALOGEN: the method

- The philosophy of HALOgen is to start from the simplest things and improve if needed.

- Main ingredients:
  - 2LPT ← Dark matter field
  - Analytic Mass function
  - Place halos with the correct 2-point function ← the tricky part

HALOGEN: the method

- The philosophy of HALOgen is to start from the simplest things and improve if needed.

- Main ingredients:
  - 2LPT ← Dark matter field
  - Analytic Mass function
  - Place halos with the correct 2-point function ← the tricky part

HALOGEN: the method

- Requirements:
  - A fast tool for generation of halo mock catalogs
  - with the right Halo Mass Function
  - and correct 2-point correlation function (real + k-space)

- Desired:
  - Correct higher order statistics

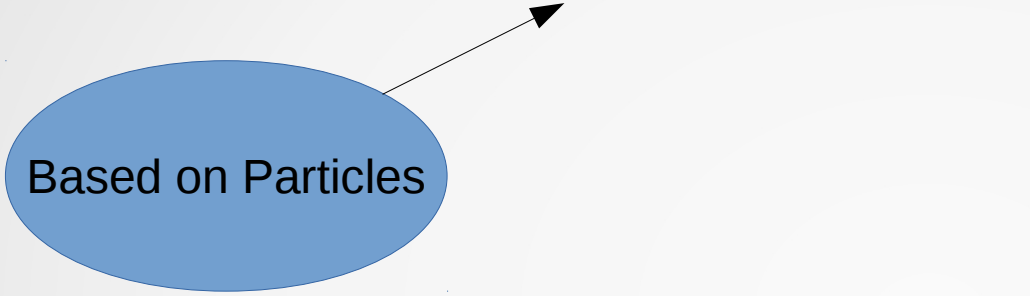
Note: we do not aim at actually *finding* halos. We *place* them with the correct statistics

2LPT

- 2LPT: **2**nd order **L**agrangian **P**erturbation **T**heory

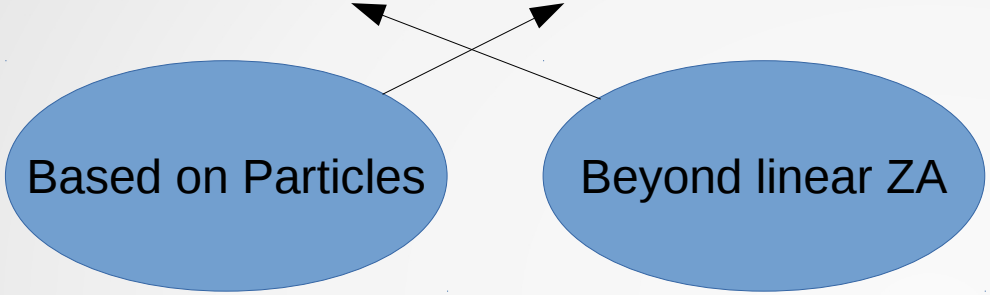
2LPT

- 2LPT: **2**nd order **L**agrangian **P**erturbation **T**heory



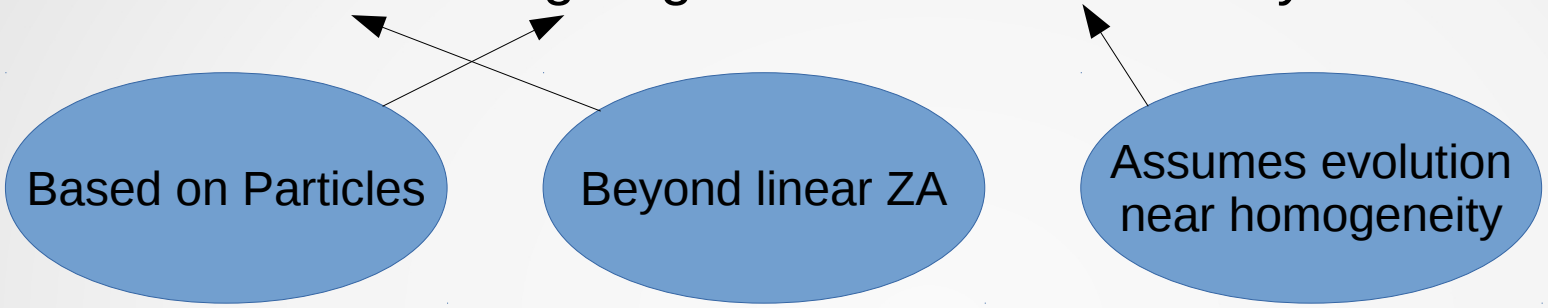
2LPT

- 2LPT: **2**nd order **L**agrangian **P**erturbation **T**heory



2LPT

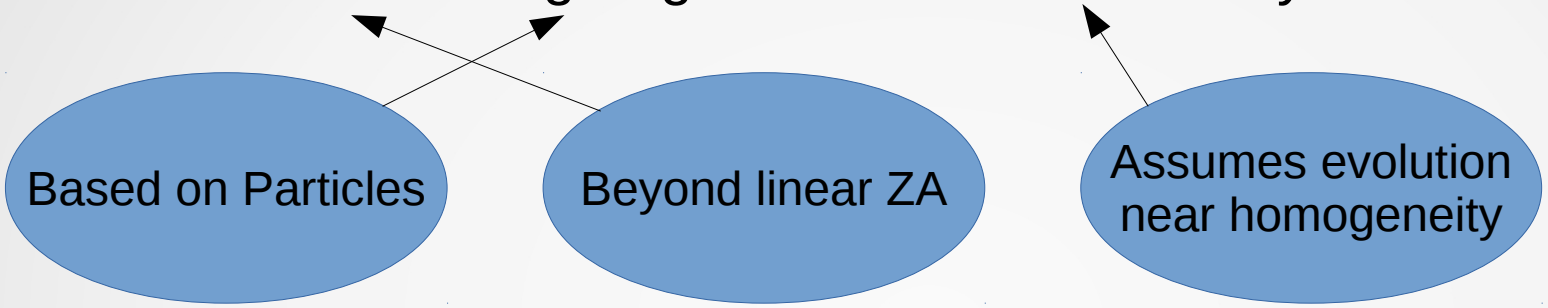
- 2LPT: **2**nd order **L**agrangian **P**erturbation **T**heory



- It reproduces well the large scales, but fails at small scales.

