Cosmological Simulations and Dark Matter

Ben Oppenheimer

University of Colorado, Boulder

previously at Leiden University

Challenges in Modern Cosmology: Dark Matter and Dark Energy

International Institution of Physics

Natal, Brazil

Three Lecture Outline

Topic Area 1: Cosmological Collisionless N-body simulations

1. Cosmological Initial Conditions
2. Collisionless N-body Simulations
3. Dark Matter Halos in Dark Matter-only Simulations

Topic Area 2: Adding in Baryons- Hydrodynamic Simulations

1. Hydrodynamics
2. Cooling and Photo-Heating
3. Star Formation
4. Feedback
5. Galaxy Formation

Topic Area 3: Deviation from and Challenges to Standard Cold Dark Matter

1. Halo Profiles
2. Missing Satellites Problem and Too Big to Fail
3. Alterntive Types of Dark Matter: Warm, Hot, and Self-Interacting
4. Void Galaxies
5. Cluster Abundance Tests

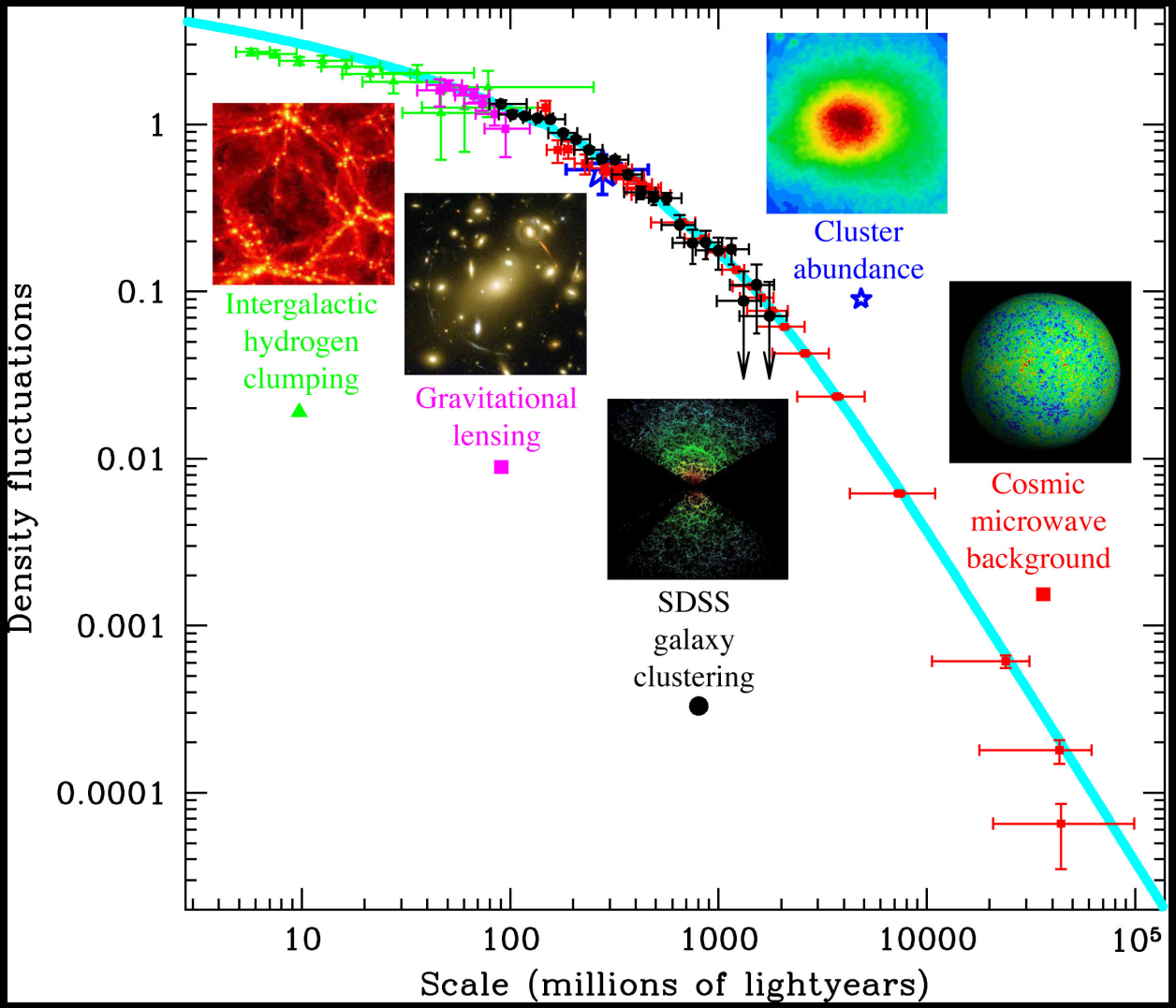
Ingredients of Cosmological Simulations

1. Cosmological Initial Conditions
2. N-body Code: Gravity-only Simulations
3. Hydrodynamics
4. Cooling and Photo-Heating
5. Star Formation
6. Feedback

Cosmological Initial Conditions

Typically, initial conditions (ICs) use ΛCDM cosmology. The ICs use a Gaussian random field to add density perturbations to a cosmological simulation box.

These density perturbations are a result of observations from the CMB at large scales to extrapolating back from galaxy clustering observed in the local Universe and the intergalactic medium via the Lyman-alpha forest on smaller scales.

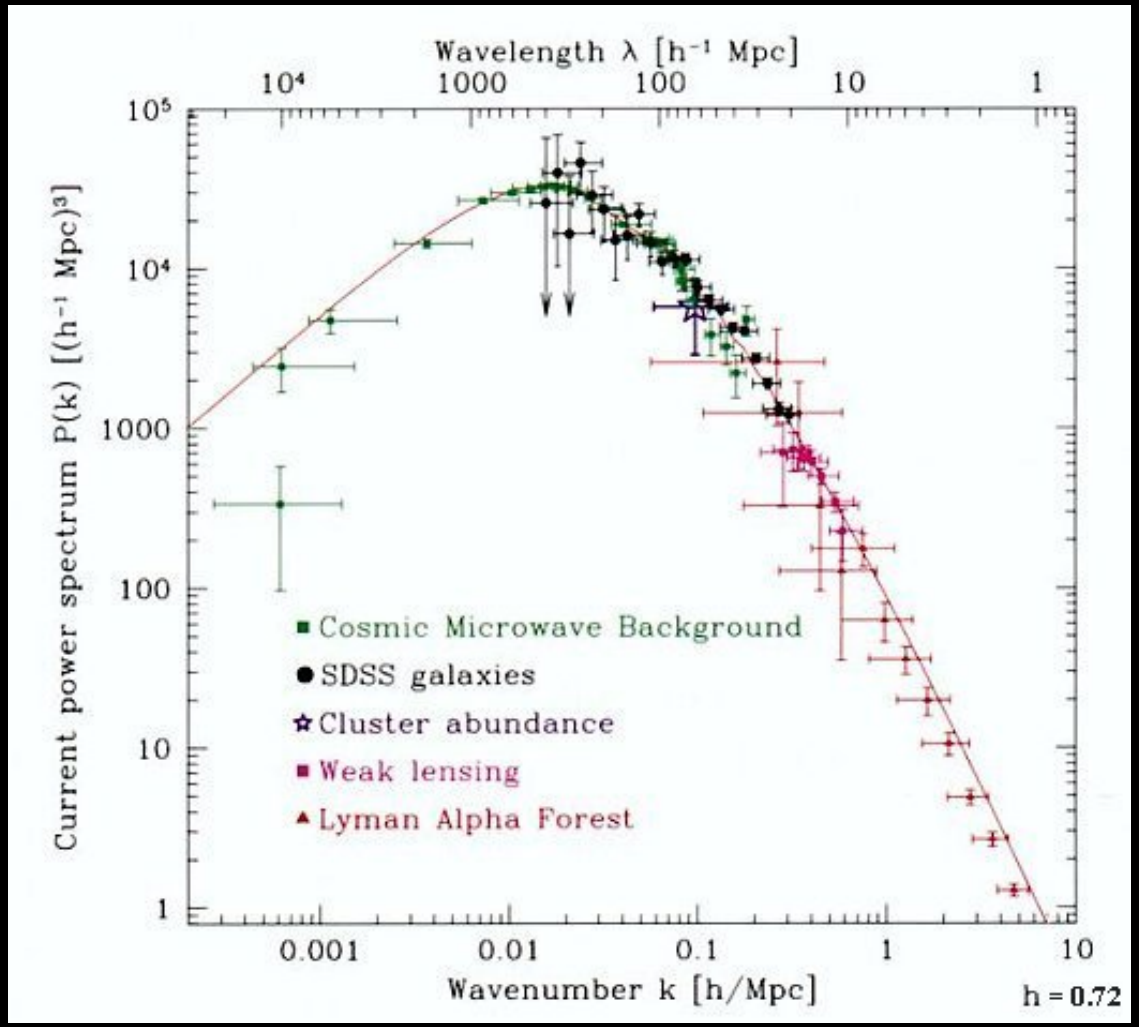


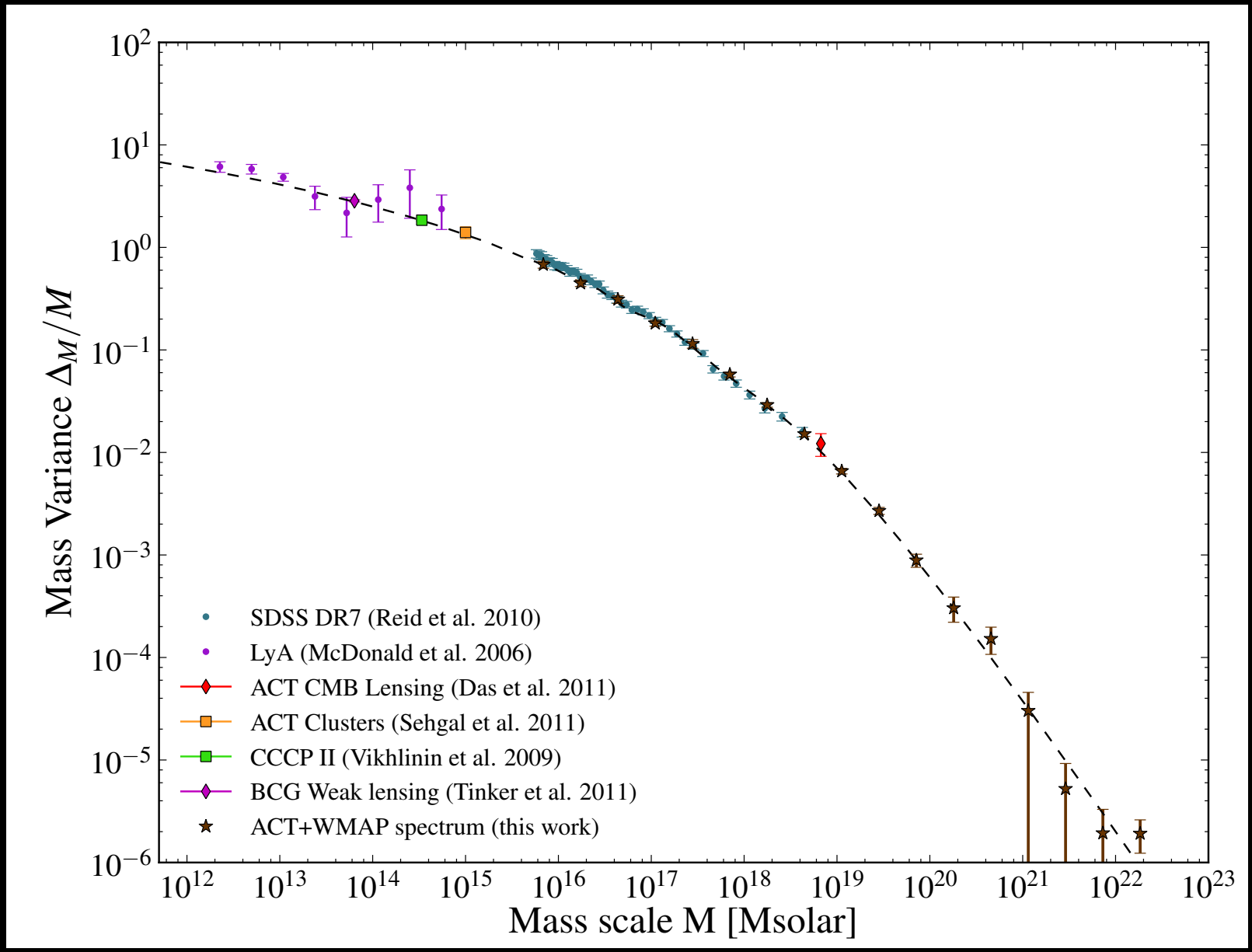
Tegmark (2002)

