Halo-matter cross correlation in cosmological simulations

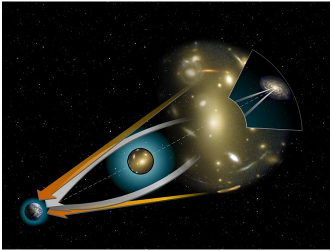
Relatore: Prof. Giuseppe Tormen
Correlatore: Prof. Ravi Sheth
Laureando: Brunetto Marco Ziosi

Dipartimento di Astronomia
Università di Padova

22 Luglio 2011

Outline

Large scale structures investigation



Compare theoretical predictions with:

- Simulations
- Observations: gravitational lensing

Halo model extension: sub-haloes investigation

Brief history

- Quantum fluctuations amplified by inflation
- Gravity let the perturbations grow and collapse
- Formation of the first structures: DM haloes
- Aggregation to form bigger haloes
- DM structures lead the formation of the baryonic structures: galaxies, . . .

Structures evolution

Linear evolution

- Friedmann's equations
- $\delta(t) \propto \frac{1}{a^2(t)\rho_{bg}(t)}$
- Continuity+Euler+Poisson equations (e.g. Cooray&Sheth)
- $\delta(k,t) = G(r)\delta(k,0) \sim \frac{1}{1+z}\delta(k,0)$

Non linear evolution

- Violent relaxation
- Secondary infall
- Excursion set
  - one-to-one relation
  - $\delta_{NL,coll} \leftrightarrow \delta_{c,lin} = 1.686$
  - Halo model from Neymann *et al.* galaxies distribution model

Simulations

GIF : see Diaferio *et al.* 1999, softening $= 30$ kpc$/h$, box side of 141 Mpc$/h$, 256^3 particles with mass $1.4 \times 10^{10}\ M_\odot$

GIF : see Gao *et al.* 2004, softening $= 6.6$ kpc$/h$, box side of 110 Mpc$/h$, 400^3 particles with mass $1.4 \times 10^{10}\ M_\odot$

Millimillennium : see Springel *et al.* 2005, softening $= 5$ kpc$/h$, box side of 62.5 Mpc$/h$, 19×10^6 particles with mass $1.7 \times 10^9\ M_\odot$

Millennium : see Springel *et al.* 2005, softening $= 5$ kpc$/h$, box side of 141 Mpc$/h$, 2160^3 particles with mass $1.4 \times 10^{10}\ M_\odot$

Millennium : see Boyland-Kolchin *et al.* 2009, softening $= 1$ kpc$/h$, box side of 100 Mpc$/h$, 2160^3 particles with mass $6.9 \times 10^6\ M_\odot$

What is it?

We need a way to characterize big spatial datasets:

Two point correlation function ξ

- It quantify the excess or defect of probability to find a pair of objects separated by a distance r
- $\mathrm{d}^2 P = n^2 \mathrm{d} V_1 \mathrm{d} V_2 [1 + \xi(r)]$

Estimators

- $\xi_n = \frac{DD}{RR} - 1$, $\xi_{DP} = \frac{DD}{DR} - 1$, $\xi_{Hew} = \frac{DD-DR}{RR}$, $\xi_{Ham} = \frac{DD\ RR}{DR^2} - 1$, $\xi_{LS} = \frac{DD-2DR+RR}{rr}, \ldots$
- The Landy&Szalay is recommended
- Good behavior: $\Delta\hat{\xi}_{LS} = \frac{1+\xi}{\sqrt{DD}}$

TPCF and spectra

The wiener-Khincine theorem

- Density fluctuation as superposition of plane waves

$$\delta(\mathbf{x}) = \frac{1}{(2\pi^3)} \int \hat{\delta}(\mathbf{k}) \exp(i\mathbf{k} \cdot \mathbf{x}) \mathrm{d}^3 k$$

- The power spectrum is the Fourier transform of the correlation function:

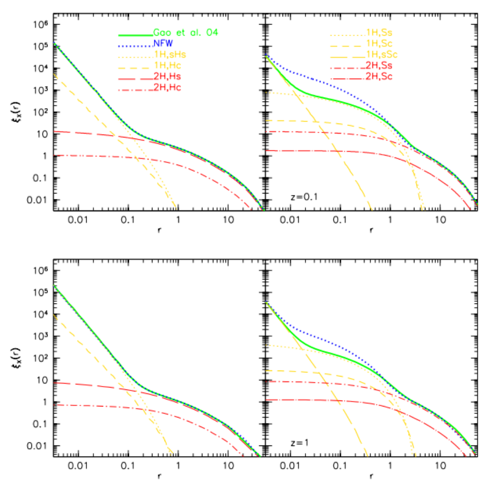
$$\xi(r) = \frac{1}{(2\pi)^3} \int P(k) \exp(i\mathbf{k} \cdot \mathbf{r}) \mathrm{d}^3 k$$