2LPT

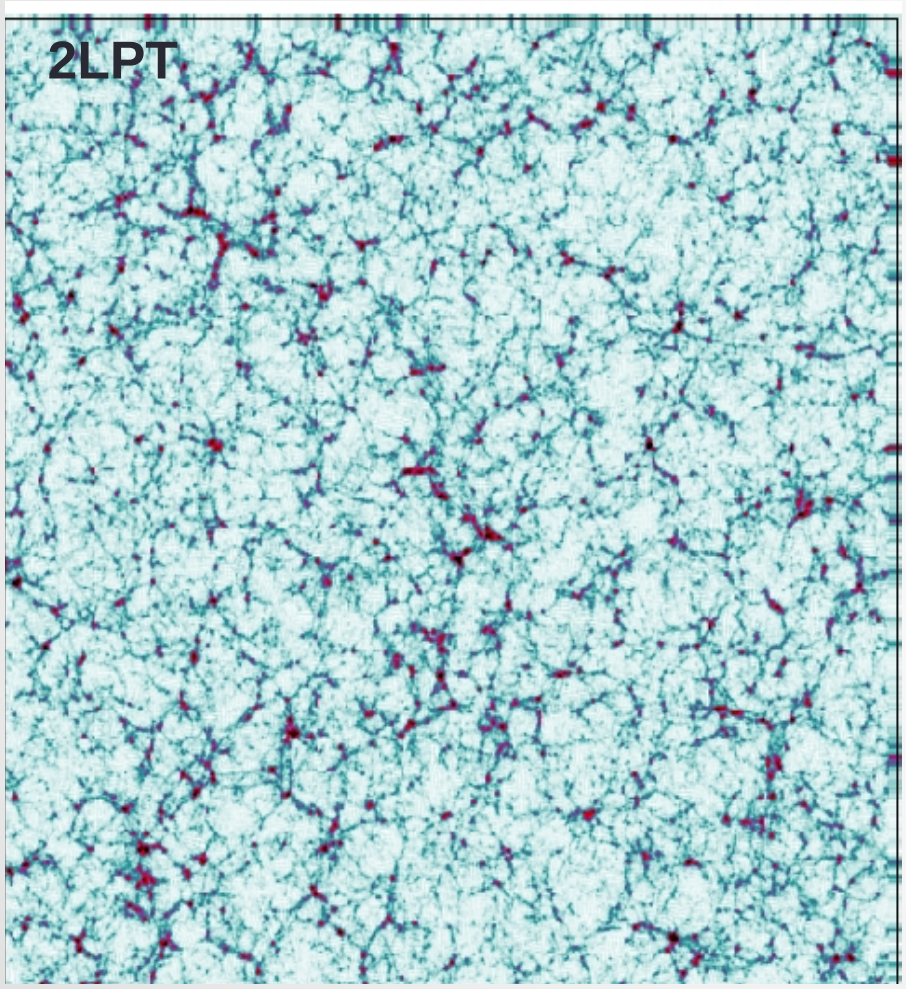
- 2LPT: **2**nd order **L**agrangian **P**erturbation **T**heory

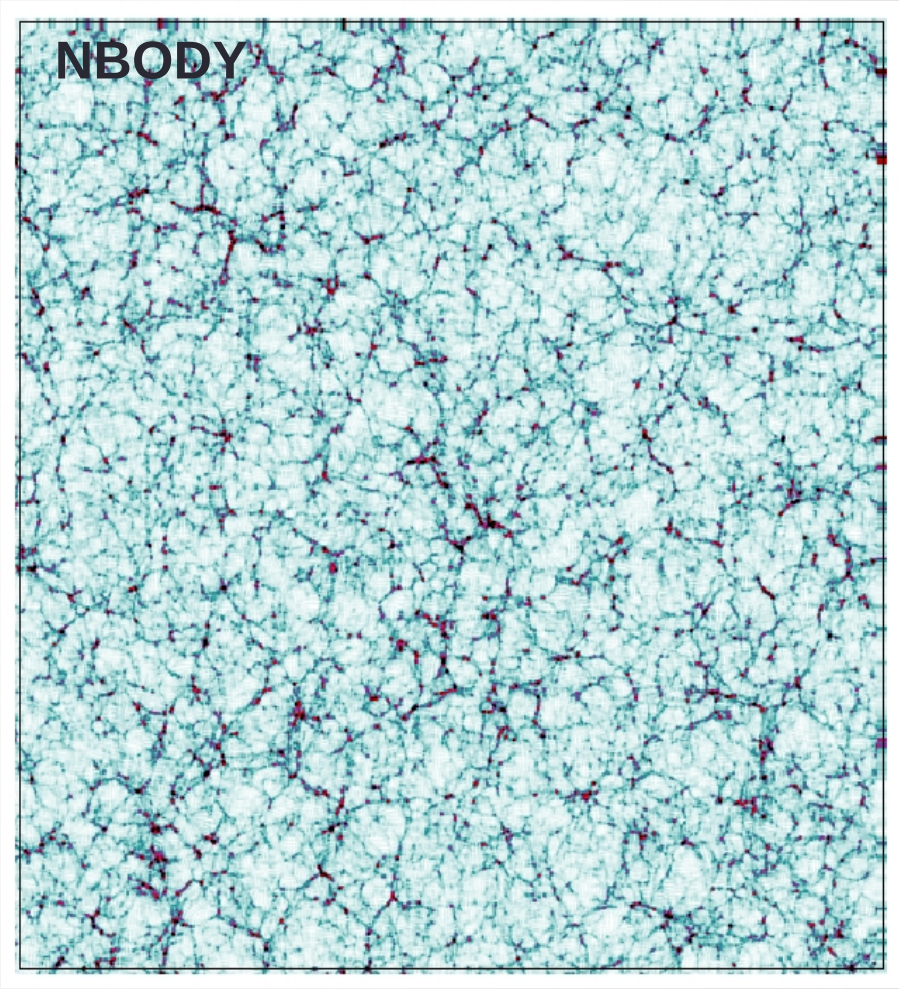


- It reproduces well the large scales, but fails at small scales.

- It can be generated from an IC generator (Scoccimarro et al.)

2LPT vs N-Body





Halo Mass Function (HMF)

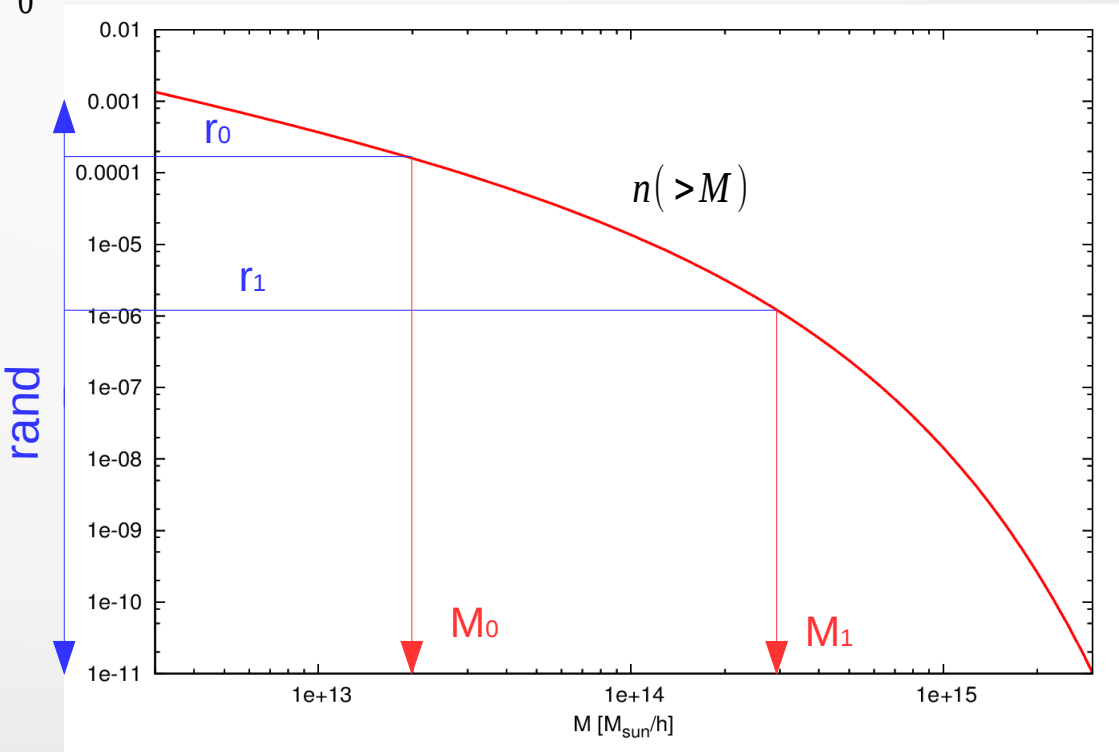
- Assume an analytic HMF fit (e.g. Tinker, Watson, etc).