Cosmological Initial Conditions

Cold Dark Matter is assumed in this power spectrum, and it agrees with constraints of the Lyman-α Forest (smallest scales probed here). We will get into this later, because simulations with warm dark matter (WDM) predict a different Lyman-α forest, which is an observations of the intergalactic medium.

CMB (largest scales, smallest k's) probe largest scales, and galaxy/cluster probes constrain the intermediate scales.

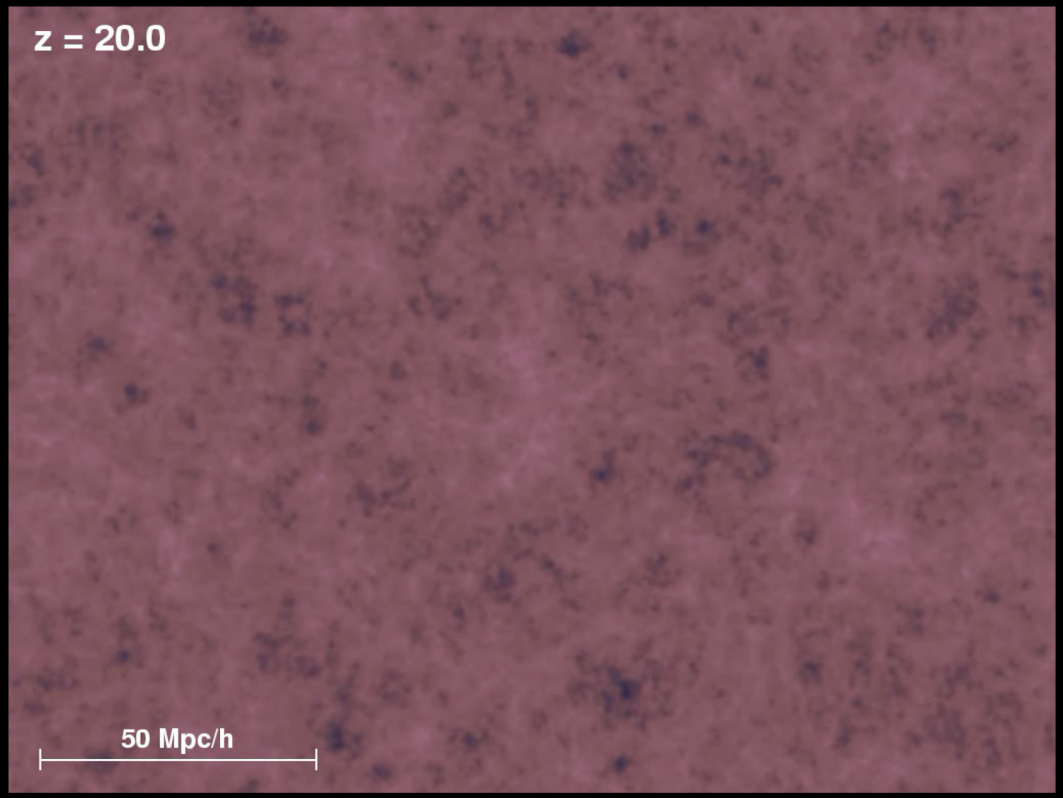




Hlozek+ (2012)- the mass power spectrum of fluctuations, derived from a variety of sources, observationally constrained over 10 orders of magnitude, which is based on the observations from the previous slides, and some new data, which agrees with ΛCDM cosmology.

Collisionless N-body Simulations

A cosmological simulation begins with the power spectrum of initial fluctuations in the linear regimes, a Gaussian random field, and calculates gravitational interactions from z>100 to z=0, while the box undergoes cosmological expansion.



Dark matter-only simulation by Volker Springel

The N-Body Code

An N-body code divide mass elements into discretized collisionless “particles” at the mass resolution of the simulation. A general N-body code will follow positions and velocities of i particles (r_i, v_i) interacting gravitationally. A cosmological simulation uses comoving coordinates wher x = r/*a* and v=dx/dt, where *a* is the scale factor of the box: *a* = 1/(1+z).

The equations of motion are solved

$$\frac{d\mathbf{x}_i}{dt} = \mathbf{v}_i,$$

$$\frac{d\mathbf{v}_i}{dt} + 2H(t)\mathbf{v}_i = -\frac{1}{a^2}\nabla_{\mathbf{x}}\Phi|_i,$$

Where $H(t) = \dot{a}/a$ and the gravitational potential, Φ, are perturbations from a uniform density field that are represented by the Poisson equation.

$$\nabla^2_{\mathbf{x}}\Phi = 4\pi G a^2 \left[\rho(\mathbf{x},t) - \overline{\rho}(t)\right]$$

Direct integration of 6-dimensional phase space (positions & velocities) is impractical, and instead the equations are integrated in a finite timestep that satisfies an error tolerance.

N-Body Method

The N-body process follows the orbits (or paths) of individual particles using the following steps.

- Begin with initial distribution of particles and velocities
- Compute the gravitational potential from the particle distribution
- Compute particle accelerations from potential gradients
- Update particle positions over the timestep
- Repeat the process

Timesteps are determined by acceptable tolerances on the force resolution, using the unitless parameter α_{tol} where $dt = \alpha_{tol}\, u/a$ (u- velocity, a- acceleration). Often local velocity dispersion is used, $dt = \alpha_{tol}\, \sigma/a$.

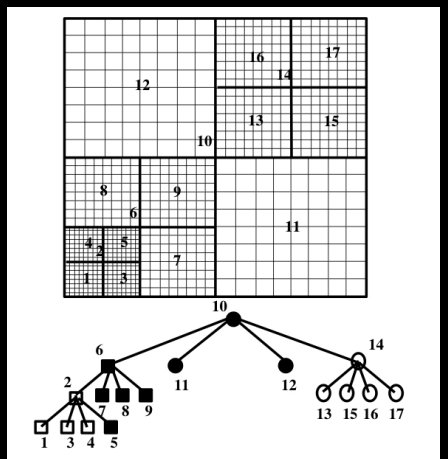
Timesteps can become infinitesimally small if particles orbit very closely to one another, which is why a "softening" length is added to suppress 2-body scattering, such that the force between two particles with mass m is not Gm^2/r^2, but instead $Gm^2/(r+\varepsilon^2)$, where ε, the softening length, is the effective spatial resolution.

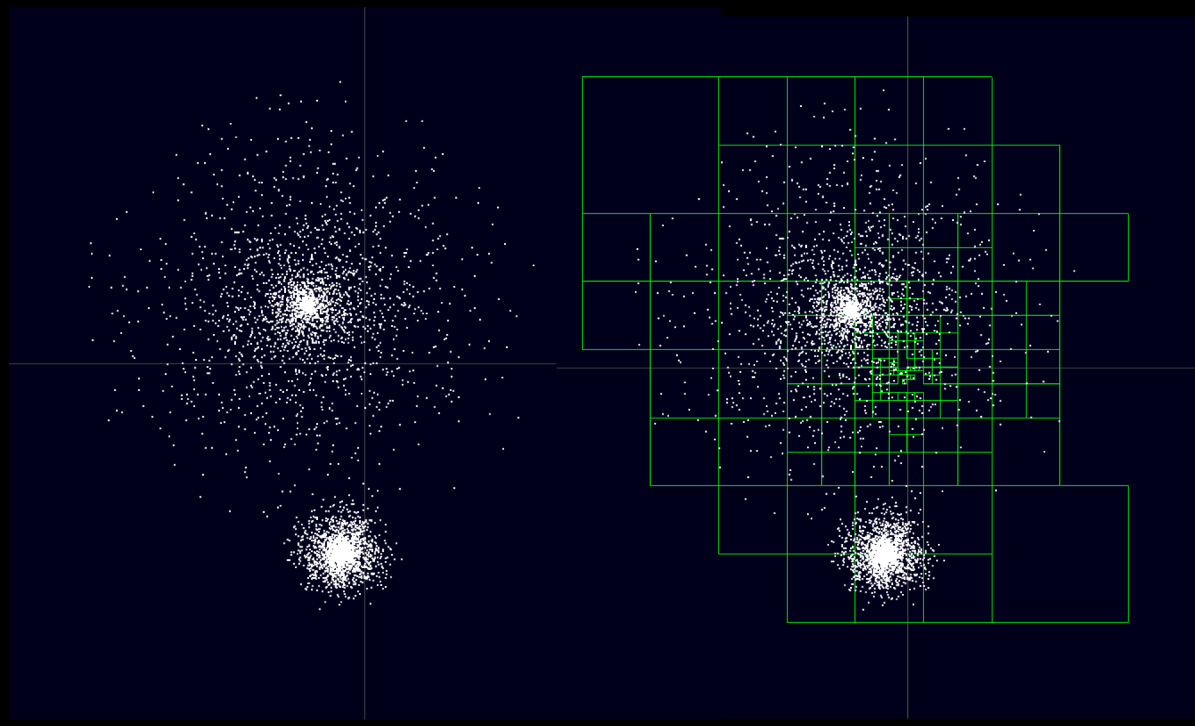
N-Body Method- Tree Codes

Calculating the actual forces is done in two main ways:

First, a tree algorithm, where particles are grouped into a hierarchical structure of cubes within larger cubes, where instead of calculating the acceleration from all N particles on each N particles (N^2 computational expense), the force can be grouped into particles in a cube acting on another particle (NlogN computational expense).

Barnes-Hut (1986) developed an algorithm to split up particles into an octree and then calculate forces based on the angular size of the cube appearing from the position of the particle.





N-Body Code- Particle-Mesh

Calculating the actual forces is done in two main ways:

The second method is a particle-mesh (P-M) algorithm where a uniform cubic grid (mesh) is formed by adding all the particles in each grid cell. The Poisson equation is solved in Fourier space and the forces on each of the N particle-mesh cells can be solved in an algorithm where computational expense scales with N. The limitation is that forces below the grid-size cannot be calculated. A 3-dimensional grid can be memory intensive, hence the size of the grid is often limited by memory

In practice, the small-scale forces within a grid can be calculated directly (i.e. direct particle-particle, P-P, calculation) while particle-mesh works on larger scales- this is called the P^3M method.