- Substitute the halo and matter density perturbation fields and it is still valid

$$\xi_{hm}(r) = \langle \delta_h(\mathbf{x}) \delta_m(\mathbf{x} + \mathbf{r}) \rangle$$

TPCF and haloes

Halo-matter, Subhalo-matter



Haloes contributes:

- 1H term:
  - Self halo-smooth
  - Self halo-clump
- 2H term:
  - Halo-smooth other
  - Halo-clump other

Sub-haloes contributes:

- 1H term:
  - Self subhalo-clump
  - Subhalo-smooth in host
  - Subhalo-clump in host
- 2H term:
  - Subhalo-smooth other hosts
  - Subhalo-clump other hosts

Python

Advantages

- Flexibility
- Fast developing
- Math libraries
- Portability
- Parallelization

Disadvantages

- Not compiled
- Not so fast

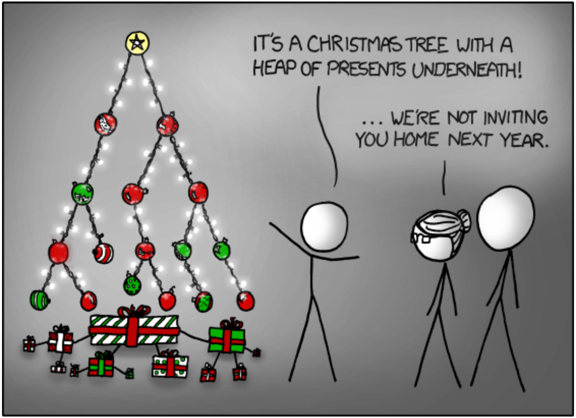
But:

- Optimization
- Use it as a glue
- With fast C/C++/Fortran core

Tree

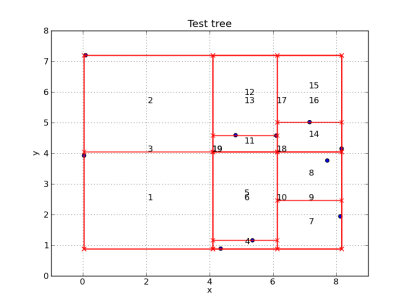
A *kd*-tree is

a hierarchical data structure in which each node represents a subset of the dataset.



Tree 2

KD-Tree



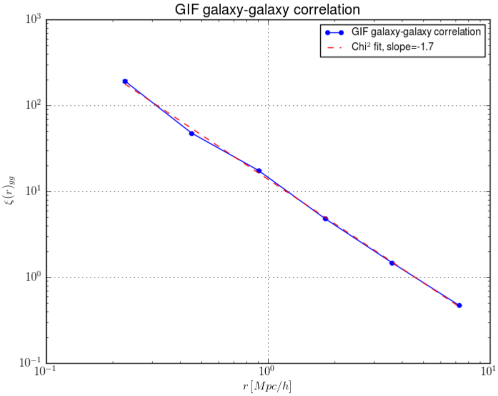
Properties

- Binary: each node has two subnodes
- Best *nearest neighbours problem* data structure
- Balanced when building it

Optimization

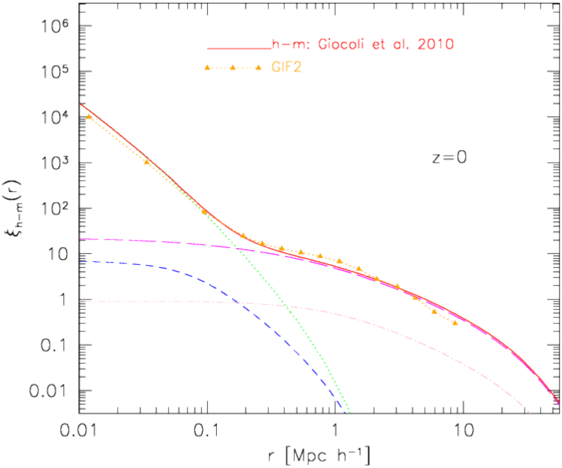
- Dual tree search
- Exclusion pruning (r_{min}, r_{max})
- Inclusion pruning
- Non redundant search
- Multiple radii
- Cached statistic
- Leaf opening strategies
- Leafsize optimization
- Possible persistence with smart node caching (PyTables)
- Time $\propto N \log(N)$ vs N^2