  In particular, we use the cumulative HMF $n(>M)$

- Select a minimum mass M_0 and derive the number of halos

  $N = n(>M_0) \cdot Vol$

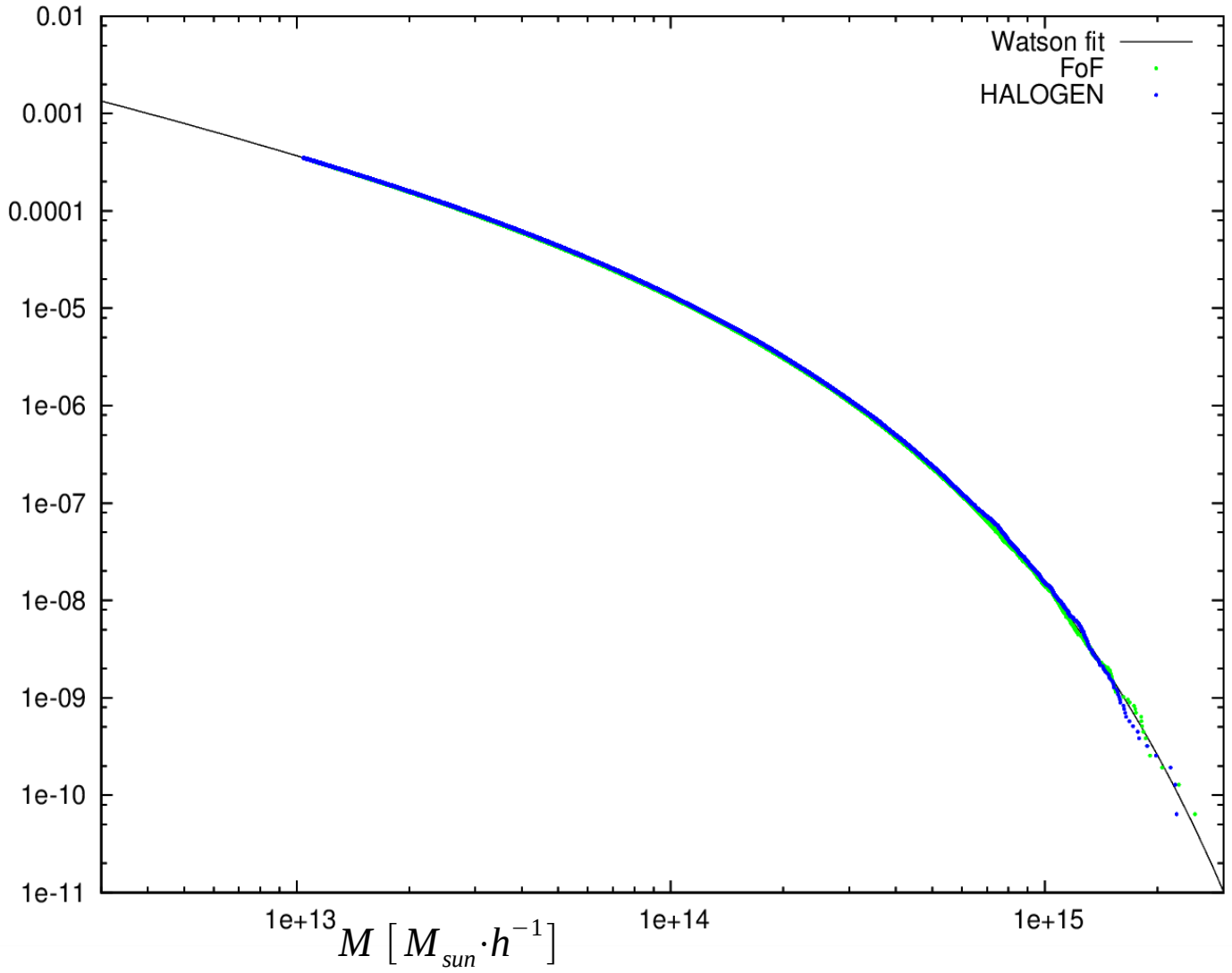
- Using Inverse-Cumulative-Distribution-Function, we generate masses for N halos.



Halo Mass Function (HMF)

$n(>M)$

$[(Mpc/h)^{-3}]$



Placing halos

- Use 2LPT particles as a scaffolding → Place halos in particles
- Select particles in a manner that reproduces the correct 2-pt correlation function of halos (implicit bias model).
- This is the difficult part...

Placing halos

- Use 2LPT particles as a scaffolding → Place halos in particles
- Select particles in a manner that reproduces the correct 2-pt correlation function of halos (implicit bias model).
- This is the difficult part...