In other codes, the tree algorithm is used for short-range forces and the particle-mesh algorithm is used on long-range forces. This allows the greatest amount of dynamic range in a large box (e.g. the Gadget-2 N-body code used to calculate the Millenium simulations).

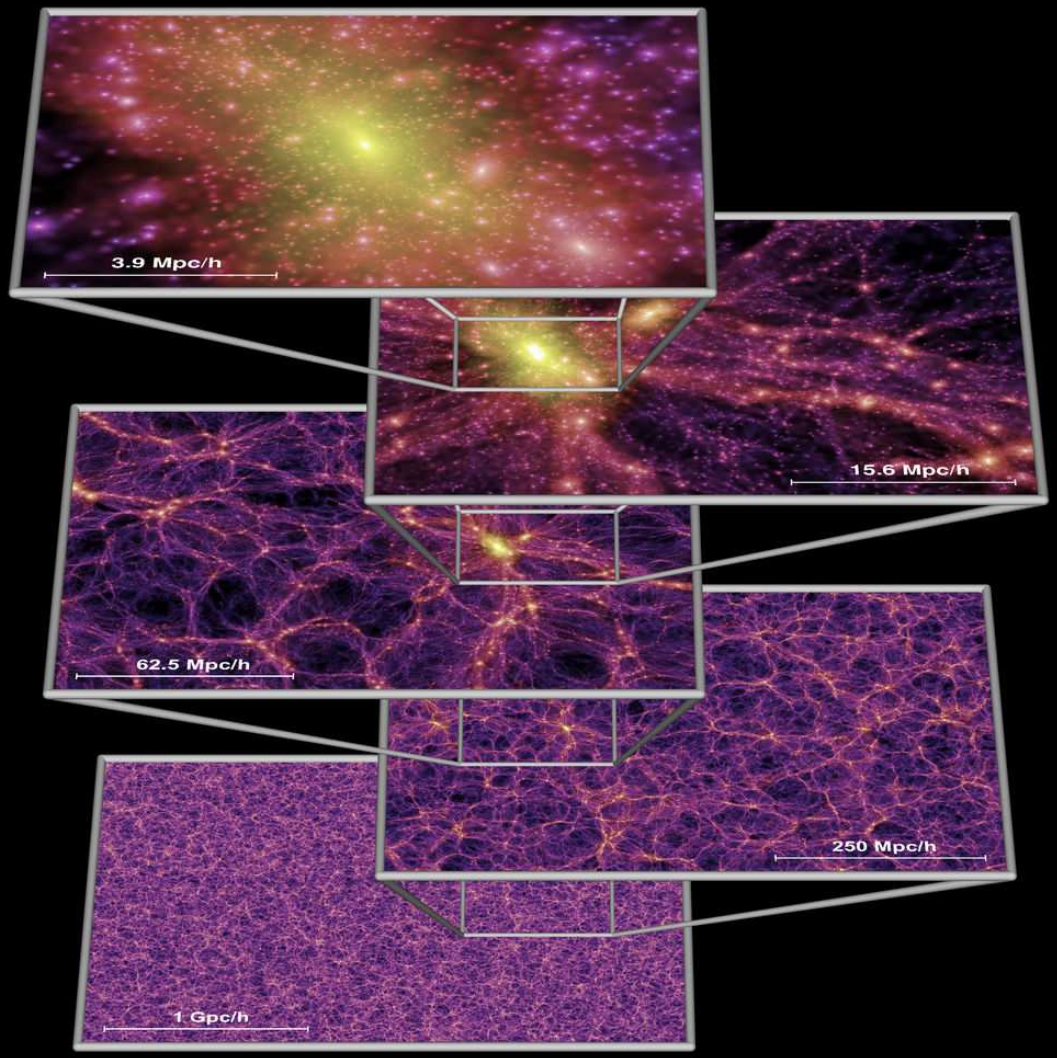
The Best-Known Collisionless DM-Only Simulation: The Millennium Simulation

The Millenium simulation is an example of a DM-only simulation with cosmological initial conditions (ICs) using the observationally constrained power spectrum.

The cosmic web on Gpc scales is resolved down to dark matter halos hosting galaxies smaller than the Milky Way.

- 2160^3 DM particles
- 8.6×10^{10} $M_\odot$ particle resolution
- 350,000 CPU hours across 512 cores in early 2000s

Simulations this size are commonplace today, but were a computational challenge just 10 years ago.



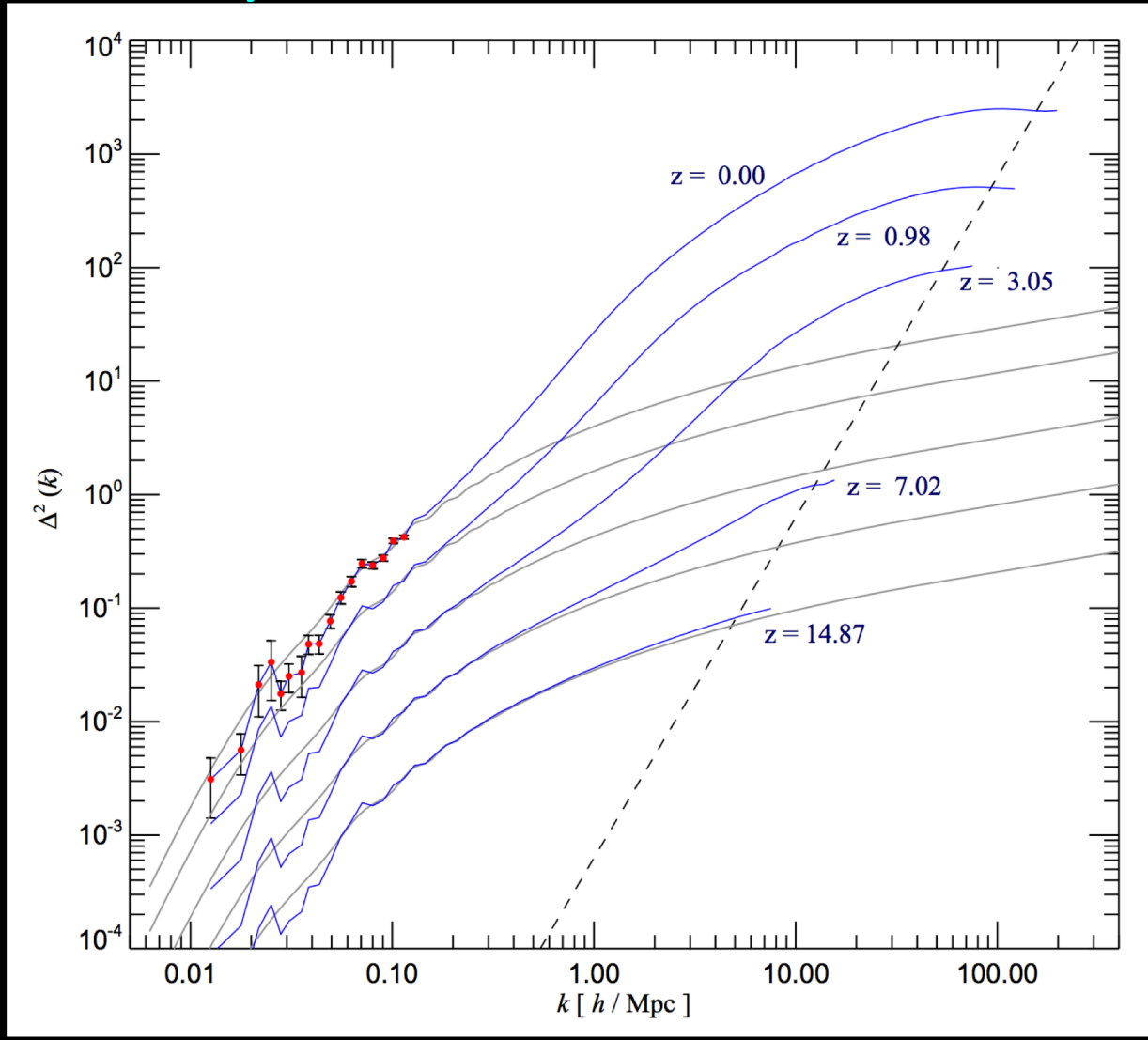
Springel (2005)- Zooms from the cosmic web to a cluster

Power Spectrum Evolution of Dark Matter of an N-body Code

The evolving power spectrum as a function of redshift in the Millennium Simulation.

Collapsing (sheets & filaments) and collapsed (halos) structures of course behave non-linearly (blue lines), and diverge from the linear spectrum (gray lines).

Error bars show cosmic variance, which has very small error bars beyond those that are plotted.



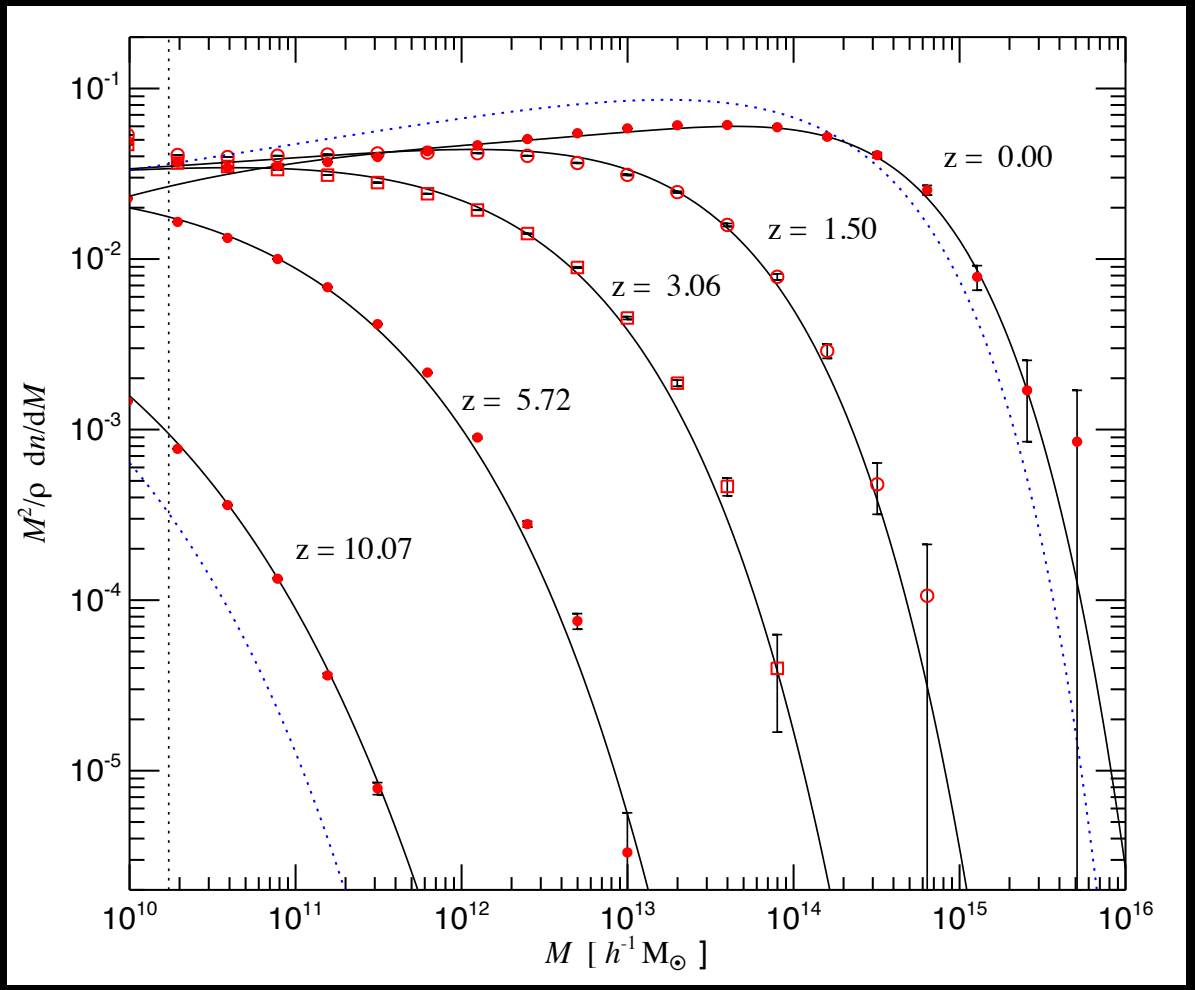
Springel (2005)- Power spectrum evolution.

