Simulations of Cosmic Surveys

Prof Shaun Cole

Institute for Computational Cosmology
Durham University
U.K.

April 2015

Introduction

What is a mock catalogue?

A fake/simulated galaxy catalogue which resembles a genuine galaxy redshift survey catalogue, but is built from a cosmological simulation.

Motivation for mocks:

- Develop and test analysis code
  (test photo-z, group finding, clustering; can be done in advance of having the observational data; you know what results you should get)
- Assess systematic errors
  (non-trivial due to non-linearity and galaxy bias)
- Estimate covariance matrices
 (Needed to get accurate parameter constraints)

Mock Ingredients

- A cosmological N-body simulation
  or shortcut (PTHaloes; PINNOCHIO; EZmocks)
  (gravity only as large scale hydro simulations of sufficient resolution are not yet feasible)
- A method of placing galaxies into the simulation
- Reproduction of the survey selection and completeness

Examples

On the right is a famous plot from the Millennium Nature paper comparing genuine surveys with simulated equivalents. Here the mocks were used to make a direct visual comparison of theory and observation, supplementing the quantitative statistical comparison.

Our focus is more on developing mocks for direct use in quantitative analysis. Some recent examples include:

- GAMA and 2dFGRS group finding
- BOSS reconstruction
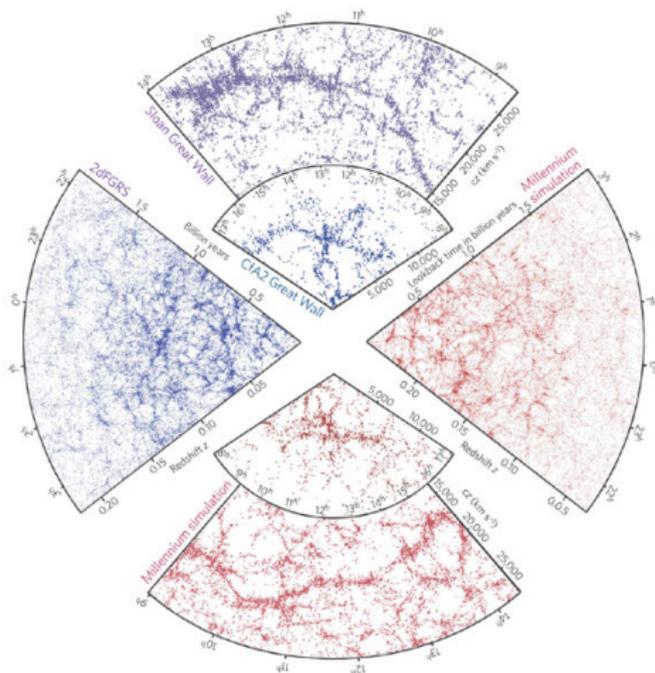
Springel et al 2006

Example: 2PIGG (2dF Group Catalogue)

The 2dFRGS team used mock catalogues to tune the parameters of a group and cluster finding algorithm. In the DM simulation from which the mock was made one can accurately identify the dark matters halos (groups and clusters) using standard methods. The mock enables:

1. The 2PIGG algorithm to be tuned so that the resulting catalogue has an optimal balance of completeness and purity (finds most of the groups and assigns the right galaxies to them)

2. Determine the likely bias in any inferred property (e.g. in the inferred cluster mass-to-light ratios)
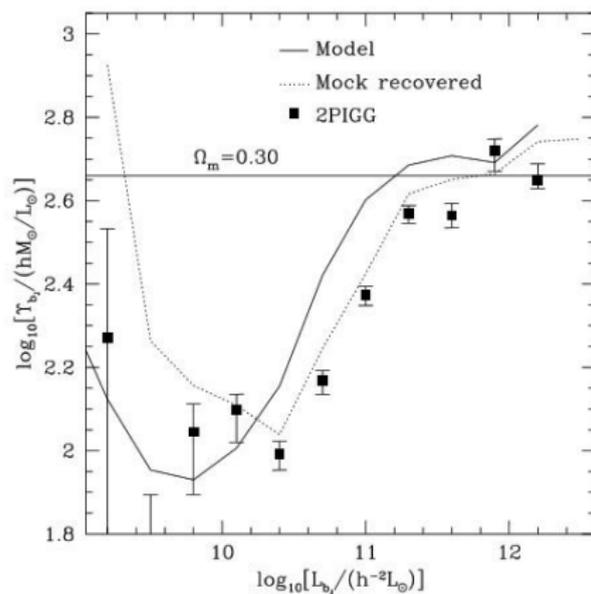
Eke et al 2004

Example: BOSS reconstruction

billion light years

Anderson et al 2012

Padmanabhan et al 2012
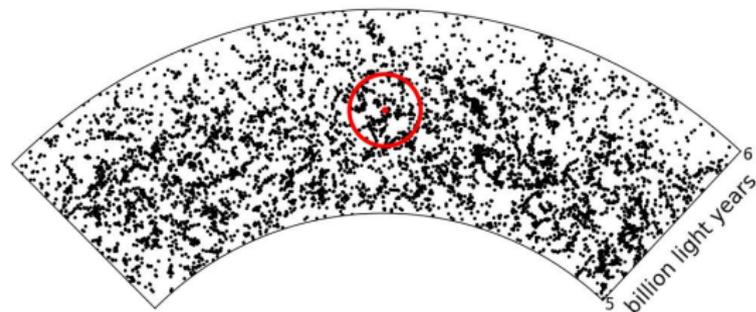
The small excess of galaxies expected at the BAO scale is distorted by the growth of large scale structure.

Reconstruction attempts to undo this distortion so that the BAO length scale can be measured more accurately. The full analysis and reconstruction process is complex and non-linear.

Mock catalogues are need to test the
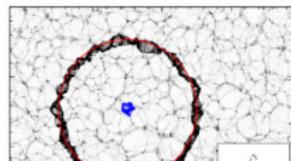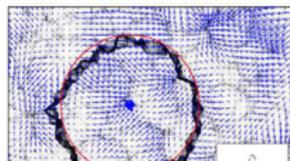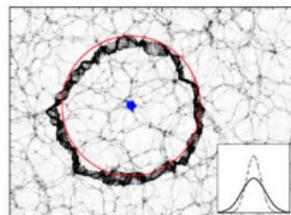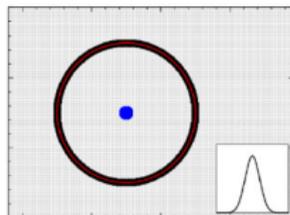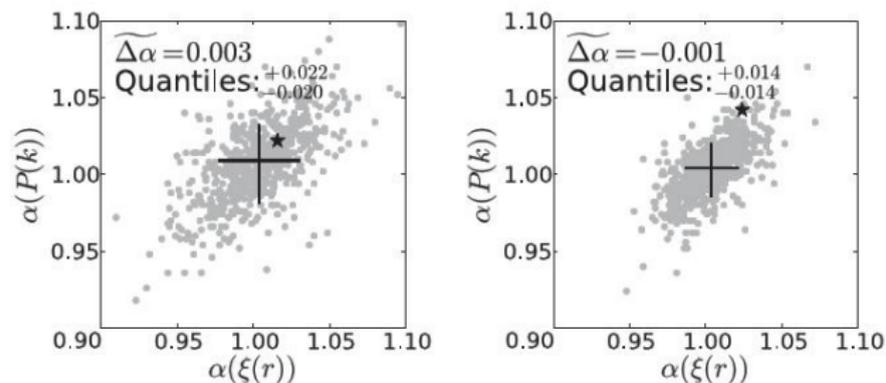
Example: BOSS reconstruction

demonstrate that the true length scale ($\alpha = 1$) is recovered with minimal bias and variance.

The mocks can also be used to determine whether analysis of the galaxy clustering in Fourier space (the power spectrum $P(k)$) or configuration space (the correlation function $\xi(r)$) yields the tightest constraints.

Anderson et al 2012

Outline

- N-body Simulation Techniques
  (force calculation; timesteps; softening; initial conditions)
- N-body Results
  (the cosmic web, halos and merger trees)
- Populating with Galaxies
  (HOD; SHAM; semi-analytics)
- Survey Construction
  (lightcones; selection functions; targeting and redshift completeness)

The N-body Problem

Evolve the orbits of system of N gravitating masses.

Only the 2-body problem (Keplerian orbits) can be solved analytically.
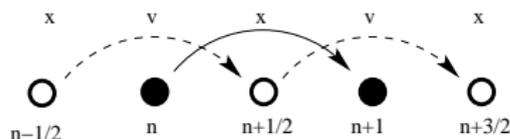
The numerical problem breaks down into

1. Calculate the resultant force on each particle due to the sum of the forces exerted on it by every other particle.
2. Integrate the orbit of each particle through a small timestep then repeat step 1.