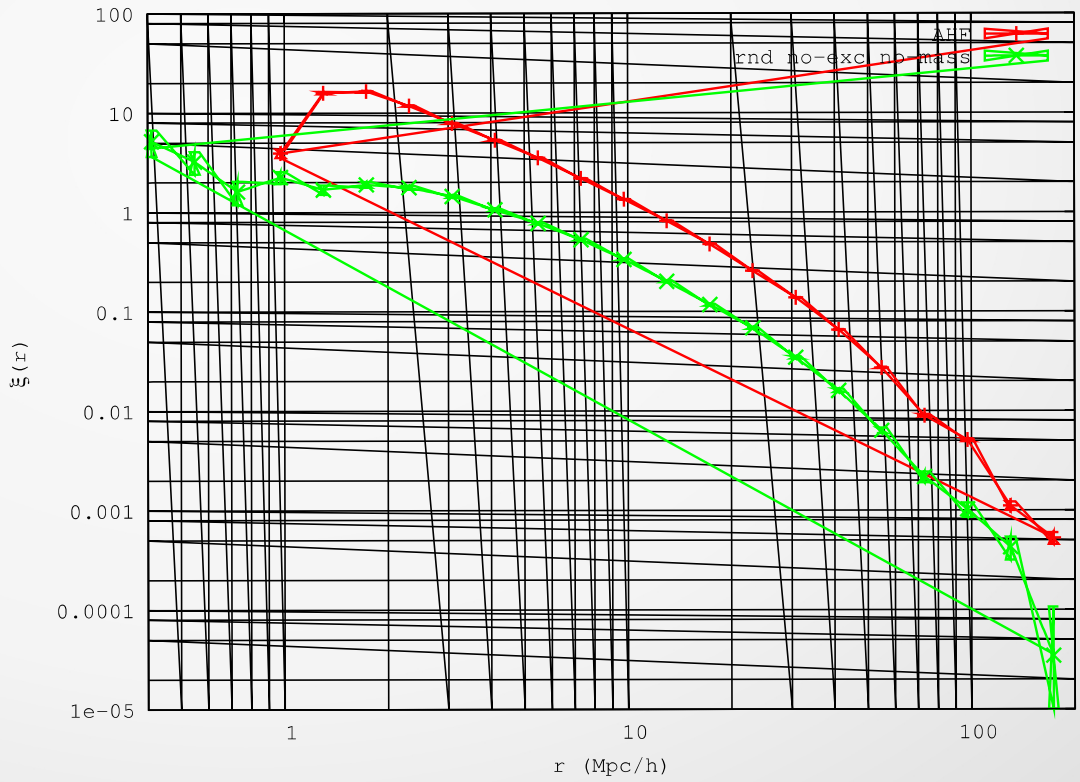
start from the simplest ideas

Placing halos

- Version 1. Select Random particles

- Too un-bias at most scales

- Positive correlation at low scale

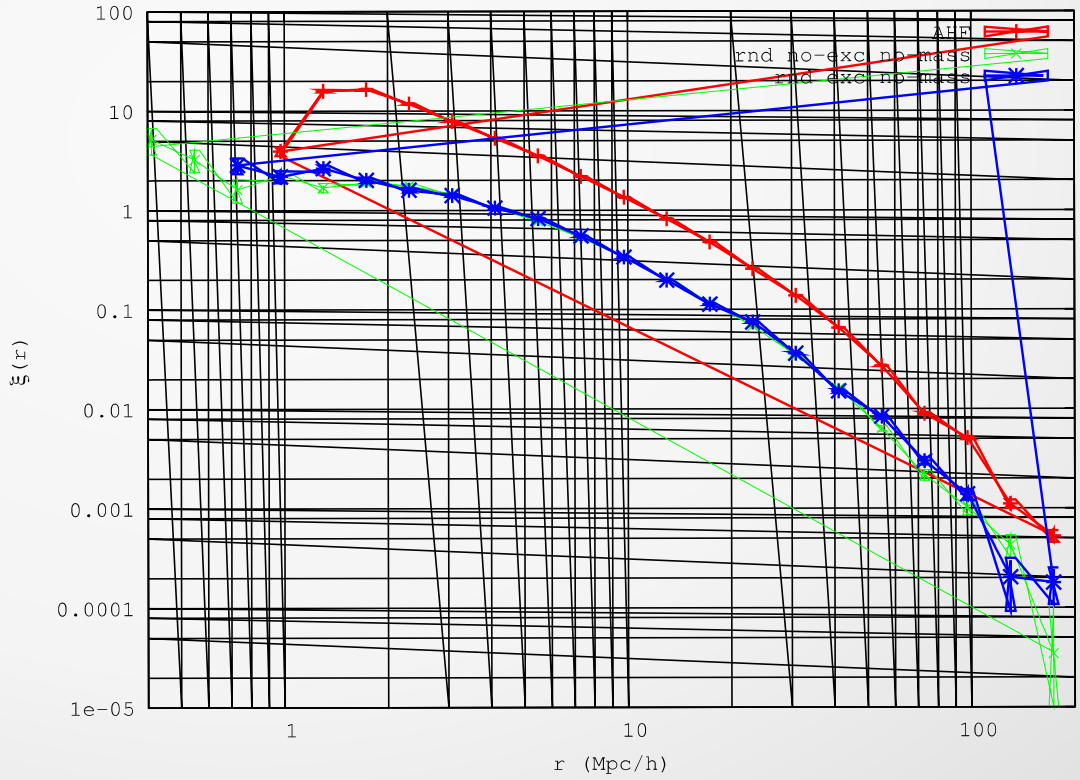


Placing halos

- Version 2. Add exclusion
  halos not allowed to overlap

+ Better, low scale cut-off

- Still low bias



Placing halos

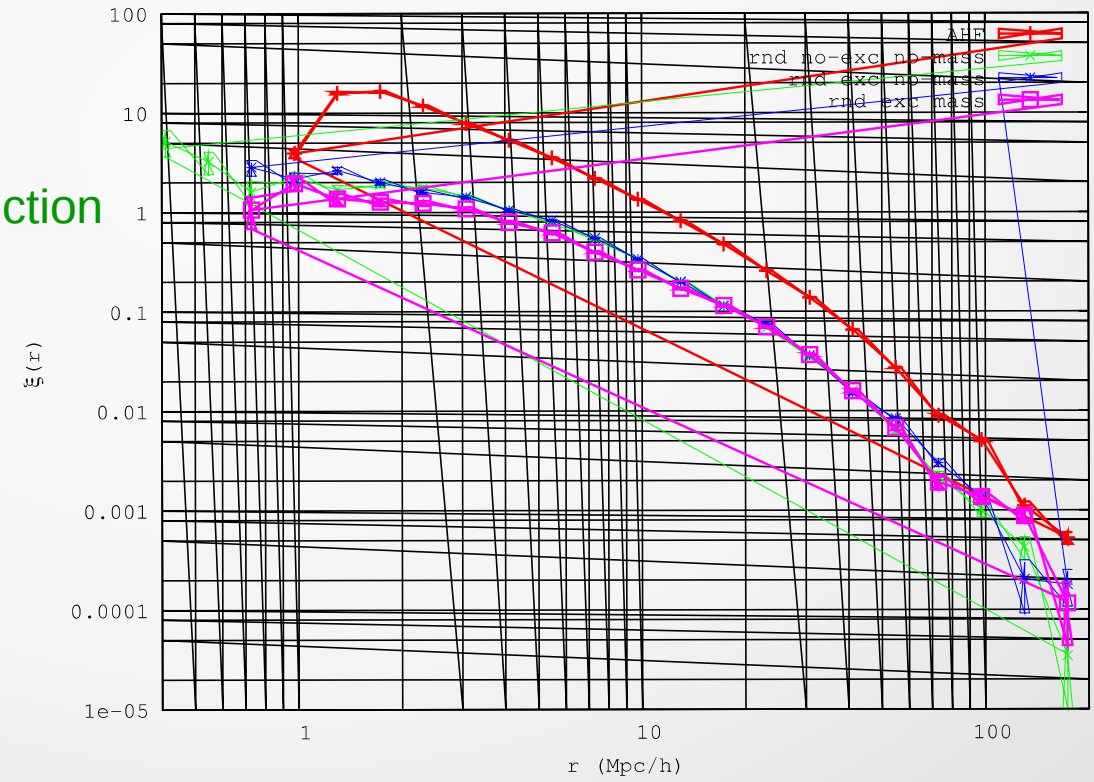
- Version 3. Conserve mass in cells
  Split the simulation box in a grid of cells
  Do not allow the mass of the halos exceed the mass of the particles in a given cell.

- Slightly better cut-off

- Better shape (tilt) of the function (difficult to see in this plot)