The Best-Known DM-Only Simulation: The Millennium Simulation

The Millenium simulation is an example of a DM-only simulation with cosmological initial conditions (ICs) from the observationally constrained power spectrum.

The cosmic web on Gpc scales is resolved down to dark matter halos hosting $2 \times 10^{10}\ M_\odot$ (DM particle mass is $8.6 \times 10^{8}\ M_\odot$.

The halo mass function goes as $N_{halos} \sim M_H^{-0.9}$ over >4 dex in mass by z=0. Hence approximately the same amount of mass in each halo mass bin M_{halo} (as it is plotted here).



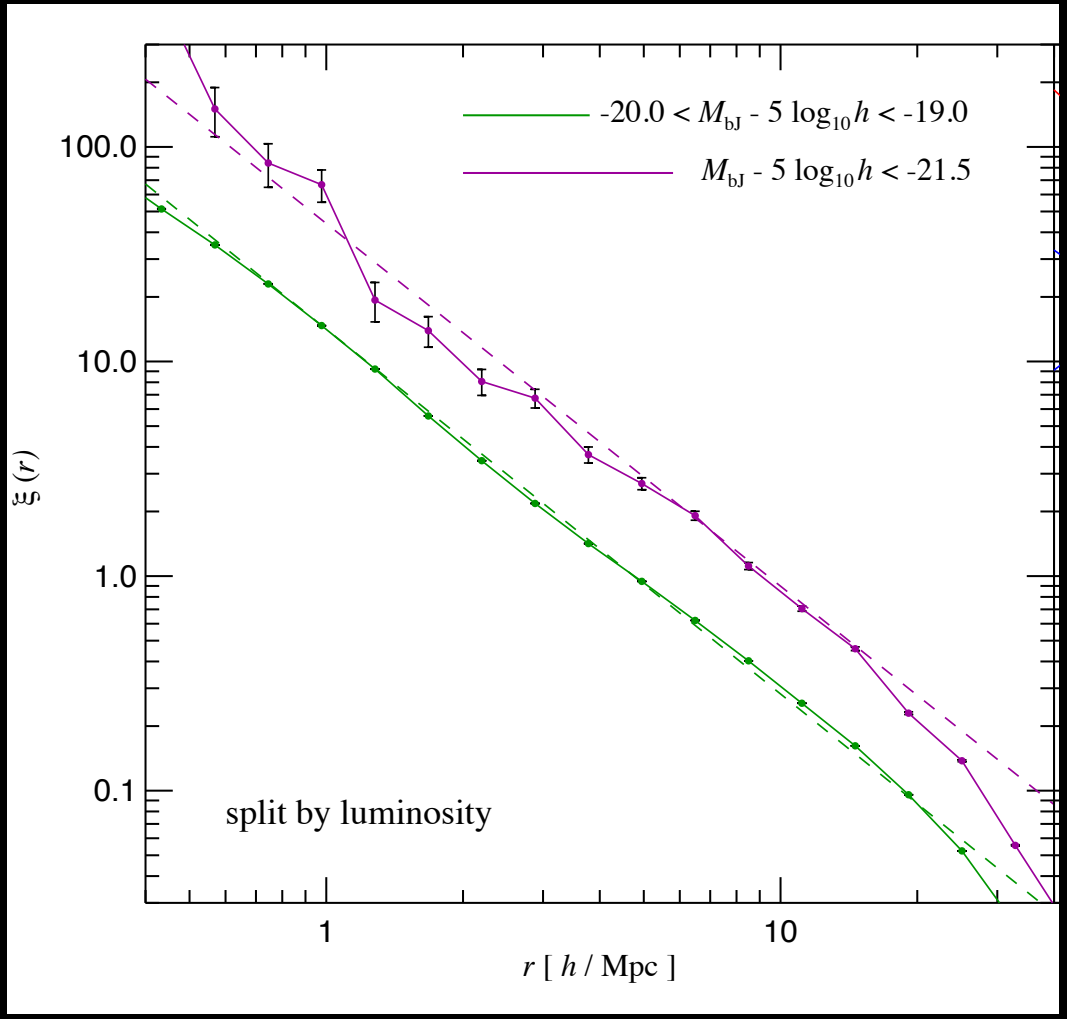
Springel (2005)- Dark matter halo function at various redshifts.

Observed Galaxy Power Spectrum Reproduced by Millennium Simulation

The Millennium Simulation was tested against the observed clustering of galaxies (in Sloan Digital Sky Survey, SDSS), and was found to reproduce the observed clustering of galaxies.

Reproducing the observed clustering of galaxies (more luminous/massive galaxies have larger clustering strengths) is a substantial success of ΛCDM cosmology.

These were some of the first tests of CDM cosmology in the 1980's (the CfA galaxy survey).



Springel (2005)- Galaxy clustering in Millennium Simulation

Dark Matter Halo Profiles

Navarro, Frenk, & White (1996, 1997) ran N-body DM-only simulations resolving the formation of DM halos. They identified a universal density profile that has 2 parameters, mass (M) and concentration (c). The profile is 2 power laws that transition at a scale radius

$$\frac{\rho(r)}{\rho_{crit}} = \frac{\delta_c}{(r/r_s)(1+r/r_s)^2},$$

The halo concentration is defined as $c = r_{200}/r_s$, and the critical density is defined as

$$\delta_c = \frac{200}{3}\frac{c^3}{[\ln(1+c) - c/(1+c)]}.$$

And r_{200} is the radius of the halo, defined as 200 times the critical overdensity of the Universe, ρ_{crit}. In terms of cosmological parameters and as a function of halo mass, r_{200} is

$$r_{200} = 1.63 \times 10^{-2} \left(\frac{M}{h^{-1}M_\odot}\right)^{1/3} \left(\frac{\Omega_0}{\Omega(z_0)}\right)^{-1/3} (1+z_0)^{-1}\, h^{-1}\mathrm{kpc},$$

Dark Matter Halo Profiles

The appendix in Navarro, Frenk, & White (1997) is a good reference for halo relationships, including the circular velocity as a function of radius.

A halo of mass M identified at $z = z_0$ can be characterized by its virial radius,

$$r_{200} = 1.63 \times 10^{-2} \left(\frac{M}{h^{-1} M_\odot} \right)^{1/3} \left(\frac{\Omega_0}{\Omega(z_0)} \right)^{-1/3} (1 + z_0)^{-1} \, h^{-1} \mathrm{kpc}, \qquad (A1)$$

or by its circular velocity,

$$V_{200} = \left(\frac{GM}{r_{200}} \right)^{1/2} = \left(\frac{r_{200}}{h^{-1} \mathrm{kpc}} \right) \left(\frac{\Omega_0}{\Omega(z_0)} \right)^{1/2} (1 + z_0)^{3/2} \, \mathrm{km/s}. \qquad (A2)$$

The density profile of this system is fully specified by its characteristic density δ_c and is given by (see eq.1)

$$\rho(r) = \frac{3H_0^2}{8\pi G} (1 + z_0)^3 \frac{\Omega_0}{\Omega(z_0)} \frac{\delta_c}{cx(1 + cx)^2}, \qquad (A3)$$

where $x = r/r_{200}$ and c is the concentration parameter, a function of δ_c given in eq. 2. The corresponding circular velocity profile, $V_c(r)$, is given by

$$\left(\frac{V_c(r)}{V_{200}} \right)^2 = \frac{1}{x} \frac{\ln(1 + cx) - (cx)/(1 + cx)}{\ln(1 + c) - c/(1 + c)}, \qquad (A4)$$

using the concentration c as a parameter.

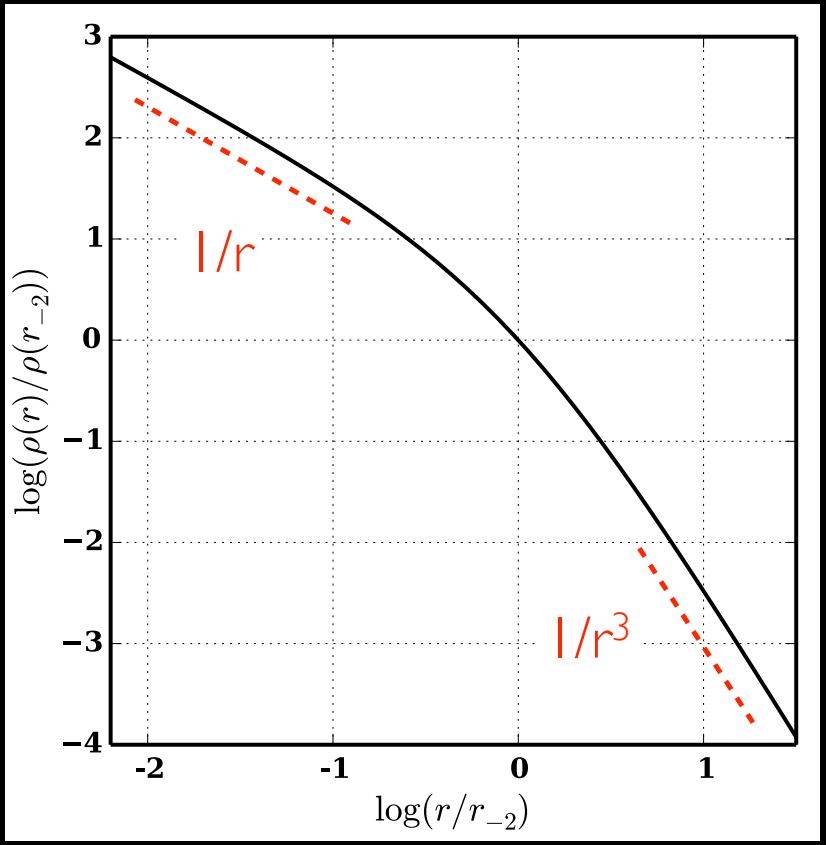
NFW found that lower mass halos collapse earlier and have higher concentrations.

NFW Profile- $1/r$ and $1/r^3$ split at $1/r^2$

$\rho(x) = \rho_0/(x(1+x^2))$

where $x = r/r_{-2}$

r_{-2} occurs where profile is at $1/r^2$ and is also referred to as the scale radius, r_s



Adapted from Mike Boylan-Kolchin slide

Dark Matter Halo Concentration

The Einasto (1965) profile has an extra free parameter beyond the NFW profile, and sometime provides a better fit to the halo profile.