Often a leapfrog integration scheme is employed:

$$x_{n+1/2} = x_{n-1/2} + v_n \, \Delta t$$

$$v_{n+1} = v_n + \dot{v}_{n+1/2} \, \Delta t$$

Methods of force calculation

- Direct summation (PP) : Flexible, but requires N^2 operations
- Tree : The space containing the particles is recursively divided into level after level of smaller and smaller regions, until there is no more than one particle in each region (there are different methods eg Barnes-Hut or KD-tree).

Adaptive quadtree where no square contains more than 1 particle
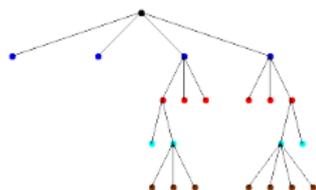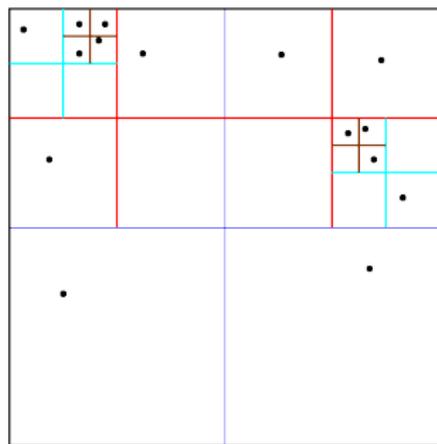
Methods of force calculation

This defines a tree structure in which the leaves contain the information on the individual particles at the lowest level and the twigs and larger branches groupings of particles from the higher levels.

The force computing algorithm is faster than PP as one exploits that the force on a given particle by a collection of distant particles can be approximated by replacing that distant collection of particles by a single more massive particle (sometimes higher order multipole corrections are taken into account).

The result is a very flexible algorithm that scales as $N \log N$.

The force accuracy is determined by an "opening angle criterion". I.e nodes subtending more than this critical angle are opened up.

Methods of force calculation

- Particle Mesh (PM): A gridded density field, $\rho(\underline{x})$, is constructed by assigning mass particles to a grid (different methods eg NGP, CIC and TSC).
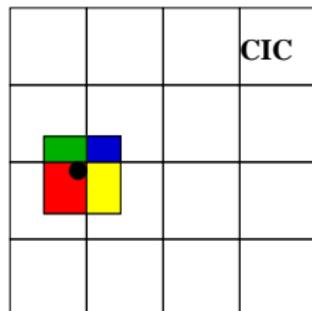
To compute the forces, $\underline{F} = -\nabla\Phi$, the potential Φ is computed on a grid by using a Fast Fourier Transform to solve Poisson's equation, $\nabla^2\Phi = 4\pi G\rho$. In Fourier Space this is $-k^2\Phi_{\underline{k}} = 4\pi G\rho_{\underline{k}}$. Thus to solve FT $\rho(\underline{x}) \to \rho_{\underline{k}}$, compute $\Phi_{\underline{k}} = -4\pi G\rho_{\underline{k}}/k^2$ and FT back to get Φ. This is all fast as the FFT takes only $N_{\mathrm{G}} \log N_{\mathrm{G}}$ operations, where N_{G} is the number of grid points.

The down side is the force resolution is determined by the cell size of the grid.

It also results in periodic boundary conditions, but this is often what one wants for cosmological simulations.

N-body Codes

- P^3M (Efstathiou) – Combined PP at short range with PM at large scales.
- GADGET (Springel) – Combines PP at short range with a tree at intermediate scales and a mesh for the largest scales. GADGET is also highly parallel.
- PKDGRAV (Stadel, Quinn) (k-D tree)
- ART (Kravstov , Klypin, Khokhlov) Adaptive Refinement Tree but actually uses meshes.

Timescales

Before deriving the equations of stellar dynamics it is useful to define two important timescales for stellar systems.

The dynamical time of a system is often defined as

$$t_{\mathrm{dyn}} = 1/\sqrt{G\rho}.$$

It is essentially the same as the crossing time

$$t_{\mathrm{cross}} = R/v,$$

as for a spherical system in virial equilibrium

$$v^2 = \frac{GM}{R}$$

$$R/v = \sqrt{\frac{R^3}{GM}} \approx 1/\sqrt{G\rho}.$$
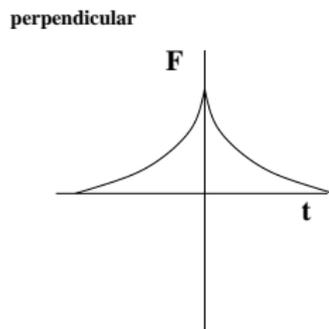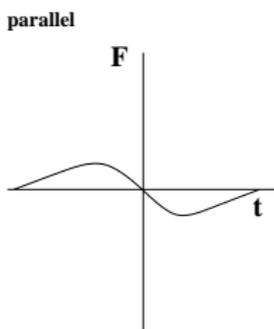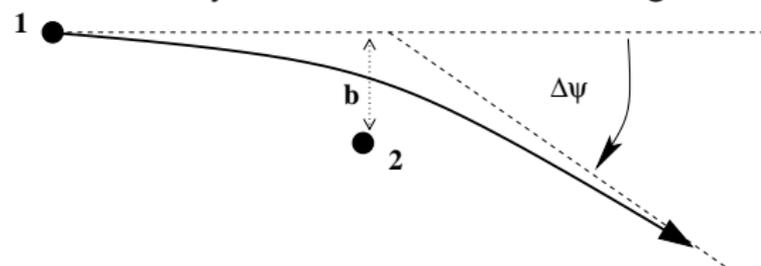
The Two Body Relaxation Time

This is the timescale over which a given star has its orbit completely changed due to two body encounters with other stars in the system.

Each time two stars pass close by each other they will bend each others orbits. The two body relaxation time can loosely be defined as the time it takes for multiple encounters to bend the original orbit through 1 radian.

The Two Body Relaxation Time

i) Consider an encounter between a pair of stars, where b is the impact parameter, v the velocity and $\Delta\Psi$ the deflection angle.

The lower plots show the time dependence of the force exerted on star 1 by star 2.

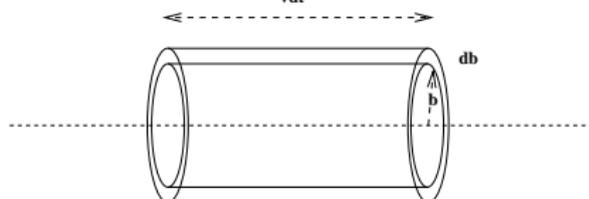By symmetry the net impulse in the parallel direction is zero.

In the perpendicular direction $F_{\max} \approx \dfrac{Gm^2}{b^2}$ and $\Delta t = \dfrac{b}{v}$ Hence

$$\Delta p_\perp = \frac{Gm^2}{b^2}\frac{b}{v}$$

The Two Body Relaxation Time

Hence the deflection angle is $\Delta \Psi = \dfrac{\Delta v}{v} = \dfrac{Gm}{bv^2}$

Now the number of encounters with impact parameter b per unit time is given by the usual kinetic theory calculation

collisions/unit time $= 2\pi \, b \, db \, v \langle n \rangle$

Hence the mean square cumulative deflection angle after time t is

$$(\Delta \Psi)^2 = \frac{G^2 m^2}{v^3} \langle n \rangle \left[\ln b \right]_{b_{\min}}^{b_{\max}} \; t$$

If we define t_{relax} as when $\Delta \Psi = 1$, i.e. when the initial stellar orbits are completely scrambled

$$\Rightarrow t_{\text{relax}} = \frac{v^3}{G^2 m^2 \langle n \rangle \ln(b_{\max}/b_{\min})}$$

There is no absolutely rigorous way of defining $b_{\min}$ and $b_{\max}$.

The Two Body Relaxation Time

Convention is to assume $b_{\max} = R$, the size of the system and
$b_{\min} = \lambda = \langle n \rangle^{-1/3}$, the typical stellar separation .
Since $t_{\mathrm{relax}} \propto 1/\ln(R/\lambda)$ we can afford to be sloppy.
Let us compare this 2-body relaxation time with the dynamical or crossing time of a stellar system.

$$t_{\mathrm{cross}} = \frac{R}{v}$$

$$\Rightarrow \frac{t_{\mathrm{relax}}}{t_{\mathrm{cross}}} = \frac{v^4}{G^2 R m^2 \langle n \rangle \ln(R/\lambda)}$$

Now we can re-express this using

1. Virial Theorem: $2T + W = 0$ i.e. $Nmv^2 - Gm^2 N^2/R \approx 0$, where N is the total number of stars. $\rightarrow v^2 \approx GmN/R$.