GIF galaxies-galaxies test



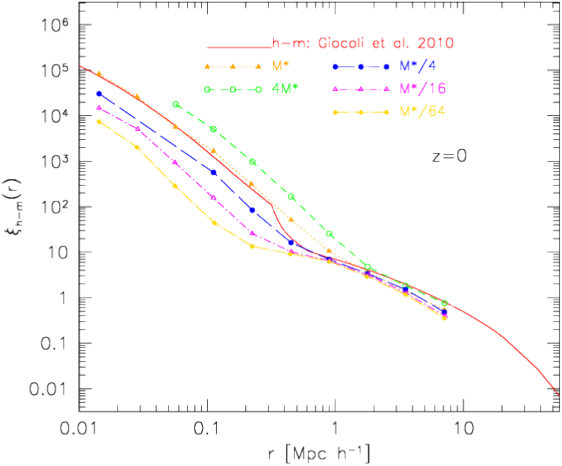
- First test of the code
- χ^2 slope of -1.7 vs -1.8 in literature
- Compare with Diaferio *et al.* 1999

GIF2 Halo-matter



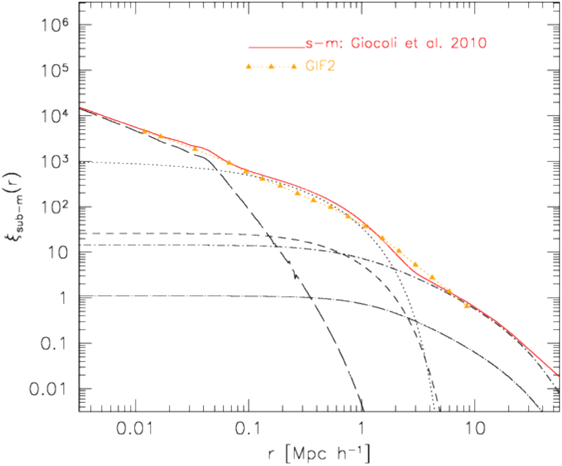
- Second, winning, test
- Small scales: NFW halo profile
- Intermediate scales: transition between 1H and 2H terms
- Large scales: bias and δ_{lin} for DM
- Compare with Hayashi&White 2008
- Compare with Giocoli *et al.* 2010

Halo-matter mass bins



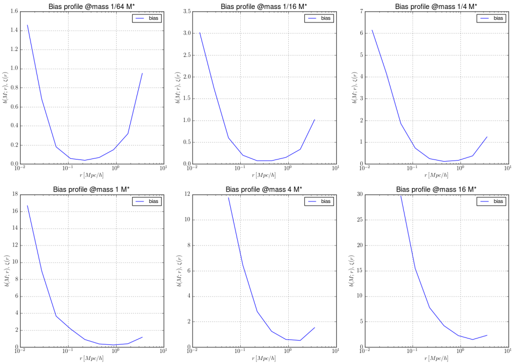
- Mass bins: $\frac{M^*}{64}$, $\frac{M^*}{16}$, $\frac{M^*}{4}$, 1, $4M^*$, $16M^*$
- $M^* = 8.9 \times 10^{12} M_\odot$
- Different signals from haloes of different masses
- Good agreement with theoretical predictions
- Poor sampling for higher masses
- Mass cut-off in the Fourier space for the predictions
- Compare with Hayashi&White 2008
- Compare with Giocoli *et al.* 2010 theoretical model

Subhalo-matter



- 1H term: subhalo with self-smooth, host-smooth and host-clump components
- 2H term: subhalo with other-smooth and other-clump components
- Compare with Giocoli *et al.* 2010

Bias



- Halo-matter divided by matter-matter
- Proportional to the bias
- Compare with Mo&White 1996
- High resolution shows halo profiles
- Smaller scales: only the first part of the profile

Conclusions

Conclusions

- Subhaloes-matter cross-correlation only for GIF2
- The Millenium II is now on with halo-matter
- Problems:
  - Code choices
  - Data retrieve
  - Data reading (Bad formats, few informations)
  - Hardware problems (failures and power off)
  - HW/SW compatibility

Further works

- Go on with the work
- Further optimization (Chyton, C/C++/Fortran), not promising
- More cached statistic
- GPUs
- Modified persistence