The Einasto profile (Einasto 1965) has an extra free parameter, the shape parameter α, and may be written as

$$\ln\left(\frac{\rho_E}{\rho_{-2}}\right) = -\frac{2}{\alpha}\left[\left(\frac{r}{r_{-2}}\right)^{\alpha} - 1\right]. \quad (4)$$

The parameter r_{-2} marks the radius where the logarithmic slope of the density profile is equal to -2. The same property

For Einasto and NFW, r_s is sometimes called r_{-2}, which corresponds to the radius where $\rho \sim r^{-2}$ (the profile of an isothermal sphere) and the transition between the inner $\rho \sim r^{-1}$ and the outer $\rho \sim r^{-3}$.

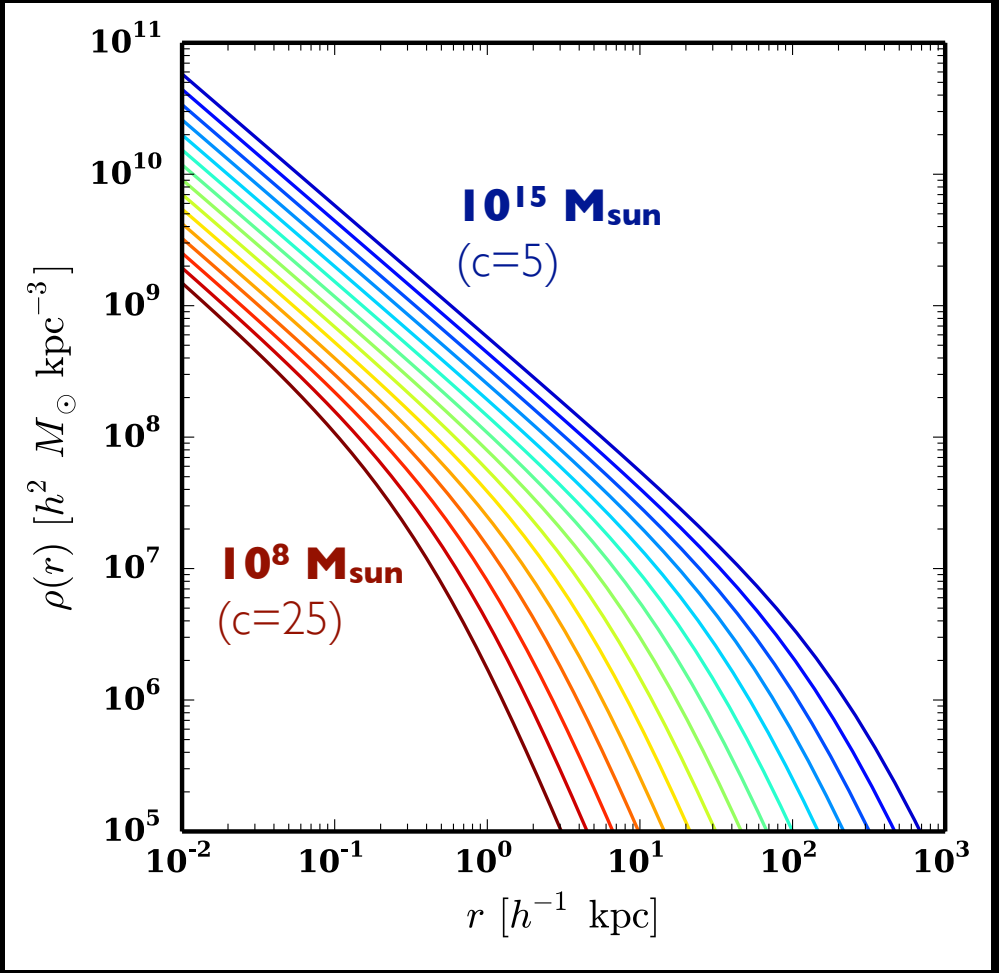
The concentration of a halo is the only free parameter of NFW profiles, which are more generally used, and understanding what concentration correlates with is an active area of research in DM-only simulations.

Dark Matter Halo Concentration

Bullock et al. (2001) showed halo concentration declined as halo mass went up.

$c = 10\ (M_H/10^{12})^{-0.1}$

Bigger cores for bigger halos!



Adapted from Mike Boylan-Kolchin slide to show Bullock+ (2001) trend.

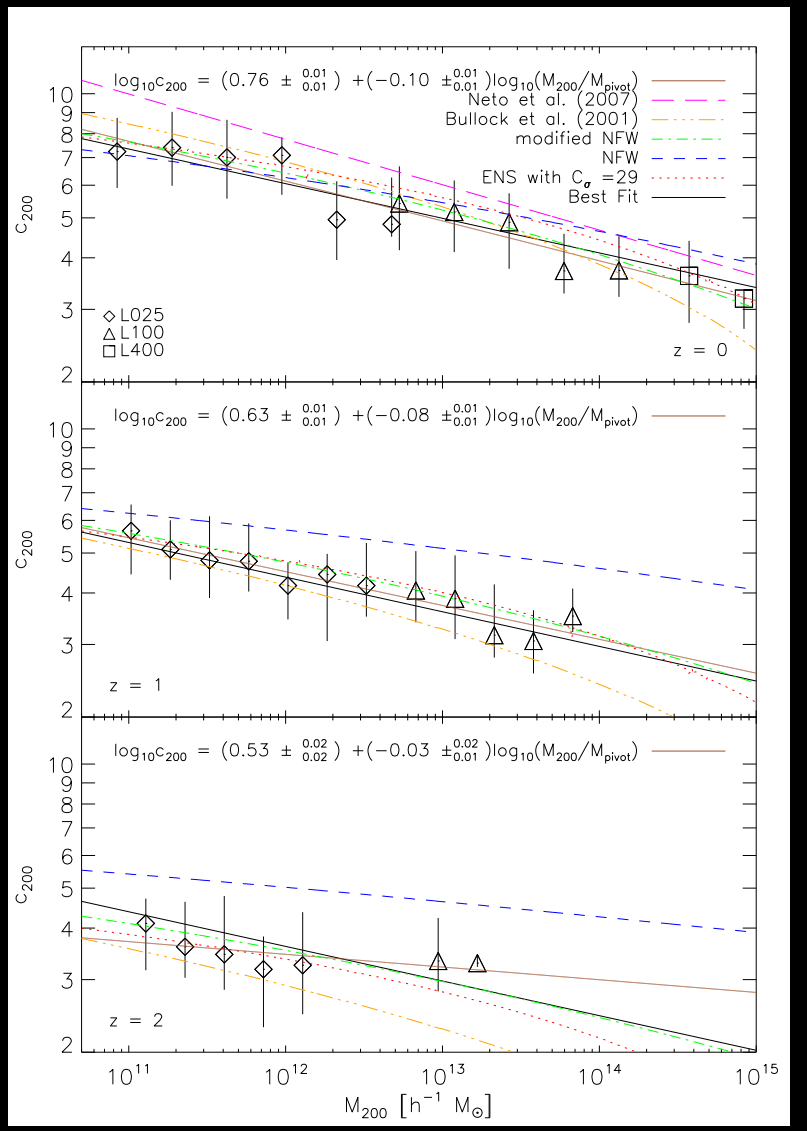
Dark Matter Halo Concentration Correlates with Mass and Redshift

Duffy et al. (2008) explored DM halo concentrations in the Millenium simulation and show the trend with mass (M_H) and redshift (z).

At a given redshift, concentration declines with halo mass.

Concentration declines with redshift (i.e. higher redshifts have lower c for a given M_H).

What is the reason for this trend?



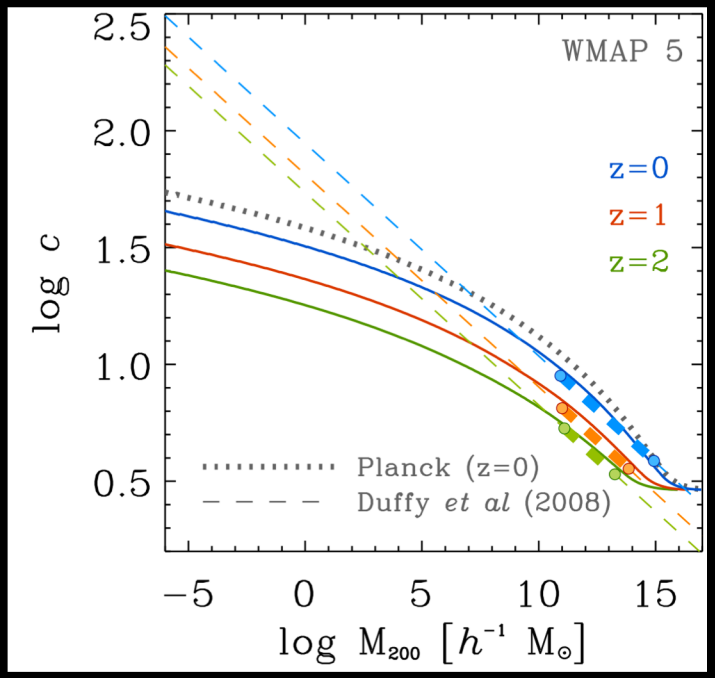
DM Halo Concentration Depends on Accretion History

Ludlow et al. (2013) showed that because lower mass halos assemble earlier, they have higher concentrations reflecting the critical density of the Universe at the time of collapse.

Mass accretion histories determine dark matter halo concentrations. A cluster-sized halo that assembles later will have a much lower concentration, while a low-mass halo will have a high concentration that forms early and a larger distance from r_s to r_{200} that has grown later at lower redshift.

Concentration increases toward lower redshift, which is mainly a result of "pseudo-growth" where the critical density of the Universe declines due to expansion leaving behind the concentrated core that evolves little once it formed.

Ludlow et al. (2014) showed how halo concentration extrapolates to lower halo masses differently than scaling the Duffy et al. (2008) relations to mini-halos (i.e. 10^6 $M_\odot$ and below).



“Zoom” Renormalization Simulations

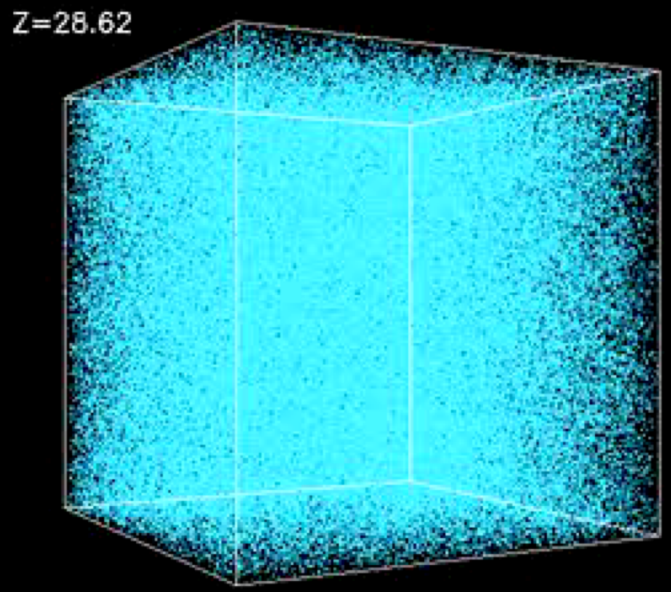
Cosmological boxes, while having cosmological structures based on ΛCDM cosmology, have the drawback of limited resolution when simulating an entire 3-dimensional periodic box.