2. $R/\lambda = (R^3/\lambda^3)^{1/3} = (R^3 \langle n \rangle)^{1/3} = N^{1/3}$

$\Rightarrow \dfrac{t_{\mathrm{relax}}}{t_{\mathrm{cross}}} = \dfrac{N}{\ln N} \left(\dfrac{1}{8} \right)$ with 1/8th coming from calibrating using numerical simulations.

The more stars in the system the less important the 2-body relaxation.

The Two Body Relaxation Time

System	N	R	t_{dyn}	t_{relax}
Globular Cluster	10^5	10 pc	10^5 years	10^7 years
Elliptical Galaxy	10^{11}	10 kpc	10^8 years	10^{16} years
Galactic DM halos	10^{67}	300 kpc	10^9 years	10^{74} years
Cluster of Galaxies	10^3	3 Mpc	10^8 years	10^9 years

Two body relaxation is not important in determining the evolution of galaxies or their DM halos, i.e. collisions can be ignored. They are also only of secondary importance in globular clusters and galaxy clusters as $t_{\mathrm{relax}} > t_{\mathrm{dyn}}$, but shorter than t_{Hubble}.

Softening

The number of dark matter particles in a galaxy halo may well exceed $N = 10^{67}$ ($\sim 10^{12} M_\odot c^2 / 1 \text{GeV}$). Hence the two-body scattering time $\tau_{2body} \gg t_{Universe}$ and so 2-body scattering should be completely neglible. However N-body simulations have been done with only thousands to millions of particles per galactic halo which would result in a 2-body relaxation time of $\sim 10^{14}$ years.

The problem is that an N-body particle doesn't represent a single point mass, but really a cloud of much less massive particles. Thus it doesn't make sense to use a $1/r$ potential down to zero separation. Hence N-body simulations (except those of point masses like stars in a globular cluster) use softened potentials/force laws. e.g

$$\Phi = \frac{-GM}{(r^2 + \epsilon^2)^{1/2}}. \qquad \text{Plummer Potential}$$

Softening

The form of softening most often employed these days is to adopt the following spline kernel as a description of the density profile of a particle

$$W = \frac{8}{\pi} \begin{cases} 1 - 6x^2 + 6x^3 & 0 < x < 1/2 \\ 2(1-x)^3 & 1/2 < x < 1 \\ 0 & x > 1 \end{cases}$$

where $x = r/(2.8\epsilon)$. The 2.8 was chosen such that at $r = \epsilon$ the force is the same as with the Plummer Potential. At $r > 2.8\epsilon$ one recovers the Newtonian Potential.

The effective resolution of a simulation is determined by an interplay of the mass resolution, force softening and time step. (Power et al 2003)

Timesteps

The timestep Δt used to advance the position of a particle needs to be small enough so that the correct orbit is followed. Many choices of timestep criteria are used. Eg. Power et al (2003) consider

$$\Delta t = \begin{cases} \eta_{a\epsilon}\sqrt{\epsilon/a}, \\ \eta_a/a, \\ \eta_{a\sigma}(\sigma/a), \\ \eta_\rho(G\rho)^{-1/2}, \\ \eta_{\sigma\rho}\min[(G\rho)^{-1/2}, (\sigma/a)], \end{cases}$$

and find the best results with $\Delta t \approx 0.2\sqrt{\epsilon/a}$, where ϵ is the softening scale and a the local acceleration.

GADGET has individual timesteps for each particle and can be based on this criteria.

Parallel Computing

The goal in parallel computing is to share the workload between N processors and aim to gain a speed up of a factor N (i.e. achieve scaling), but this is hampered by many issues.

TROUBLING ASPECTS OF PARALLELIZATION

▶ **Incomplete parallelization**
The residual serial part in an application limits the theoretical speed-up one can achieve with an arbitrarily large number of CPUs ('Ahmdahl's Law'), e.g. 5% serial code left, then parallel speed-up is at most a factor 20.

▶ **Parallelization overhead**
The bookkeeping code necessary for non-trivial communication algorithms increases the total cost compared to a serial algorithm. Sometimes this extra cost increases with the number of processors used.

▶ **Communication times**
The time spent in waiting for messages to be transmitted across the network (bandwith) and the time required for starting a communication request (latency).

▶ **Wait times**
Work-load imbalances will force the fastest CPU to idly wait for the slowest one.
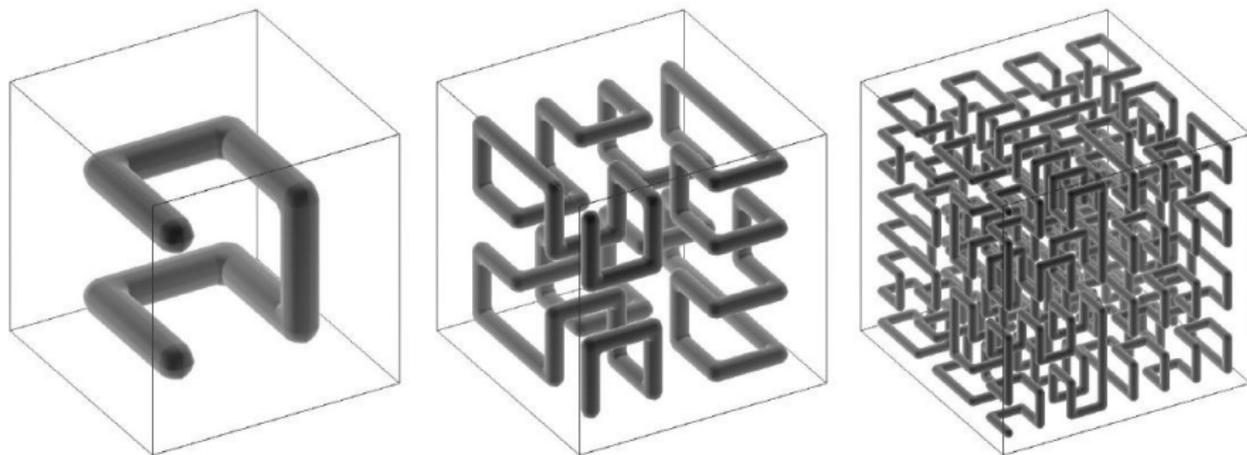
Gadget 2

The public Gadget2 code (Springel 2005) combines a variety of techniques to achieve good scaling.

The large scale force is computed by an FFT mesh. One can use standard libraries that are fully parallel and scale very well.

This leaves only a short range force which allows the forces in separated regions to be calculated independently.

The computational volume is decomposed into separate domains using a space filling Peano-Hilbert curve.

Gadget2

Particles are gridded up into cells and the cells assigned to sections of the P-H curve. The curve is then broken into N contiguous sections each containing 1/Nth of the particles. These are the domains that are passed to each processor.

Domains are obtained by cutting the Peano-Hilbert curve into segments

Tree on Process 1

Tree on Process 3
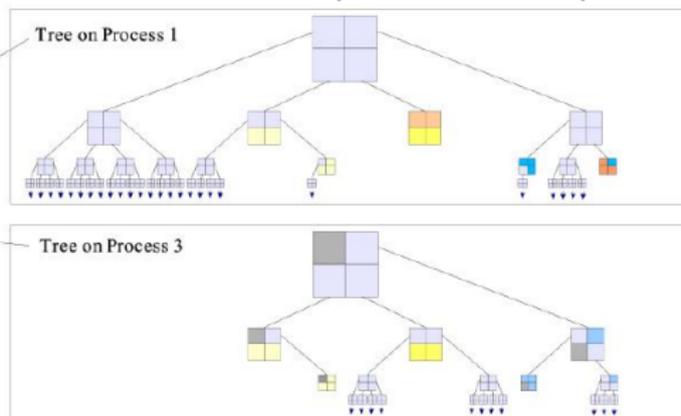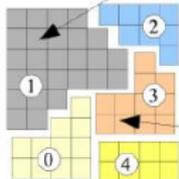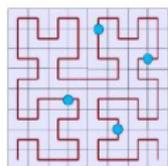
A tree structure is used to work out which neighbouring domains/processors information has to be exchanged with in order to compute the total force.

Gadget2

This works well sharing the load in simulations of cosmological volumes.
It is not so good for simulations of individual objects (e.g. a single galaxy halo of galaxy cluster) as it can end up with all the particles that need the shortest timestep being on the same processor.