- Slows down the method

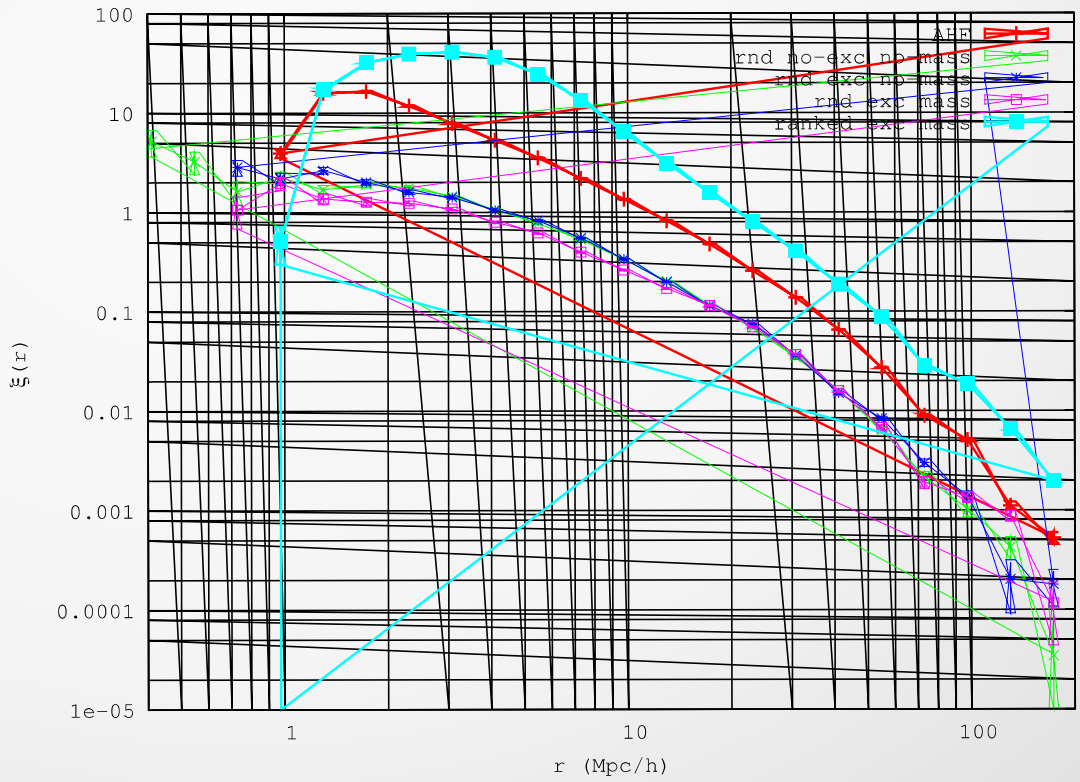
- Still low bias



Placing halos

- Version 4: Ranked-ordered placement
  place the n^{th} most massive halo in the n^{th} densest cell and pick up a random particle within cells

- Too much bias



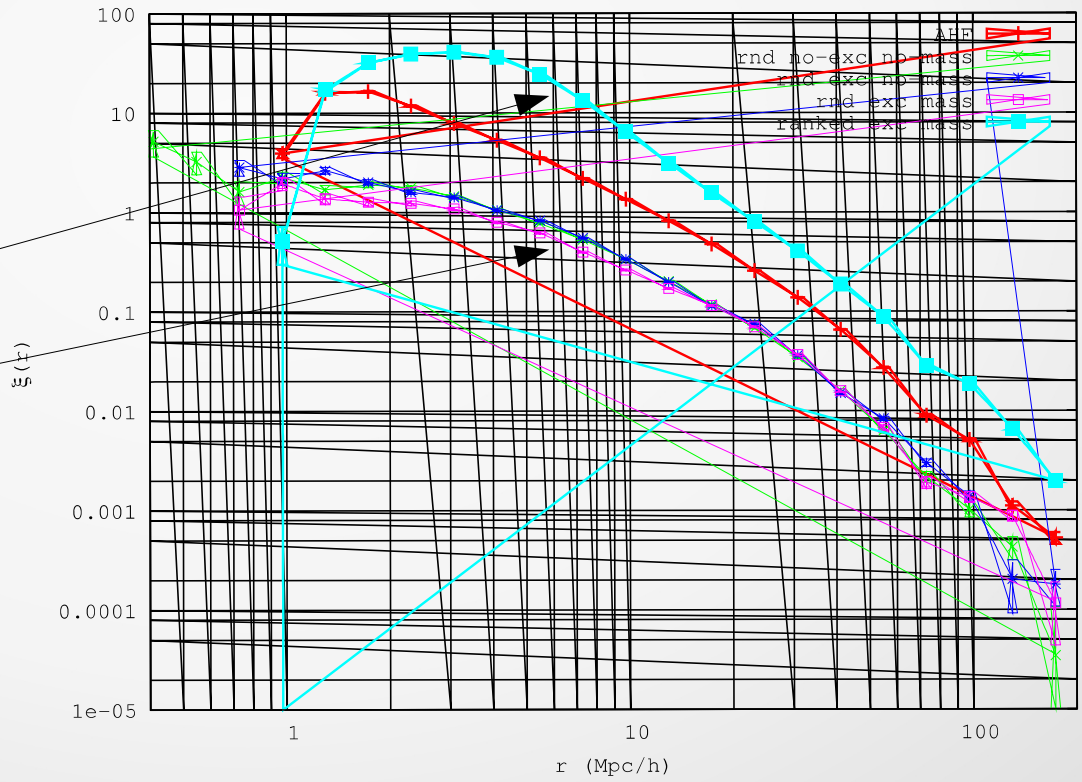
Placing halos

- Version 4: Ranked-ordered placement
  place the n^{th} most massive halo in the n^{th} densest cell and pick up a random particle within cells

- Too much bias

  - Ranked: too bias
  - Random: too stochastic

Anything in between?



Placing halos

- Version 5: Control bias-stochasticity

  - Weight the probability of picking up a cell by $P_{cell} \propto \rho_{cell}^{\alpha}$

Placing halos

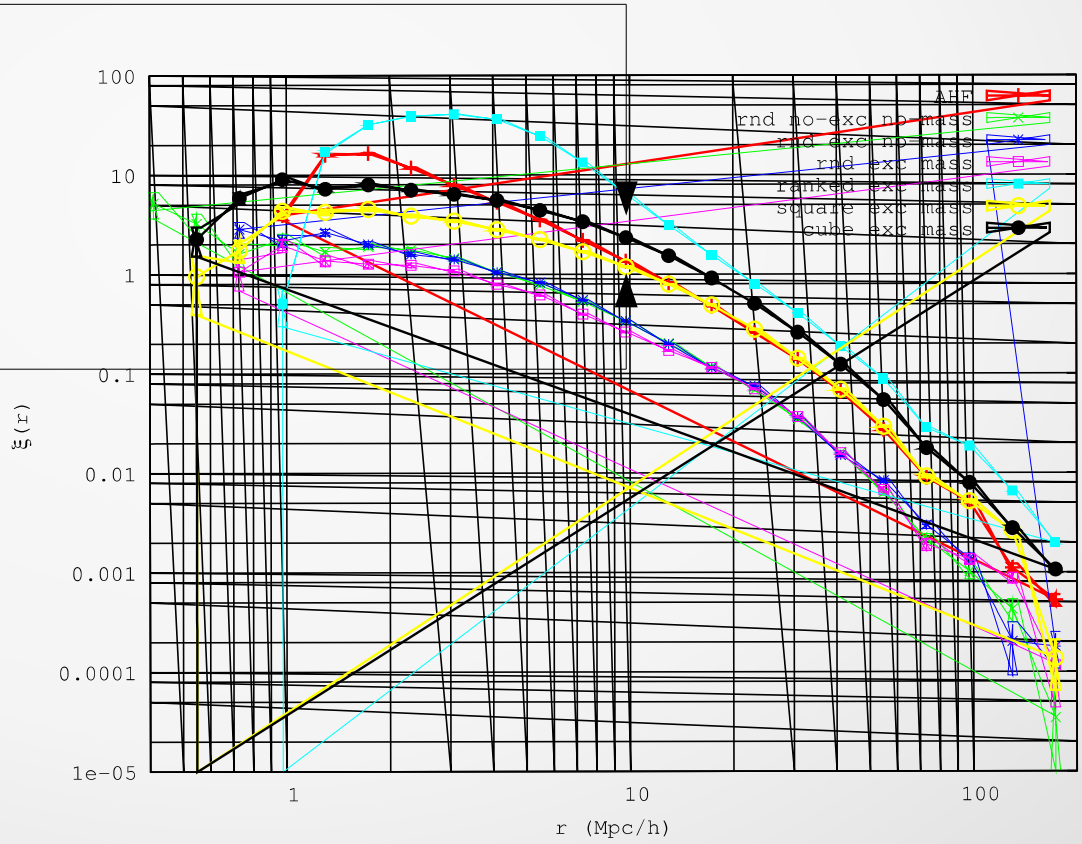
- Version 5: Control bias-stochasticity

  - Weight the probability of picking up a cell by $P_{cell} \propto \rho_{cell}^{\alpha}$