Simulating a single galactic halo can follow particles only in that halo at a much higher resolution for the same computation time. However, the initial conditions of ΛCDM cosmology is missing, and the need to follow the power spectrum on small scales is not possible using a general halo profile as an IC.

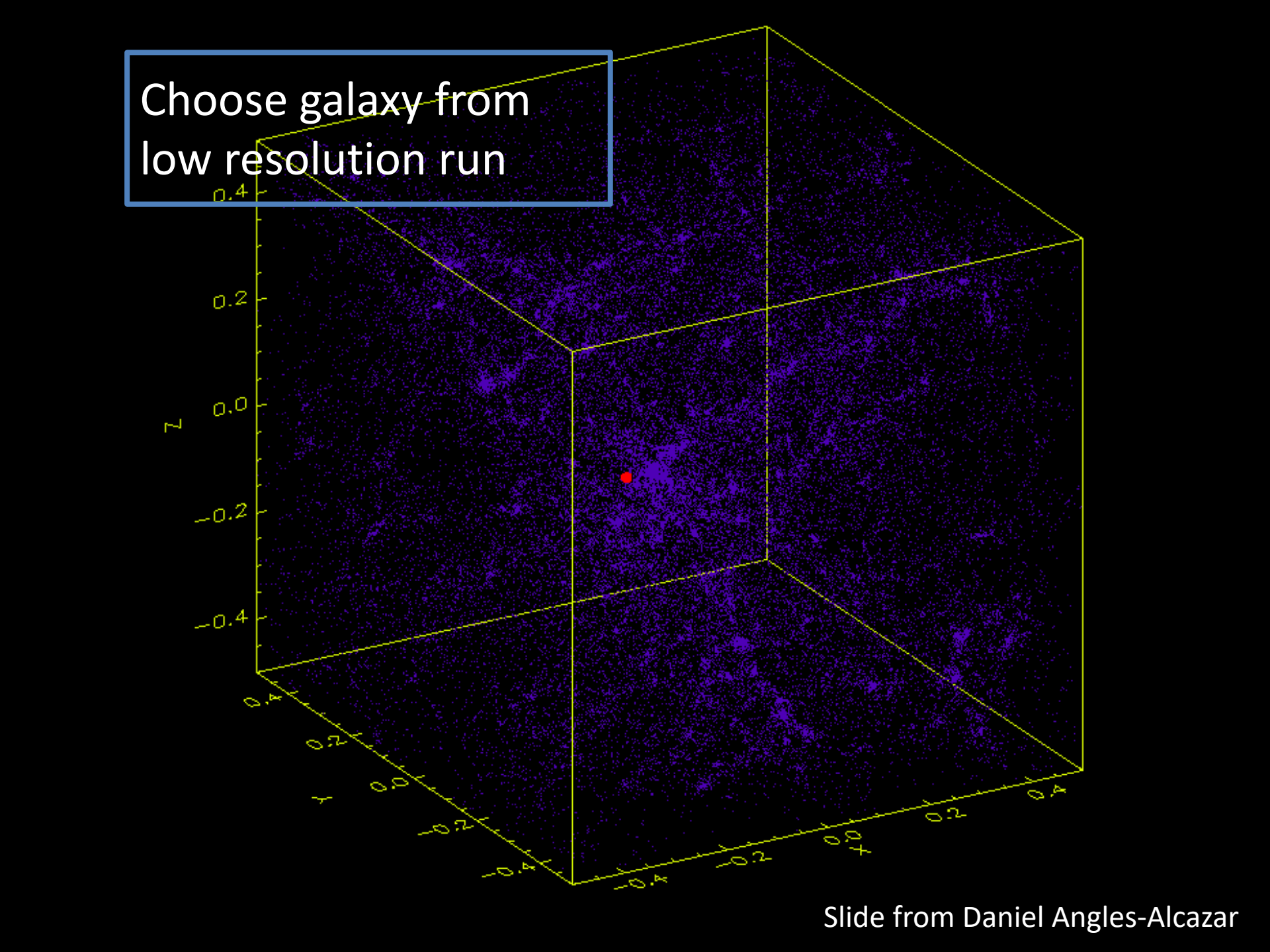
Renormalization simulations allow one to select the particles that form a halo in a later simulation using cosmological initial conditions, but also adding the small-scale power spectum fluctuations that will evolve into the subhalos and substructure in the parent halo. Now you get both:

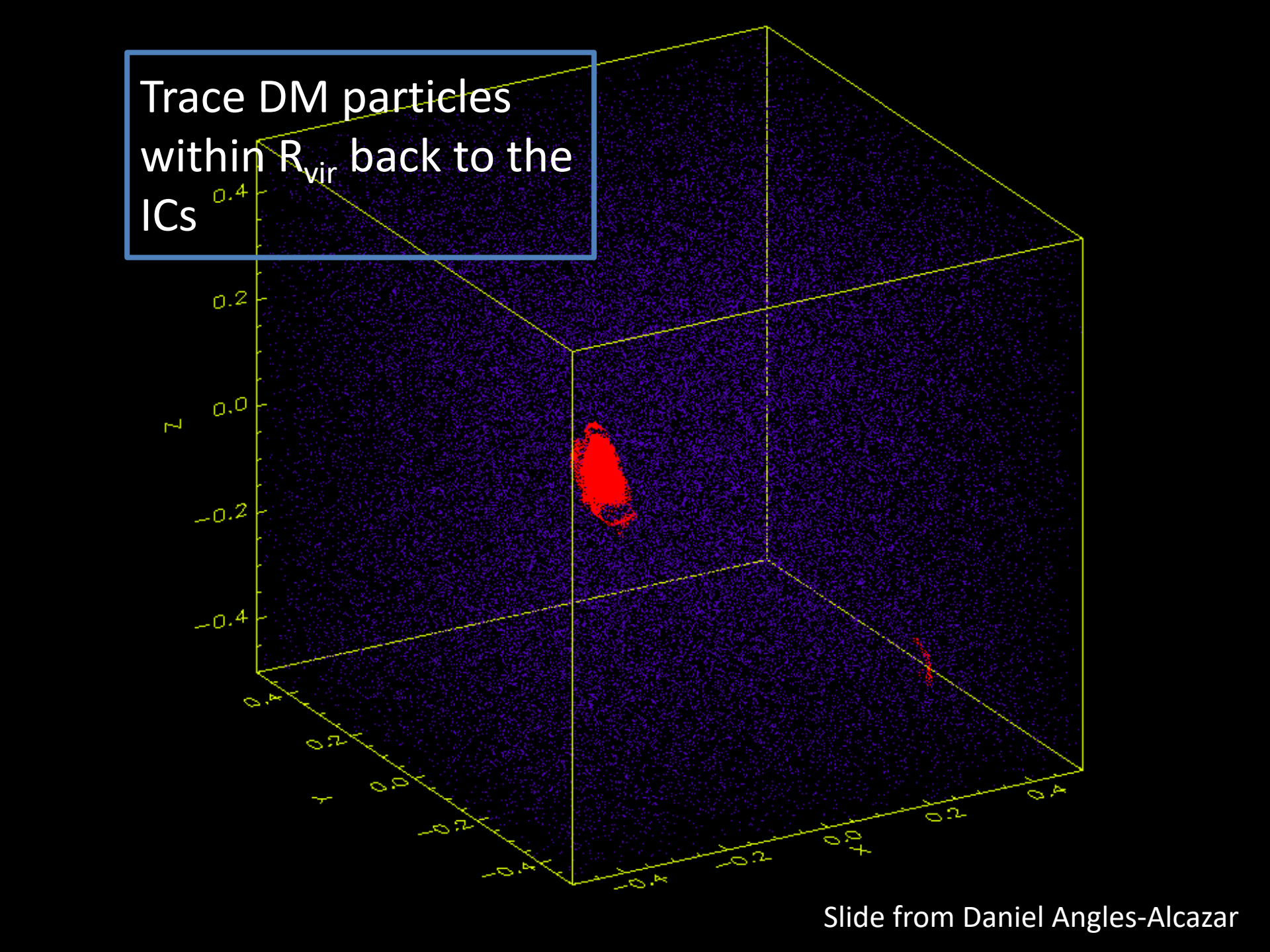
-Higher resolution concentrating on a single halo where you want it.

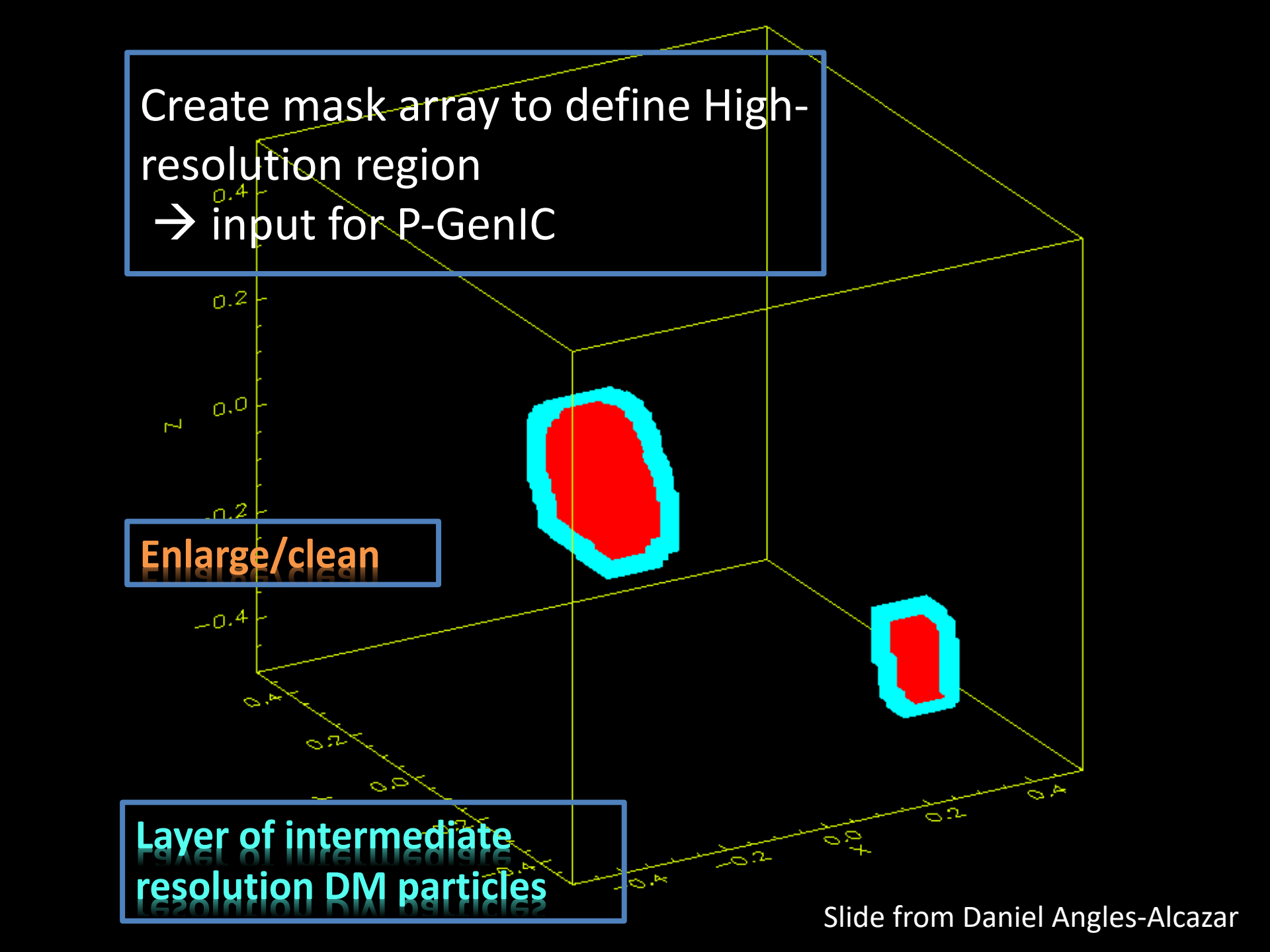
-Cosmologically-based initial conditions resolving small-scale structure- important for resolving sub-halos in a larger halo.