Gadget3 subdivides the P-H curve further to avoid this
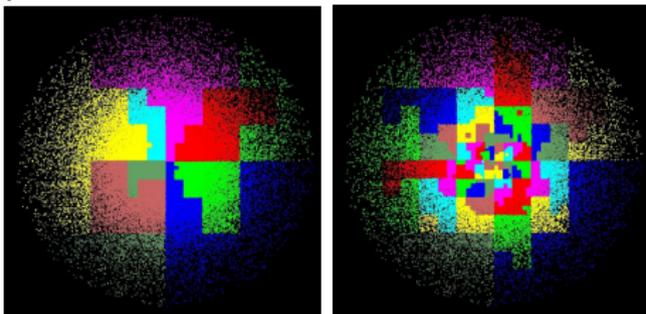
Good scaling it achieved with hundreds of processors.
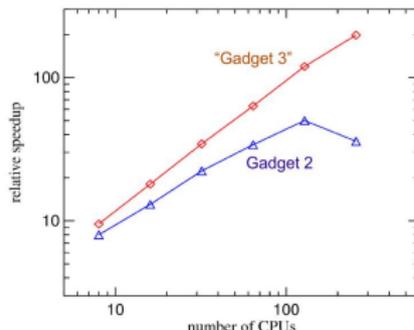
Other parallel algorithms

There is active research into other parallelization strategies to achieve scaling to much high processor numbers.
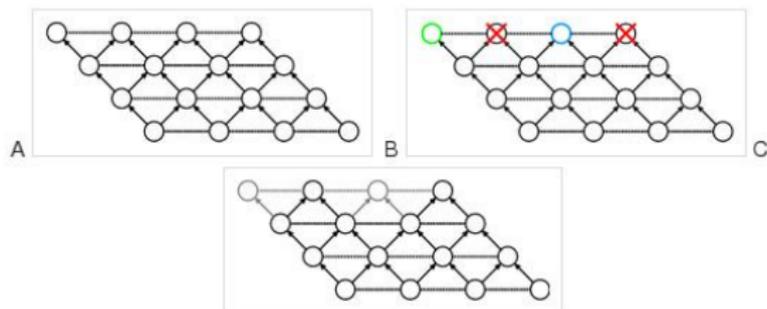
One approach is task based parallelism.

Figure 1: Task-based parallelism. (**A**) Decomposition of a computation into individual tasks (circles) with dependencies (arrows) and conflicts (dashed lines). (**B**) Two threads select two tasks with no unresolved dependencies or overlapping conflicts (coloured circles) and block the tasks with which they conflict (red crosses). (**C**) When the threads are done executing the tasks, they are removed from the task queue, along with their conflicts and dependencies.

SWIFT

SWIFT is N-body and hydrodynamic code (Pedro Gonnet, Matthieu Schaller & Tom Theuns) that uses graph-based domain decomposition. It uses information from a task graph to decompose the simulation domain such that the work and not just the the data, as in the case in the GADGET schemes, is equally distributed amongst all processor nodes.
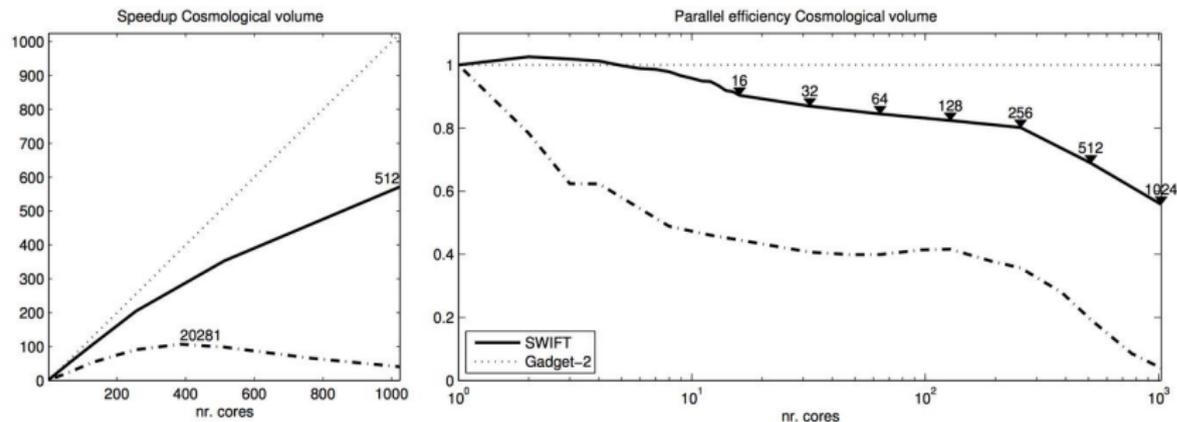
Strong scaling test comparison of GADGET and SWIFT.

Cosmological Initial Conditions

How do we set up an initial particle distribution with positions and velocities consistent with the CMB (I.E. a Gaussian random field with the power spectrum of a cosmological model whose parameters have been constrained by the CMB)?
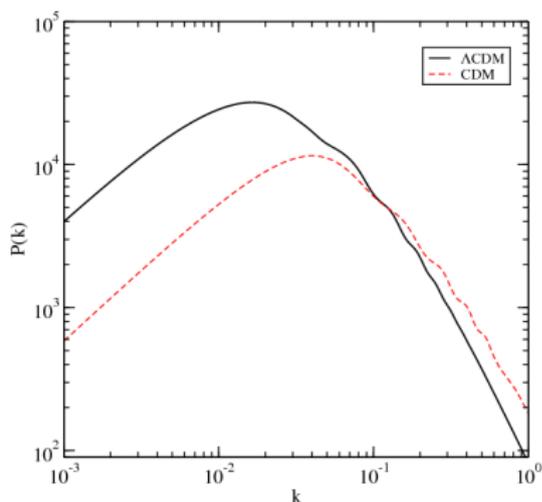
CAMB or CMBFAST model fit to the CMB data provides a well specified model for the mass power spectrum $P(k) = \langle |\delta_k|^2 \rangle$. From this one can define a particular realization[a] of a Gaussian random field $\delta(\mathbf{x}) = \sum_{\mathbf{k}} \delta_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{x}}$, but how do we set up consistent particle positions and velocities.

---

[a]Real and Imaginary components of $\delta_{\mathbf{k}}$ chosen independently from a zero mean Gaussian with variance $P(k)(\Delta k)^3 / 2$

Fluid Equations

The Zel'dovich approximation provides a very convenient way of setting up the initial conditions of a cosmological N-body simulation. At high redshift the amplitude of density perturbations are very low (10^{-5} at recombination) and so one can attempt to use perturbation theory to solve the full set of equations describing a continuum of particles (the fluid approximation).

$$\frac{\partial \rho}{\partial t} + \underline{\nabla}.(\rho \underline{v}) = 0 \qquad \text{Continuity}$$

$$\frac{\partial \underline{v}}{\partial t} + (\underline{v}.\underline{\nabla})\underline{v} = -\underline{\nabla}\Phi \qquad \text{Euler}$$

$$\nabla^2 \Phi = 4\pi G \rho \qquad \text{Poisson}$$

Fluid Equations in Comoving Coordinates

We want to work instead in comoving coordinates, $\underline{r} = a(t)\underline{x}$, and look at perturbations to the density in the mass dominated, $\rho = \langle\rho\rangle(1+\delta) = (\rho_0/a^3)(1+\delta)$ and potential, $\nabla^2\hat{\Phi} = \nabla^2\Phi - 4\pi G\langle\rho\rangle a^2$.

Applying the chain rule gives the following substitutions

$$\underline{\nabla} \to \frac{\underline{\nabla}}{a} \quad \text{and} \quad \frac{\partial}{\partial t} \to \left.\frac{\partial}{\partial t}\right)_{\underline{x}} - \dot{a}\underline{x}.\frac{\underline{\nabla}}{a}$$

and so the fluid equations become:

$$\frac{\partial\delta}{\partial t} + \underline{\nabla}.\dot{\underline{x}} + \underline{\nabla}.(\dot{\underline{x}}\delta) = 0 \qquad \text{Continuity}$$

$$\ddot{\underline{x}} + 2\frac{\dot{a}}{a}\dot{\underline{x}} + (\dot{\underline{x}}.\underline{\nabla})\dot{\underline{x}} = \frac{\nabla\hat{\Phi}}{a^2} \qquad \text{Euler}$$

$$\frac{\nabla^2\hat{\Phi}}{a^2} = 4\pi G\langle\rho\rangle\delta \qquad \text{Poisson}$$

Fourier Transformed Fluid Equations

Take the Fourier Transform of each term e.g.

$$\delta = \sum_{\underline{k}} \delta_{\underline{k}} e^{-i\underline{k}.\underline{x}}$$

and using $\underline{v}$ to now denote the comoving velocity $\underline{\dot{x}}$

$$\underline{\nabla}.\underline{v} = \underline{\nabla}.\sum_{\underline{k}} \underline{v}_{\underline{k}} e^{-i\underline{k}.\underline{x}} = \sum_{\underline{k}} -i\underline{k}.\underline{v}_{\underline{k}} e^{-i\underline{k}.\underline{x}}.$$