  - Example: $\alpha = 3$
    $\alpha = 2$

  - With $\alpha = 1$ we recover *rand*

  - With $\alpha = \infty$ we recover *rank*



Placing halos

- Version 5: Control bias-stochasticity

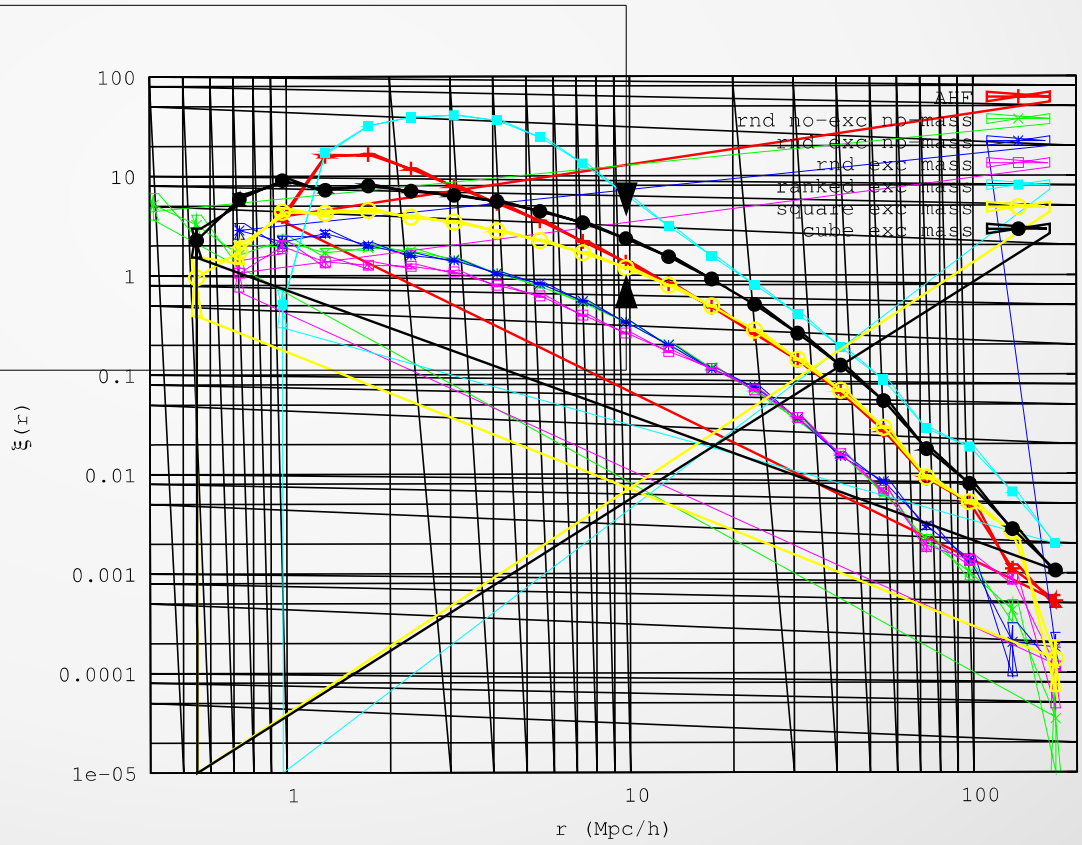
  - Weight the probability of picking up a cell by $P_{cell} \propto \rho_{cell}^{\alpha}$

  - Example: $\alpha = 3$
    $\alpha = 2$

  - With $\alpha = 1$ we recover *rand*

  - With $\alpha = \infty$ we recover *rank*

  + Can find the right bias
  - Introduce a free parameter to be fitted α



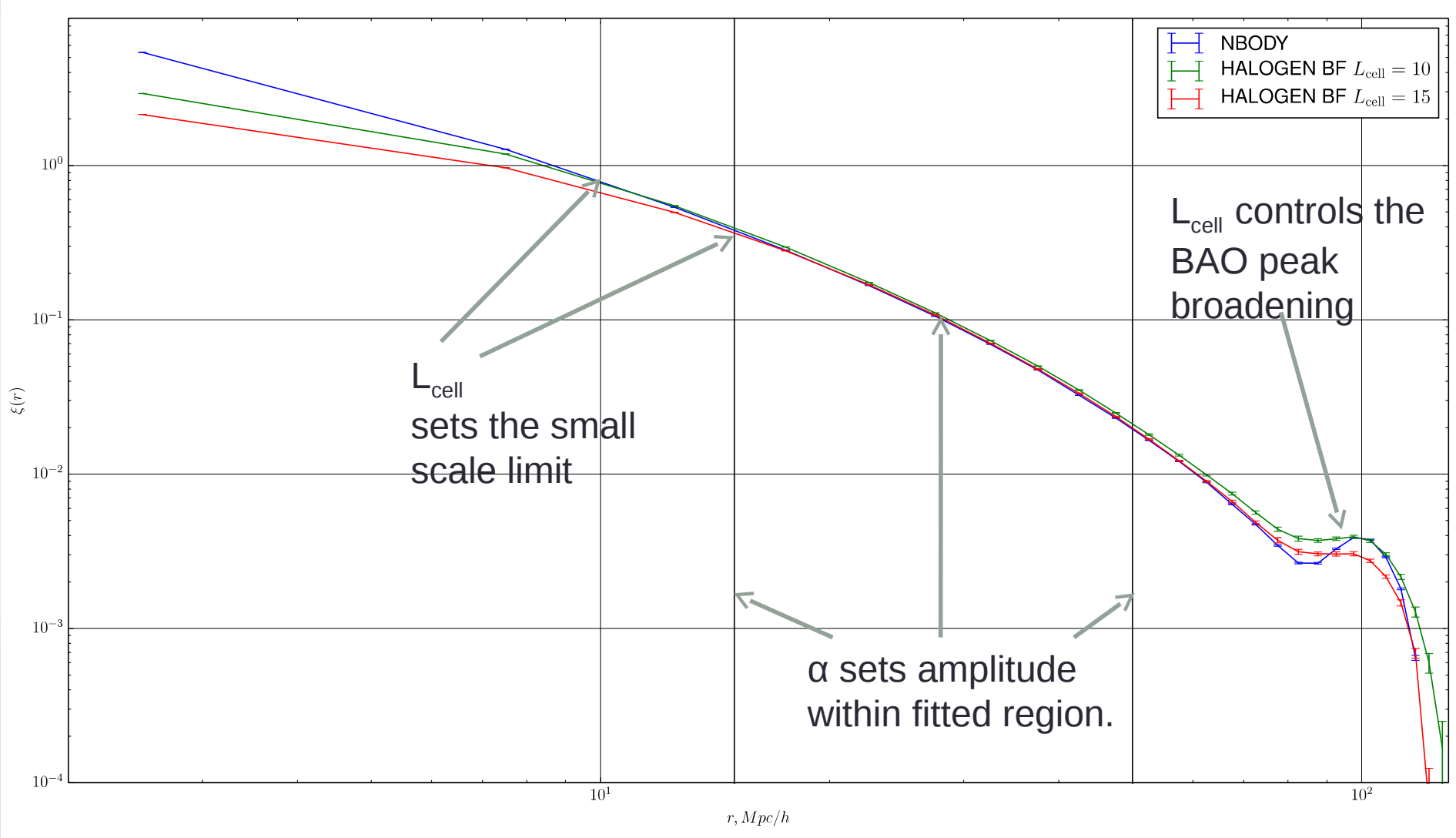
HALOGEN: recapitulation

- Run a 2LPT dark matter field
- Choose a Halo Mass Function fit and sample it with halos
- Place halos in dark matter particles:
  - Select a halo of mass M_h
  - Select a cell with probability $P_{cell} \propto \rho_{cell}^{\alpha(M_h)}$
  - Select a random particle within cell making sure the new halo does not overlap a previous one
  - Reduce the mass in the cell by M_h