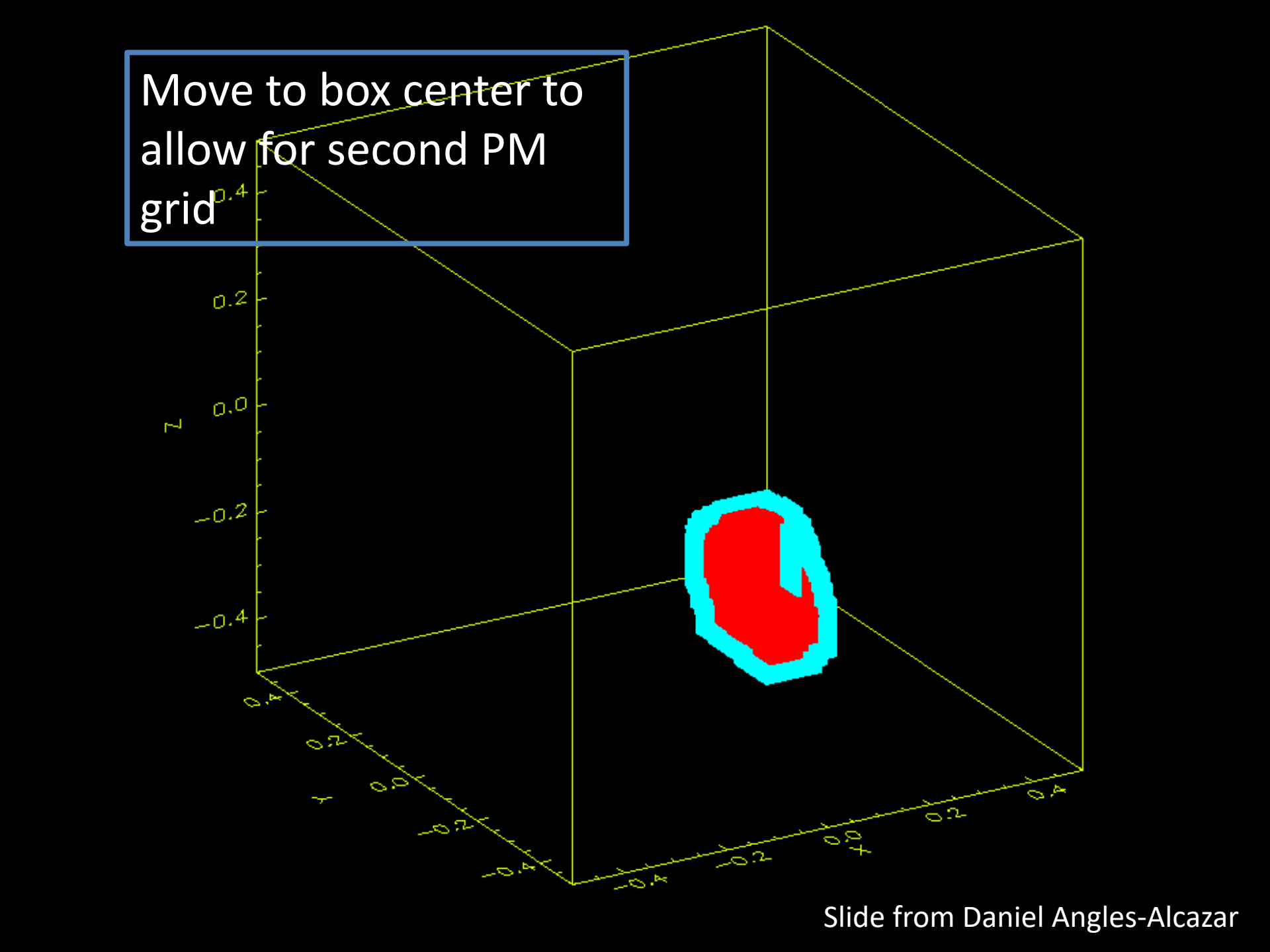


Kravtsov DM-only simulation movie.


Choose galaxy from
low resolution run
0.4
0.2
0.0
−0.2
−0.4
Z
Y
X
Slide from Daniel Angles-Alcazar


Trace DM particles within R_{vir} back to the ICs
0.4
0.2
0.0
−0.2
−0.4
Z
Y
X
Slide from Daniel Angles-Alcazar


Create mask array to define High-resolution region
→ input for P-GenIC
Enlarge/clean
Layer of intermediate resolution DM particles
Slide from Daniel Angles-Alcazar



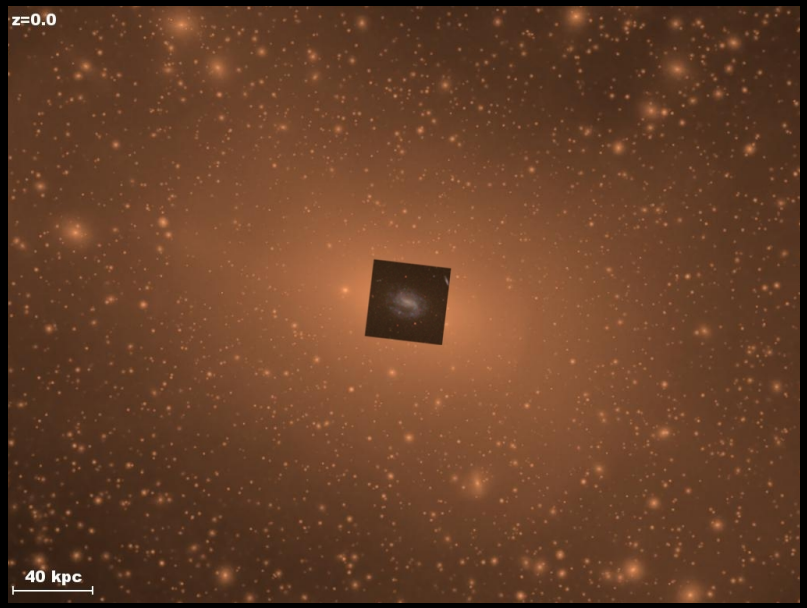
Slide from Daniel Angles-Alcazar

Zoom Simulations Allow the Simulation of a Milky Way-Mass Halo at High Resolution

Several large simulations of Milky Way mass halos reveal the substructure of dark matter halos, which are expected to be many in the cold dark matter model.

The simulations are very expensive to run, because they have up to 10^9 dark matter particles that have to be followed on very small timesteps- softening lengths are <100 pcs.

The Aquarius simulations have DM particle masses below 10^4 M_{sol}, which means that subhalos with masses ~10^6 M_{sol} can be resolved (need at several dozen to a hundred DM particles to resolve a halo) in a 10^{12} M_{sol} halos capable of hosting the Milky Way (MW). Others would argue more to really resolve the dynamics.



Diemand+ (2007)- Via Lactae simulation- another super-high resolution MW halo simulation.

The Effect of Baryons on Dark Matter Halos

Much work goes into understanding dark matter halos in DM-only simulations. Following the formation of mini-halos in a cosmological simulation is computationally expensive, because the softening lengths are small meaning the timesteps are small and many more are taken.

Adding baryons to simulations increases the cost of simulations even more due to hydrodynamics, cooling/photo-heating, and star formation. When understanding dark matter, baryons are critical to follow for two main reasons:

1. They are what we observe, hence they are observational traces of dark matter. Until we can clearly directly detect dark matter, we are relegated to baryons.

2. Baryons cause deviations from expected halo profiles from DM-only simulations. Despite being 16% of the total mass (Ω_{baryon}~0.046, Ω_{DM}~0.23), baryons dominate the mass in galaxies. Hence they change the dark matter profiles.

Hydrodynamics Following Baryons

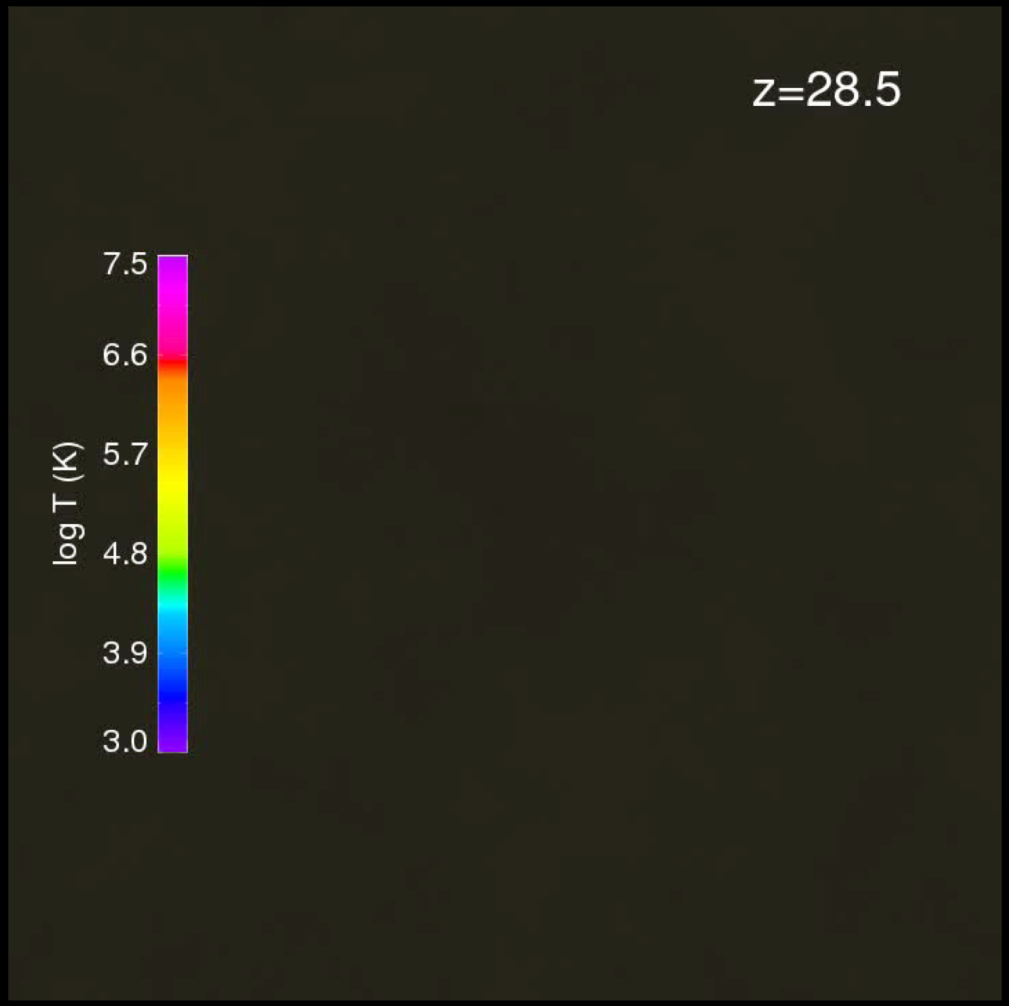
About $1/6^{th}$ of mass is baryons (i.e. normal matter– protons, electrons, neutrons) according to the favored ΛCDM cosmology. $\Omega_{baryon}=0.046$, $\Omega_{matter}=0.28$, $\Omega_{darkmatter} = 0.28-0.046 = 0.234$.