The fluid equations become

$$\frac{\partial \delta_{\underline{k}}}{\partial t} - i\underline{k}.\underline{v}_{\underline{k}} - \sum_{\underline{k}'} i\underline{k}.\underline{v}_{\underline{k}-\underline{k}'}\, \delta_{\underline{k}'}\, e^{-i\underline{k}.\underline{x}} = 0 \qquad \text{Continuity}$$

$$\frac{\partial \underline{v}_{\underline{k}}}{\partial t} + 2\frac{\dot{a}}{a}\underline{v}_{\underline{k}} - \sum_{\underline{k}'} i[\underline{v}_{\underline{k}'}.(\underline{k}-\underline{k}')]\,\underline{v}_{\underline{k}-\underline{k}'} e^{-i\underline{k}.\underline{x}} = \frac{i\underline{k}\hat{\Phi}_{\underline{k}}}{a^2} \qquad \text{Euler}$$

$$\frac{\hat{\Phi}_{\underline{k}}}{a^2} = \frac{-4\pi G \langle\rho\rangle \delta_{\underline{k}}}{k^2} \qquad \text{Poisson}$$

Linearized Fourier Transformed Fluid Equations

If we make the linear approximation and discard terms that are second order in δ and/or $\underline{v}$ we find

$$\frac{\partial \delta_{\underline{k}}}{\partial t} - i\underline{k}.\underline{v}_{\underline{k}} = 0 \qquad \text{Continuity}$$

$$\frac{\partial \underline{v}_{\underline{k}}}{\partial t} + 2\frac{\dot{a}}{a}\underline{v}_{\underline{k}} = \frac{i\underline{k}\hat{\Phi}_{\underline{k}}}{a^2} \qquad \text{Euler}$$

$$\frac{\hat{\Phi}_{\underline{k}}}{a^2} = \frac{-4\pi G\langle\rho\rangle\delta_{\underline{k}}}{k^2} \qquad \text{Poisson}$$

Eliminate $\underline{v}_{\underline{k}}$ from Euler and Continuity equations and $\hat{\Phi}_{\underline{k}}$ using the Poisson equation gives

$$\frac{d^2\delta_k}{dt^2} + 2\frac{\dot{a}}{a}\frac{d\delta_k}{dt} = 4\pi G\langle\rho\rangle\delta_{\underline{k}} \quad \text{Linear Perturbation}$$

Each $\underline{k}$-mode evolves independently.

Growing and Decaying Modes

The solution to the linear perturbation equation depends on the background cosmology (Ω and Λ) through the dependence of $\langle \rho(t) \rangle$. For Einstein-de Sitter, $\Omega_{\mathrm{m}} = 1$, $a \propto t^{2/3}$ and

$$4\pi G \rho_0 t_0^2 = \frac{2}{3}.$$

giving

$$\frac{d^2 \delta_{\mathrm{k}}}{dt^2} + \frac{4}{3}\frac{1}{t}\frac{d\delta_{\mathrm{k}}}{dt} = \frac{2}{3}\frac{\delta_{\underline{k}}}{t^2}.$$

Its general solution is

$$\delta_{\underline{k}} = A_{\underline{k}} t^{2/3} + B_{\underline{k}} t^{-1}$$

The $A_{\underline{k}}$ are the growing modes and $B_{\underline{k}}$ are the decaying modes.
To determine both one needs both initial densities, $\delta_{\underline{k}}$, and velocities $\underline{v}_{\underline{k}}$.

Zel'dovich Approximation

Generally we can ignore the decaying modes as even if they exist they decay away at early times. Hence if we assume pure growing mode we only need the initial $\delta_{\underline{k}}$ to specify the initial conditions.

$$\delta_{\underline{k}} = A_{\underline{k}} t^{2/3} \quad \text{and} \quad \underline{v}_{\underline{k}} = -\frac{2}{3} \frac{i\underline{k}}{k^2} A_{\underline{k}} t^{-1/3}$$

Inverse Fourier Transforming

$$\delta(\underline{q}) = \sum_{\underline{k}} \delta_{\underline{k}} \, e^{-i\underline{k}\cdot\underline{q}}$$
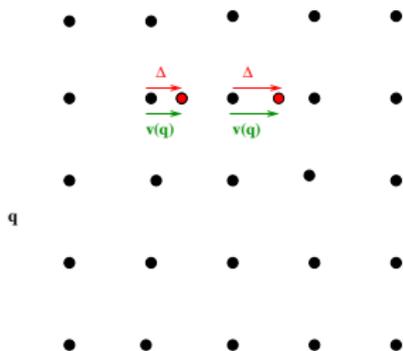
$$\underline{v}(\underline{q}) = \sum_{\underline{k}} \frac{2}{3} \frac{-i\underline{k}}{k^2} A_{\underline{k}} \, t^{-1/3} \, e^{-i\underline{k}\cdot\underline{q}} = \sum_{\underline{k}} \frac{2}{3} \frac{-i\underline{k}}{k^2} \frac{\delta_{\underline{k}}}{t} \, e^{-i\underline{k}\cdot\underline{q}}$$

Integrating with time to get the corresponding displacement

$$\Delta\underline{q} = t^{2/3} \sum_{\underline{k}} \frac{-i\underline{k}}{k^2} A_{\underline{k}} \, e^{-i\underline{k}\cdot\underline{q}} = \sum_{\underline{k}} \frac{-i\underline{k}}{k^2} \delta_{\underline{k}} \, e^{-i\underline{k}\cdot\underline{q}}$$

Zel'dovich Approximation

These equations allow one to set up the initial conditions of an N-body simulation by displacing particles from an initial grid of coordinates $\underline{q}$ by the displacements $\Delta\underline{q}$ and the corresponding velocities $\underline{v}(\underline{q})$.

Sometimes particle positions are perturbed away from a random glass configuration rather than a uniform grid.
Jenkins (2009) has extended these methods to 2nd order Lagrangian Perturbation theory, allowing simulations to start at lower redshift.
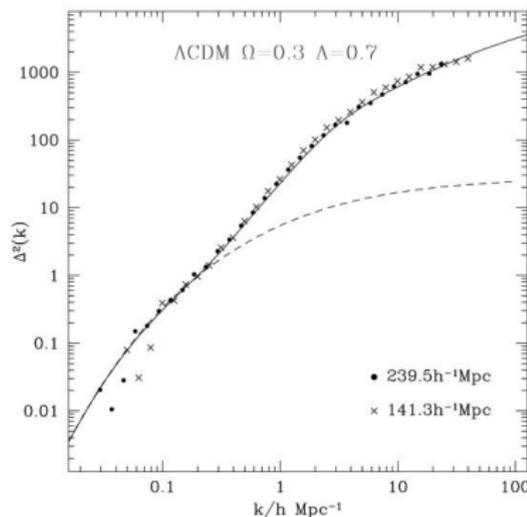
Linear and Nonlinear evolution

When all mode amplitudes $|\delta_{\mathbf{k}}| \ll 1$, mode coupling is negligible and all modes grow at the same rate.

$$\delta_{\mathbf{k}} = \delta_{\mathbf{k},0}\, D(a).$$

In hierarchical models such as CDM small scales go non-linear first. The non-linear power spectrum differs from the initial linear shape first on the smallest scales and then progressively on larger scales. Here
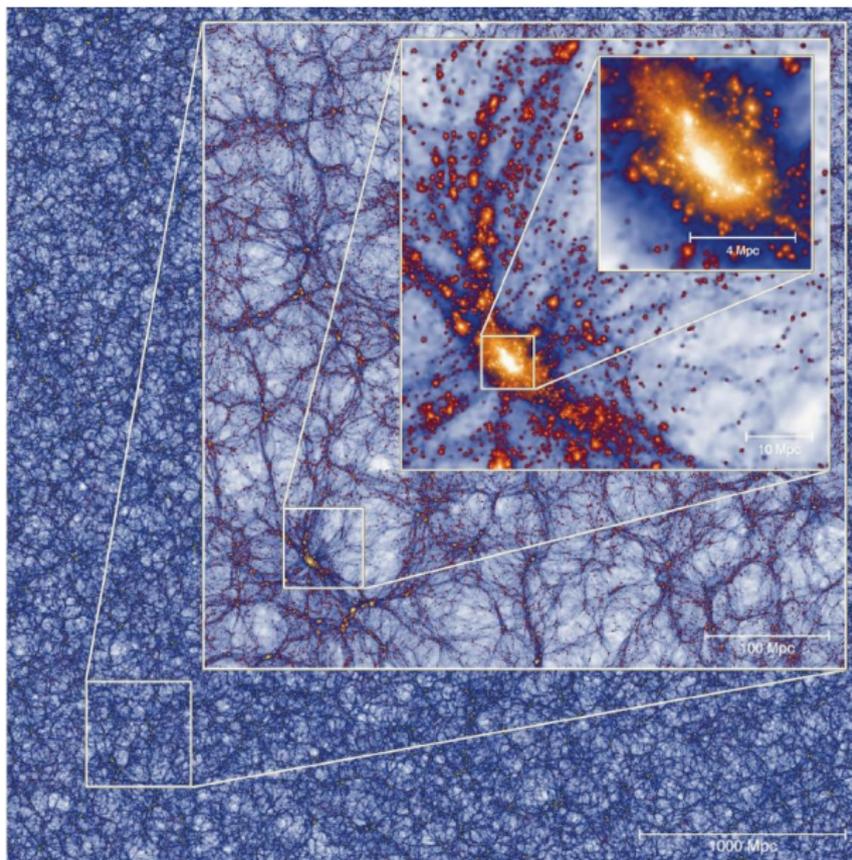
$$\Delta^2(k) = k^3 P(k)$$

is the matter power spectrum expressed in dimensionless form. $\Delta^2(k)$ is the contribution to the variance in the density fluctuations from each logarithmic interval of wavenumber $\Delta^2(k) = \mathrm{d}\sigma^2/\mathrm{d}\ln k$.