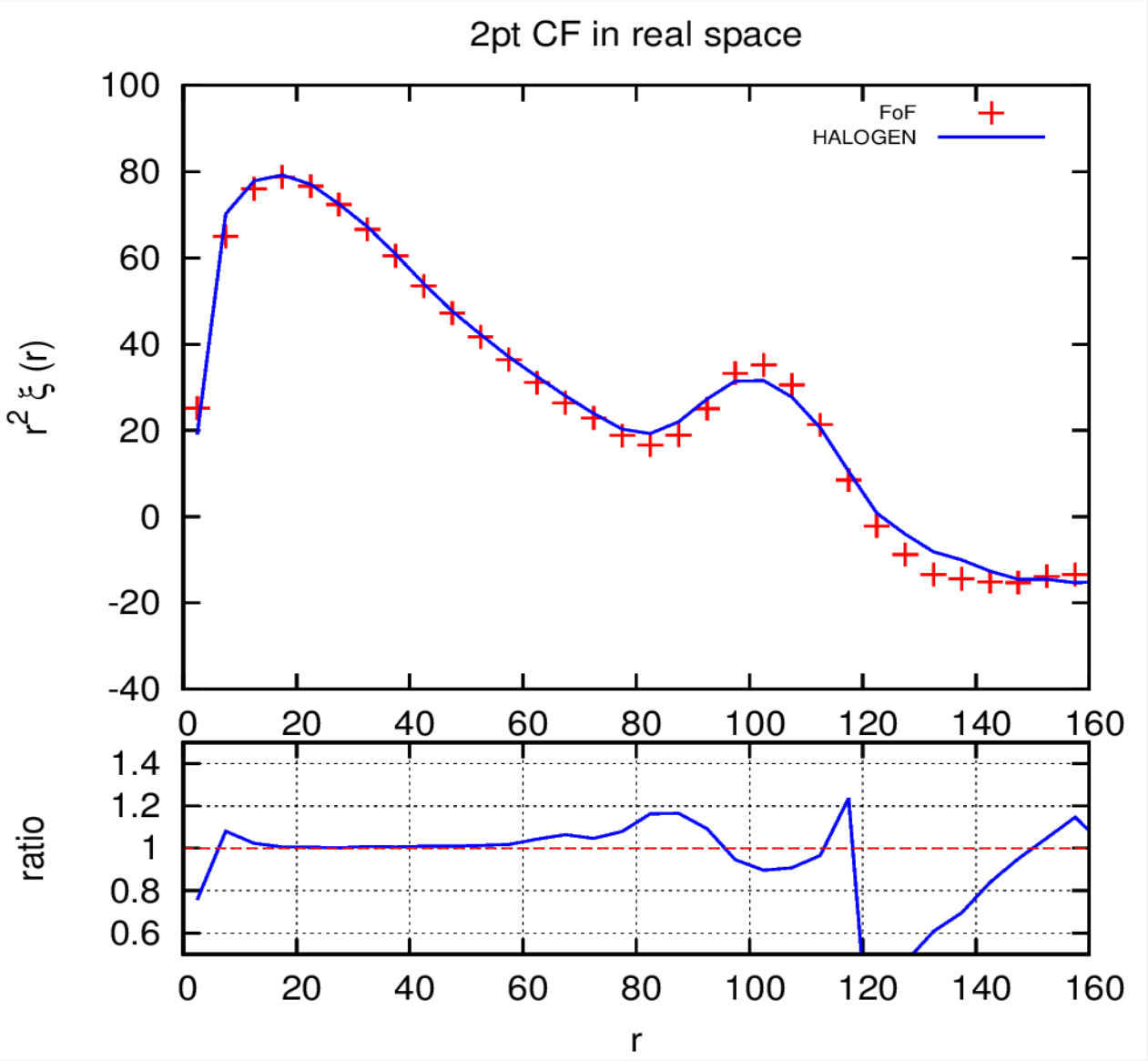
HALOGEN: recapitulation

- Run a 2LPT dark matter field
- Choose a Halo Mass Function fit and sample it with halos
- Place halos in dark matter particles:
  - Select a halo of mass M_h
  - Select a cell with probability $P_{cell} \propto \rho_{cell}^{\alpha(M_h)}$
  - Select a random particle within cell making sure the new halo does not overlap a previous one
  - Reduce the mass in the cell by M_h
- NOTE: 2 parameters introduced:
  - α : has to be fitted in mass bins
  - l_{cell} : desirable as small as possible before 2LPT breaks down