(Jenkins et al 1998)

The Cosmic Web

Nonlinear gravitational evolution leads to the formation of the Cosmic Web, a network of sheets, filaments, voids and clusters. Within the filaments are dense virialized halos of dark matter which flow along the filaments and merge at their intersections.
The growth of these dark matter halos is hierarchical.

MXXL: Angulo et al (2012)

The Halo Mass Function

Dark matter halos which can be
identified with a number of group
finding algorithms are the locations
where galaxies and galaxy clusters form.
Thus an important basic statistic is the
number density of halos as a function of
their mass.

This evolves with time and depends on
the shape of the initial power spectrum,
but when expressed in terms of the
variance of the linear mass field is
essentially universal.

$$f(\sigma) = \frac{M}{\rho_0} \frac{dn}{d\log \sigma} = \frac{M}{\rho_0} \frac{dn}{d\log M} \frac{d\log M}{d\log \sigma}$$


Jenkins et al 2002
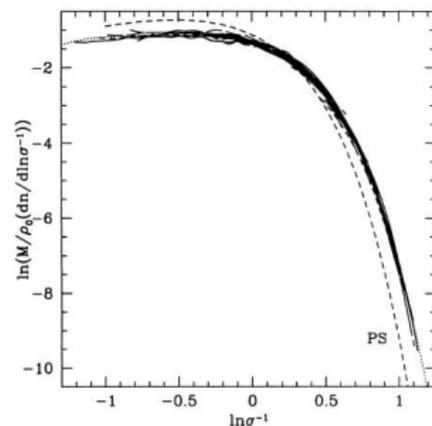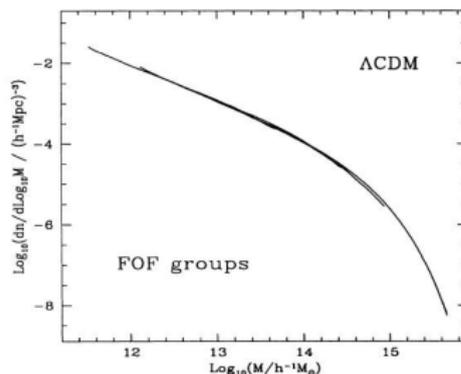
Halo Merger Trees and Subhalos

Halos are not featureless monolithic entities with NFW density profiles:

$$\rho(r) = \frac{4r_{\rm s}^3 \, \rho_{\rm s}}{r(r_{\rm s} + r)^2}$$

where the concentration $c = r_{200}/r_{\rm s}$. They contain substructure in the form of subhalos which are remnants of their hierarchical formation and that can be identified by codes such as SUBFIND. The hierarchical formation can be described in terms of a merger tree .

Virgo Aquarius simulation AqA2
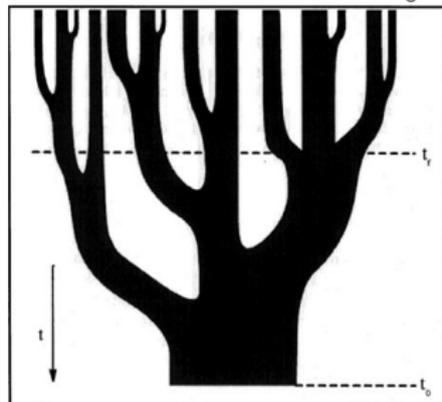
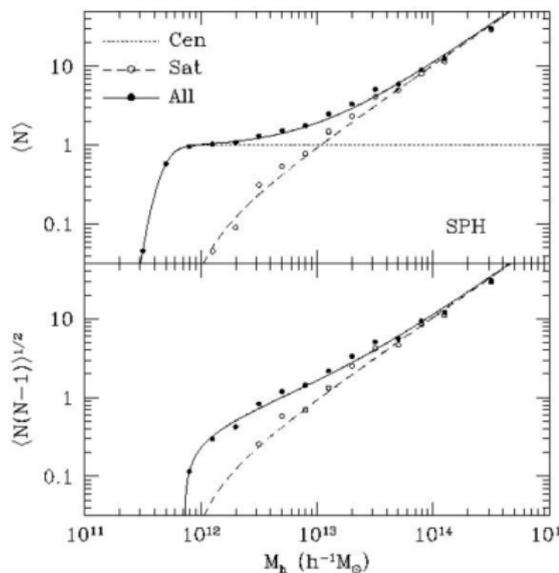# The Halo Occupation Distribution (HOD)

Assuming galaxies reside in halos the simplest
way of populating a DM N-body simulation
with galaxies (of a particular type) is to specify
the Halo Occupation Distribution (HOD):
$P(N|M)$.

The probability that a halo of mass M hosts N
galaxies (e.g. Weinberg & Berlind 2002). This
is normally specified in terms of the mean
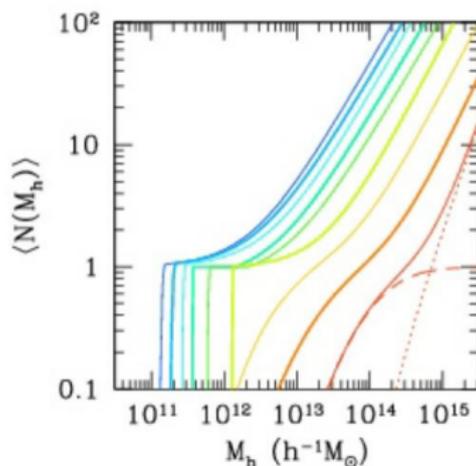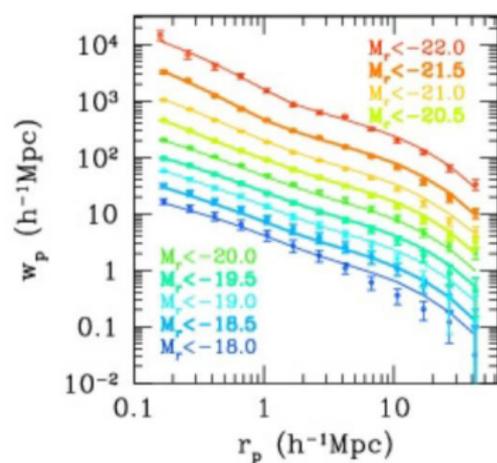number of satellites and central galaxies per
halo

$$\langle N(M_h) \rangle = \frac{1}{2} \left[1 + \mathrm{erf} \left(\frac{\log M_h - \log M_{\min}}{\sigma_{\log M}} \right) \right]$$
$$\times \left[1 + \left(\frac{M_h - M_0}{M_1'} \right)^\alpha \right],$$

and assuming the satellites number is Poisson
distributed about this mean. (There can only
be 0 or 1 central per halo).

Zehavi, Zheng, Weinberg ... (2011)

The Halo Occupation Distribution (HOD)

Zehavi, Zheng, Weinberg ... (2011)

The observed galaxy clustering and its dependence on luminosity can be fit by an HOD whose mean halo mass increases with luminosity.

This framework provides a very simple way of populating an N-body simulation with a particular class of galaxies e.g. LRGs. All that is required is a simulation in which the halos hosting the relevant galaxies are resolved.

Subhalo Abundance Matching (SHAM)

The HOD technique ignores the differing formation histories of individual DM halos and assumes galaxy properties only depend on halo mass. The SHAM technique requires higher resolution simulations in which subhalos are resolved but then makes use of the subhalo populations in individual halos.
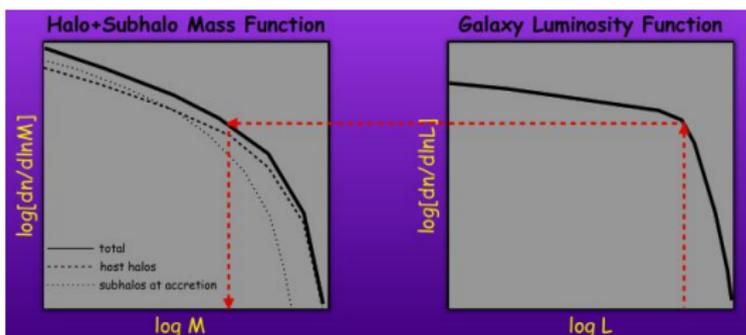
van den Bosch

Abundance matching

$$n_{\mathrm{gal}}(> M_*) = n_{\mathrm{halo}}(> M_{\mathrm{halo}})$$

defines the relation $M_*(M_{\mathrm{halo}})$ or $L(M_{\mathrm{halo}})$ between galaxy stellar mass or luminosity and subhalo mass.

For subhalos the mass used here is not the current mass of the subhalo but the mass it had at infall. (Popular variants exist including using V_{max} of the halo rather than mass.)

Kravtsov et al 2004, Vale & Ostriker 2004, Conroy et al 2006, Moster et al 2010, Behroozi et al 2010

Subhalo Abundance Matching (SHAM)

The parameter free (or one parameter is one allows scatter in the relation) SHAM technique has been remarkably successful at modelling a large range of observations.


Source: Conroy, Wechsler & Kravtsov (2006)
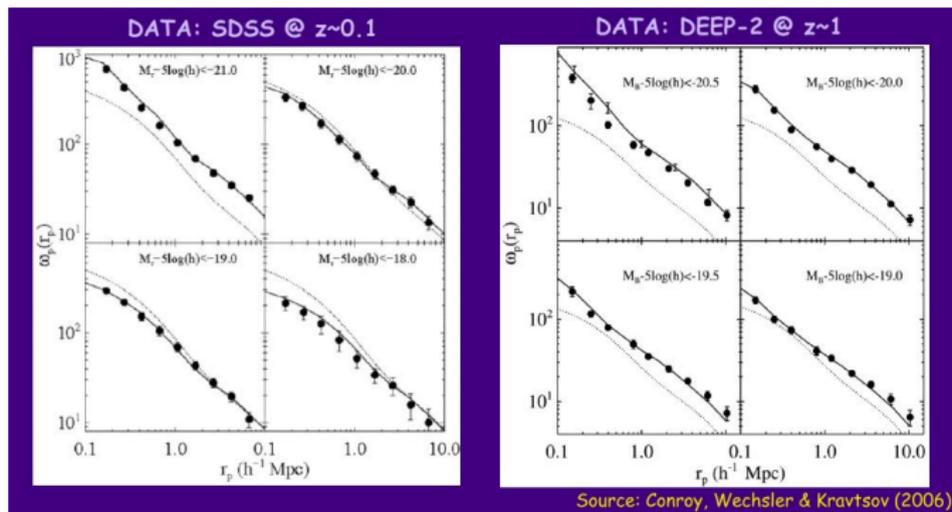
Variants of the SHAM model now exist in which a second subhalo property is correlated with a second galaxy property. E.g. at fixed luminosity and subhalo mass galaxy colour is ranked against accretion redshift (Hearin & Watson 2013)

Semi-Analytic Modelling

Here halo merger trees taken from DM simulations are taken as the framework in which to build an analytic model of galaxy formation.

These models provide much more than the statistical descriptions given by the HOD and SHAM approaches. As well as the location of each galaxy the model provides the evolutionary history of each galaxy and observable properties such as broad band luminosities.

Cole et al 2000
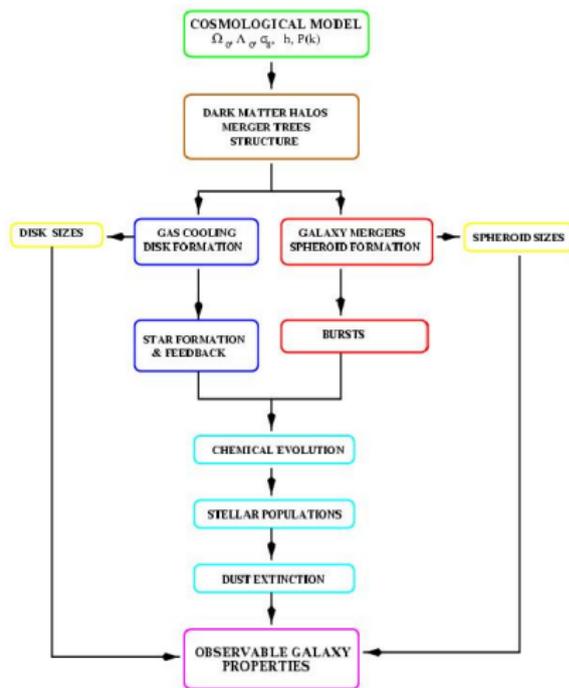
Semi-Analytic Modelling

Semi-analytic models include prescriptions for:

- Gas shock heating
- Cooling and disk formation
- Star formation
- Stellar feedback
- Galaxy mergers
- Merger induced star bursts
- AGN feedback
- Chemical Evolution
- Stellar Population Synthesis
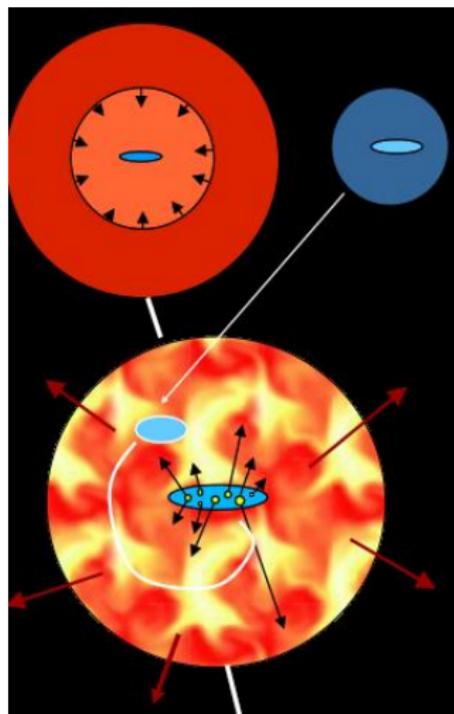
Wechsler et al 2001

Lightcone construction

Once an N-body simulation has been populated by galaxies then one must cut out the survey volume. The survey might be deeper than the simulation box, requiring the simulation to be replicated.

Typically one should not neglect evolution with lookback time. Hence the galaxies should be extracted on the past lightcone of the observer.