HALOGEN: parameters



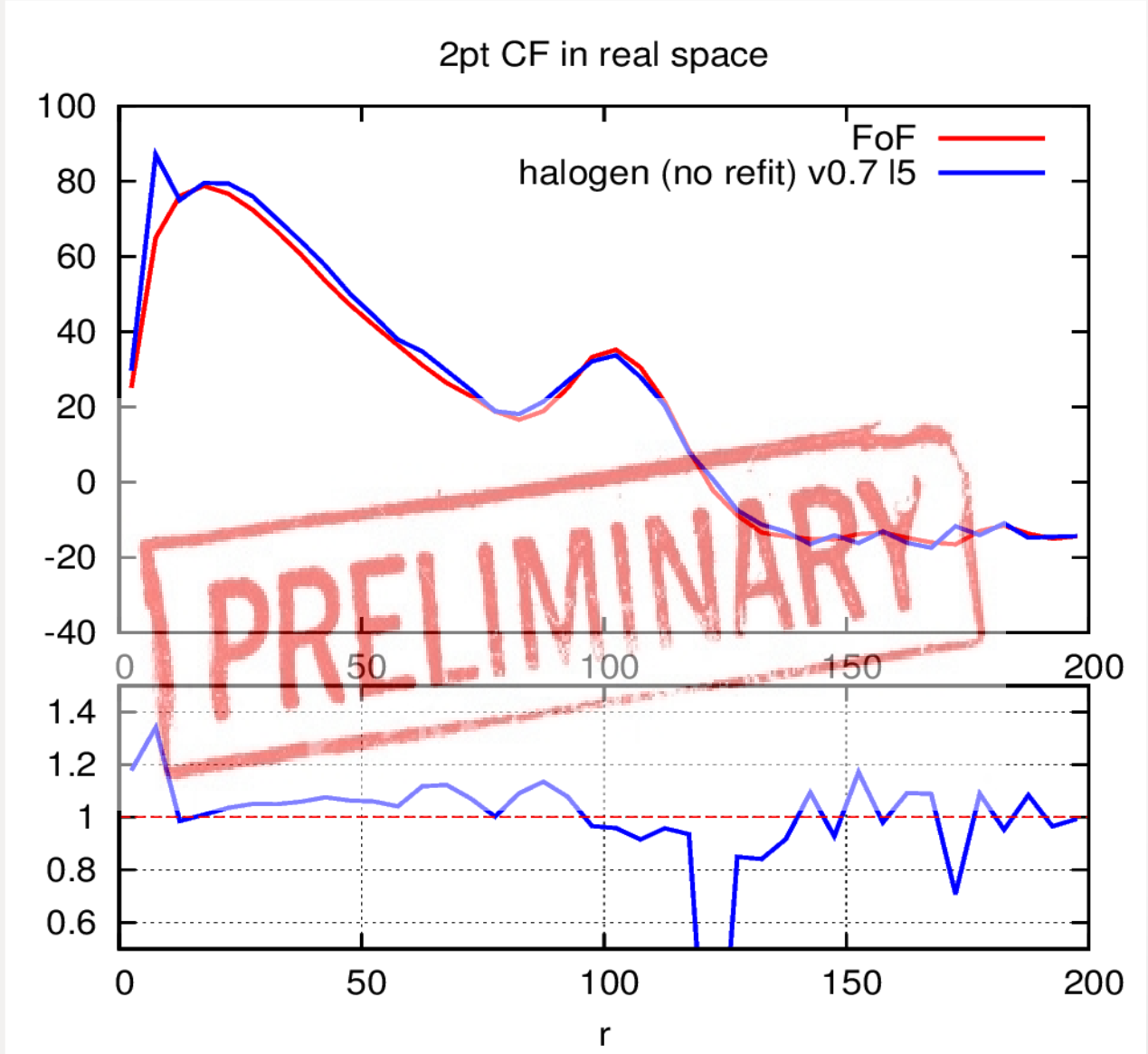
HALOGEN: results



- After using the fitting routine for $\alpha(M_h)$
- And setting $l_{cell} = 5\,Mpc/h$

(before 2LPT limitations start)

HALOGEN: Work in progress



- Working in sub-cell selection

Conclusions

- We created a tool HALOGEN capable of:
  - Creating halo catalogs with the right mass function

Conclusions

- We created a tool HALOGEN capable of:
  - Creating halo catalogs with the right mass function
  - And the 2-point correlation function correct

Conclusions

- We created a tool HALOGEN capable of:
  - Creating halo catalogs with the right mass function
  - And the 2-point correlation function correct
  - It can create a mock catalog in a 2.5 Gpc/h box with ~5 million halos in ~5min.

Conclusions

- We created a tool HALOGEN capable of:
  - Creating halo catalogs with the right mass function
  - And the 2-point correlation function correct
  - It can create a mock catalog in a 2.5 Gpc/h box with ~5 million halos in ~5min.
  - The k-space, redshift space and 3-point statistics are (preliminarily) also recovered in a slightly wider margin of error. (Chuang et al., in prep.)
  - We do not need to resolve halos → less RAM and faster

Conclusions

- We created a tool HALOGEN capable of:
  - Creating halo catalogs with the right mass function
  - And the 2-point correlation function correct
  - It can create a mock catalog in a 2.5 Gpc/h box with ~5 million halos in ~5min.
  - The k-space, redshift space and 3-point statistics are (preliminarily) also recovered in a slightly wider margin of error. (Chuang et al., in prep.)
  - We do not need to resolve halos → less RAM and faster
- We introduced a parameter $\alpha(M_h)$ that controls the bias.
  - It has to be fitted → Fitting routine
  - Once fitting, the 1000 mock catalogs can be run very quickly
  - We aim at implementing a fit from theory predictions
- Future: introduce galaxies

HALOGEN

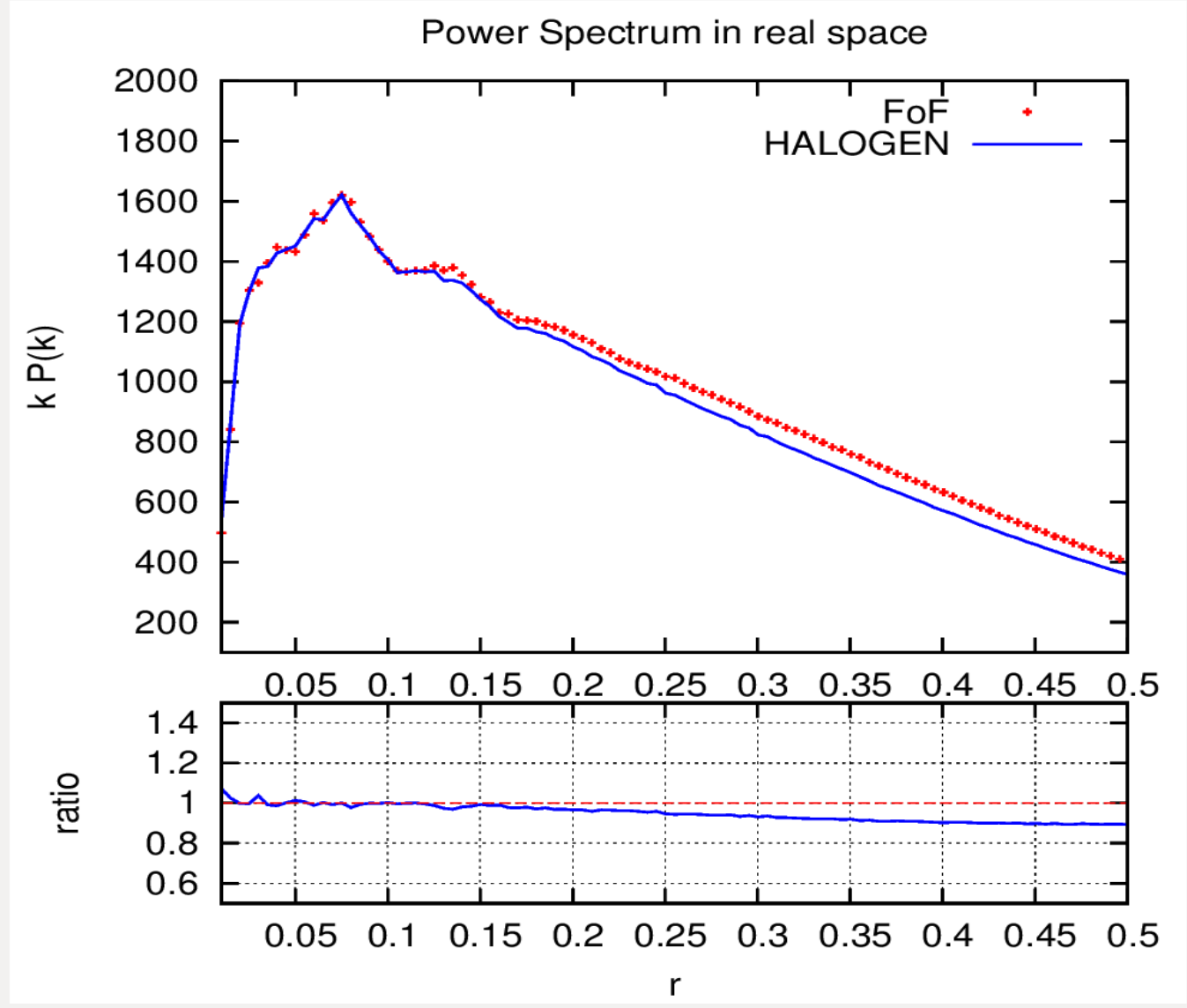
Thank you for your attention

Santiago Avila

UAM UNIVERSIDAD AUTONOMA DE MADRID

ift Instituto de Física Teórica UAM-CSIC

HALOGEN: P(k)


Power Spectrum in real space
FoF
HALOGEN
k P(k)
2000
1800
1600
1400
1200
1000
800
600
400
200
0.05
0.1
0.15
0.2
0.25
0.3
0.35
0.4
0.45
0.5
ratio
1.4
1.2
1
0.8
0.6
r