Thus there is an issue of how to interpolate galaxy positions and properties or deal with discontinuities.
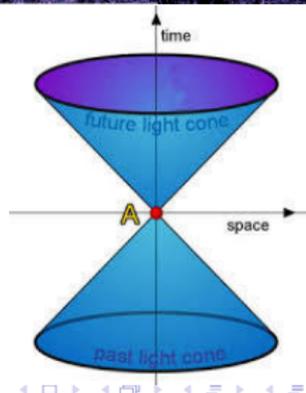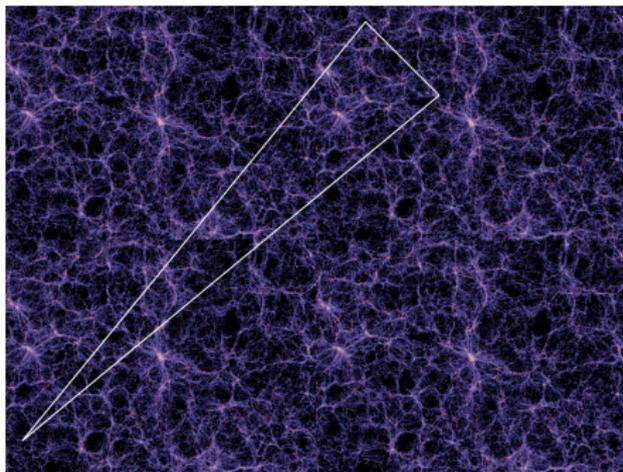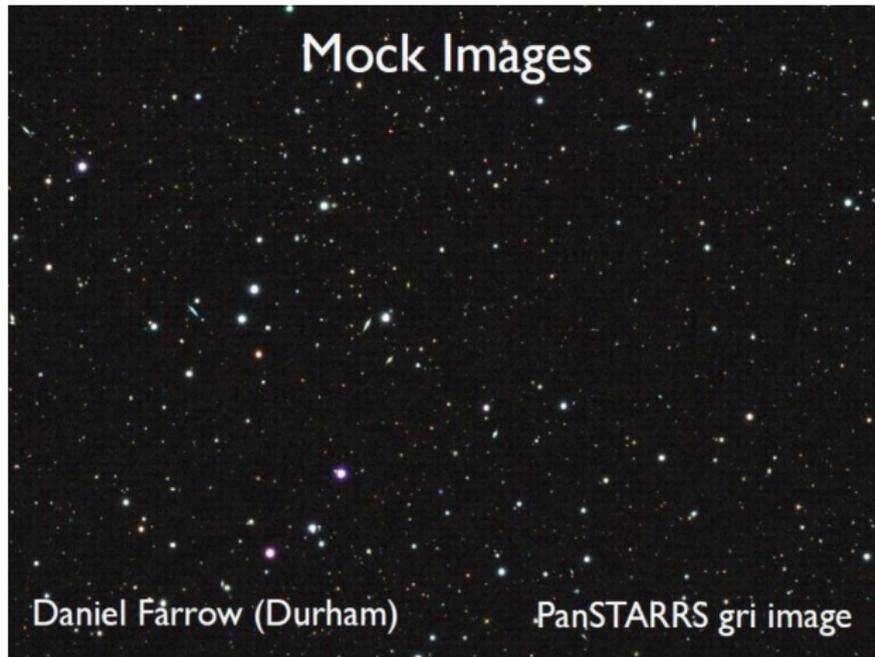
Lightcone construction
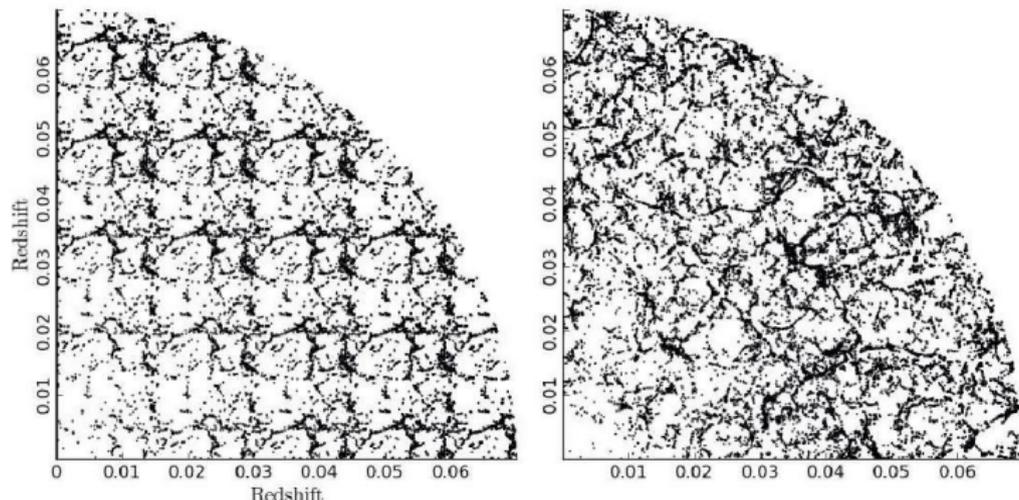
Bernyk et al 2014

Some constructors simply repeat the periodic structures in the way assumed during the N-body calculation. Others rotate and shift the replicas to remove the periodic structure at the expense of creating discontinuities.

Simulated Images

Mock Images

Daniel Farrow (Durham)   PanSTARRS gri image

Normally the output of these synthesized surveys is a mock catalogue listing galaxy properties, but some now go further and create mock images (e.g. Bernyk et al 2014 https://tao.asvo.org.au/tao/mock_galaxy_factory/). Ultimately these can be used in end-to-end testing of data reduction pipelines (e.g. Euclid).

Survey Selection Function

Selecting galaxies from the volume limited light cone that meet the selection criteria for being included in the observational catalogue.

In principle this is easiest for the semi-analytic models as one can directly impose the same observational criteria. However if the semi-analytic does not perfectly reproduce the observed properties then this can produce large systematic differences.

By construction, SHAM models match the observations for the galaxy property(ies) that were abundance matched but in general these are not the properties on which the observed catalogue was selected. One must additionally model the relation between the SHAMed and observed property.

HOD models most naturally produce only volume limited catalogues. They are mainly used to model volume limited subsets of an observed catalogue. (LRG catalogues are close to volume limited.)

Targeting Completeness

All modern multi-object spectrographs have targeting limitations. In SDSS fibres could not be placed closer than $55''$. In DESI the efficiency of targeting depends on the target density and scale on which it fluctuates.

Mock catalogues can be used to test the efficiency of various observing and targeting strategies, though this can also be done using real data if the photometric catalogue which will be targeted already exists.

Clustering analysis must employ a scheme which attempts to compensate for the targeting incompleteness. Mock catalogues are very useful for testing the effectiveness and accuracy of such schemes.
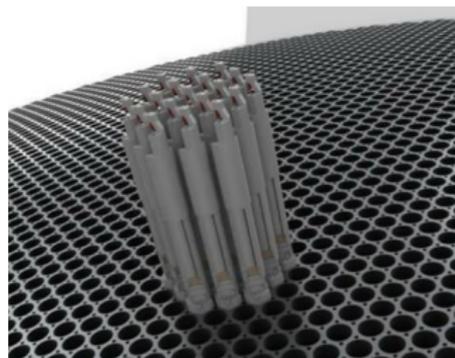
The 5000 fibre DESI focal plane

Redshift Completeness
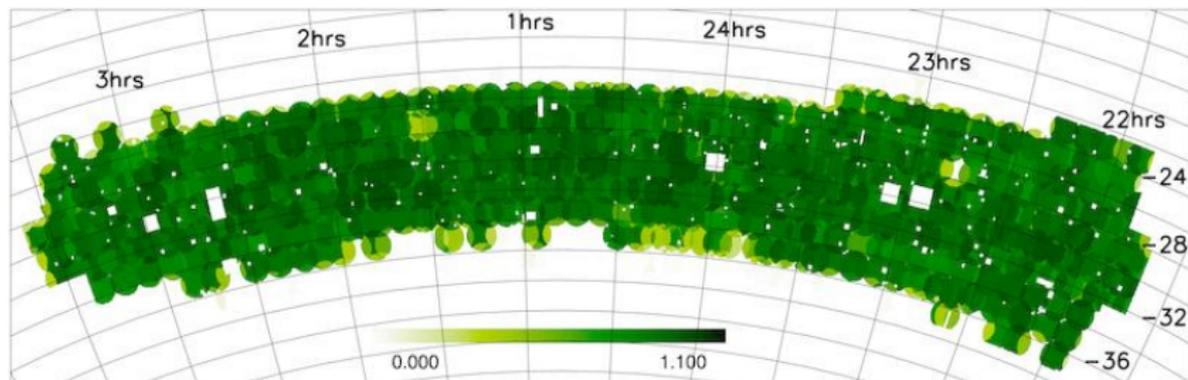
Map of the redshift completeness in 2dFGRS SGP strip

Targeting incompleteness is one contribution to redshift incompleteness. It is relatively easy to correct for as for it one can assume the incompleteness is independent of redshift, z.

More problematic is incompleteness due to redshift measurement failures. This will correlate with the photometric properties of the galaxies (e.g. surface brightness) and can depend directly on redshift (e.g. where emission lines overlap sky lines).

Pros and Cons

Method	Pros	Cons
HOD	Simple Does not need to resolve subhalos	Does not provide galaxy properties
SHAM	Simple	Needs to resolve subhalos Does not provide all galaxy properties
Semi-analytic	Provides **all** galaxy properties	Complex Hard to tune to match observations

Summary

1. Mock galaxy catalogues are important in the planning, analysis and exploitation of large scale galaxy redshift surveys.

2. They require exploiting state-of-the-art techniques and computers to generate the necessary cosmological N-body simulations.

3. As it is not yet possible to simulate the physics of galaxy formation within cosmological volumes, different techniques have to be used to "paint" galaxies onto Dark Matter simulations.

4. Different compromises have to be made depending on the primary goal of the mocks. E.g.

   1. where large scale statistics is foremost, HOD population of low resolution simulations;
   2. where smaller scales and survey selection function is important SHAM models of higher resolution simulations;
   3. where clustering is used to test the physics of galaxy formation models one needs semi-analytic models or full hydro simulations.

Resources

- Merson et al 2013 http://adsabs.harvard.edu/abs/2013MNRAS.429..556M semi-analytic lightcome mocks
- J. Carretero et al 2014 arXiv:1411.3286 mocks from the MICE simulations
- Bernyk et al 2014 arXiv:1403.5270 provide a web interface to a system that can generate tailor made mock catalogues https://tao.asvo.org.au/tao/ Theoretical Astrophysics Observatory
- McBride et al LasDamas mocks http://lss.phy.vanderbilt.edu/lasdamas/
- Chuang et al 2015 http://adsabs.harvard.edu/abs/2015MNRAS.446.2621C EZ mocks using extended Zel'dovich approximation
- Manera et al http://www.marcmanera.net/mocks/ PTHaloes galaxy mocks
- Monaco et al 2013 http://adsabs.harvard.edu/abs/2013MNRAS.433.2389M Pinnochio algorithm for generating dark matter halo catalogues
- Ancient sets of 2dF and SDSS mocks http://astro.dur.ac.uk/~cole/mocks/