We will also see that despite being $1/6^{th}$ the mass, baryons dominate the mass in regions where probes of dark matter are important. The Milky Way galaxy inside the solar radius (~8 kpcs) is dominated by baryonic mass (although it is close).

The initial conditions of baryons before galaxy formation is gas. Gas is different than dark matter in several ways:

- Gas does not interpenetrate like dark matter, which is treated as collisionless. Instead it shocks turning kinetic energy into thermal energy, leading to dissipation.
- Gas equations of motion include pressure gradients of the gas.
- Gas dissipates energy by cooling and can also be heated by other sources like radiation, i.e. photo-ionization.
- Gas can be converted into stars via star-formation. Stars are collisionless, much like dark matter, and treated the same way in a simulation as gravitationally interacting particles.

A Hydro Simulation



Gadget-2 simulation used in Oppenheimer+ (2010)- Color is gas temperature.
384^3 dark matter and hydrodynamically interacting gas particles

Equation of Hydrodynamics

Cosmological hydrodynamic simulations have an additional layer of gas on top of dark matter. Gas is treated as an idea fluid where the continuity, momentum, and energy conservation equations- referred to as the Euler equation- written in physical coordinates (r) are solved:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = 0, \qquad \text{Mass conservation}$$

$$\frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v} \cdot \nabla)\,\mathbf{v} = -\left(\nabla \Phi + \frac{\nabla P}{\rho}\right), \qquad \text{Momentum conservation}$$

$$\frac{\partial}{\partial t}\left[\rho\left(\frac{v^2}{2} + \mathscr{E}\right)\right] + \nabla \cdot \left[\rho\left(\frac{v^2}{2} + \frac{P}{\rho} + \mathscr{E}\right)\mathbf{v}\right] - \rho \mathbf{v} \cdot \nabla \Phi = \mathscr{H} - \mathscr{C}. \qquad \text{Energy conservation}$$

where ρ is density, v is velocity, *P* is pressure, $\mathscr{E}$ is specific internal energy of the fluid, and $\mathscr{H}$ and $\mathscr{C}$ are heating and cooling, which we will consider later. An ideal gas with an adiabatic index γ defines the the equation of state $P = \rho(\gamma - 1)\mathscr{E}$

In addition the gravitational potential satisfies the Poisson equation $\nabla^2 \Phi = 4\pi G \rho_{\rm tot},$

Where ρ_{tot} includes baryonic and DM matter density.

Smooth Particle Hydrodynamic Codes

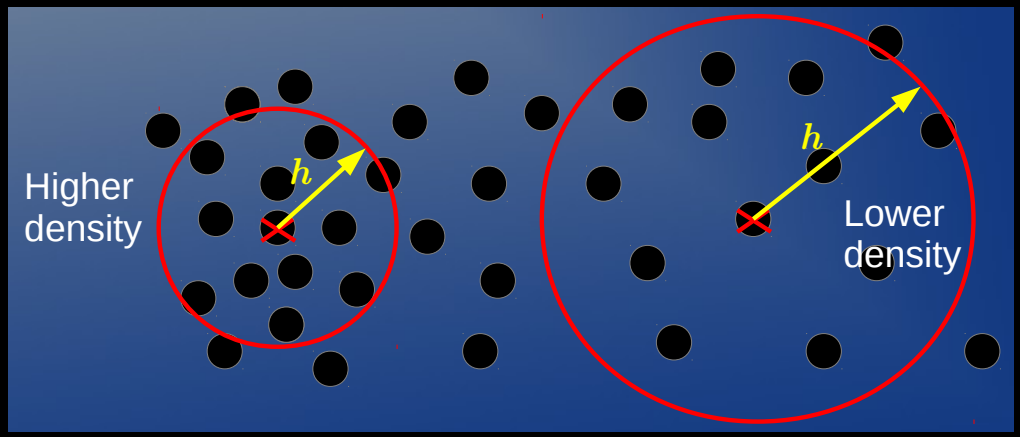
Smooth particle hydrodynamics (SPH) is a particle-based method for solving hydro, meaning that individual gas particles are followed. It is *Lagrangian* in nature meaning that these gas mass elements can move and have a velocity.

A smooth particle has a smoothing length, *h*, over which it interacts with its neighboring smooth particles hydrodynamically. A smoothing kernel is defined such that its density falls off according to a kernel, W(r,*h*), where R=r/*h*.

$$W(r,h) = \frac{1}{\pi h^3} \begin{cases} 1 - 3R/2 + 3R^3/4 & (0 \le R \le 1) \\ (2-R)^3/4 & (1 < R \le 2) \\ 0 & (\text{otherwise}), \end{cases}$$

A key advantage of SPH is its ability to resolve high density regions better than a mesh code with a single grid size.

Although these are discrete particles, density and other relevant physical quantities can be determined at any point by integrating over the smoothing kernels of all overlapping SPH particles.

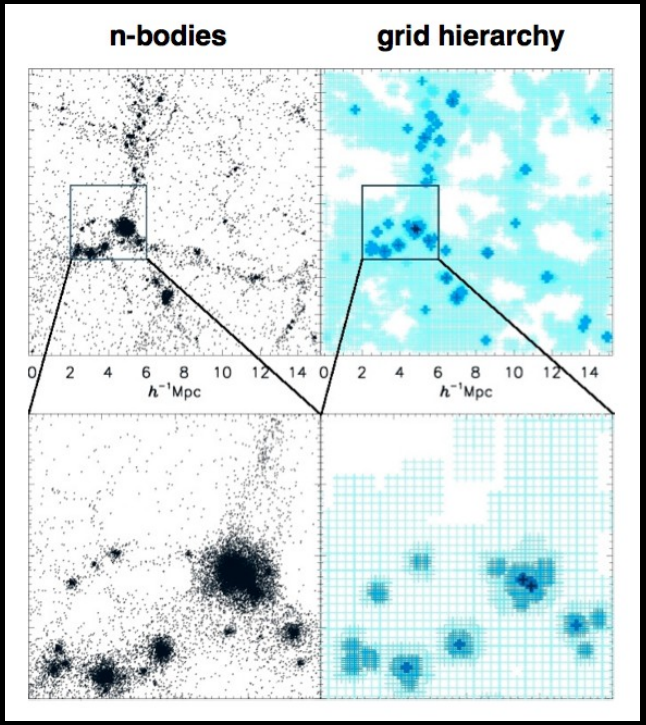


Grid-based Hydrodynamic Codes

Grid-based mesh codes divide gas into cells. This is an Euler method where cells are static in space and fluids flow between cells. The conservation equations of hydrodynamics are integrated across grid cell boundaries. A problem can arise if there are significant discontinuities across a boundary, however *adaptive mesh refinement* (AMR) can be used to subdivide grid cells into sub-grid cells at a refinement level *l,* thus splitting the spatial dimensions of the cell into $1/2^l$ cells. Hence, the dynamic range of an adaptive mesh refinement simulation can be greatly increased from a uni-grid.

The adaptive mesh refinement criterion is based on what the code specifies. It can be based on having a mass limit per cell element, but often is based on resolving shocks with sufficient resolution elements.

The levels of refinement in an AMR can effectively be infinite, but is limited by computational speed (i.e. smaller cells require smaller time-stepping), and available memory (one cannot save an infinite number of cells). Nevertheless, the number of levels of refinement can be great, $l > 10$, and even more.

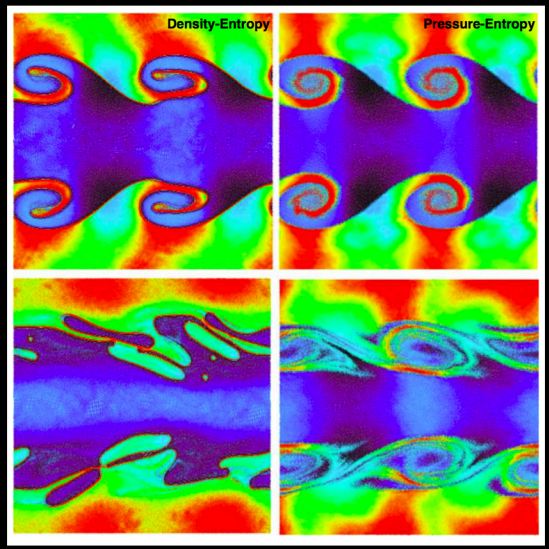


SPH versus AMR Simulations

Agertz+ (2007)- Normal SPH (top 2) cannot form Kelvin-Helmholtz instabilities that disrupt gas. AMR shreds this gas cloud as expected.

***AMR also has some key advantages over SPH simulations*:** 1) The key hydrodynamic test of forming Kelvin-Helmholtz instabilities fails in normal SPH. This leads to too much gas clumping in dense gas propagating through low-density background. Entropy profiles in clusters are wrong in SPH (too much cold gas sinks to center). 2) SPH cannot resolve shock boundaries like AMR, and shock capturing can suffer, especially if there is in-shock cooling (which should not occur). 3) The discrete nature of SPH particles means that artificial viscosity needs to be added, that mixing of fluid elements is absent, and the smooth length limits the spatial resolution and creates artifacts.

***SPH has key advantages over AMR*:** 1) Lagrangian nature of SPH allows one to follow fluid elements, and there is no preferred directions due to a grid. 2) Large density contrasts are resolvable in SPH due to particles following where the mass is, whereas AMR requires many-level refinement. 3) Conservation of angular momentum done better in SPH. 4) Short-comings of SPH for Kelvin-Helmholtz and artificial viscosity can be solved by adjusting the parameters and nature of SPH.



Hopkins (2013)- modified SPH right) can handle Kelvin-Helmhotlz instabilities

Hydro vs. DM-only Simulations

Hydro simulations are much more expensive than dark matter-only simulations, because of the need to track a layer of baryons. They also require far more additional coding. One needs to track the hydrodynamics of this set of these baryons in addition to their gravitational forces. Cooling and star-formation can be significantly expensive as well, and something that I will not talk about is radiative transfer, the propagation of light, which can easily dominate computational expense over the N-body and hydro portions of the code.

SPH simulation size is often quoted as $2x512^3$ where there are 512^3 DM particles and 512^3 baryonic particles. Up to a few years ago, this was considered a state-of-the-art size, but now some simulations are exceeding $2x1000^3$ and take 10's of millions of CPU hours using 10's of thousands of nodes.

State-of-the-art grid-based codes use adaptive mesh refinement, with state-of-the art simulations using many levels of refinement (>10) from starting with a base grid that is of significant size, e.g. 256^3.

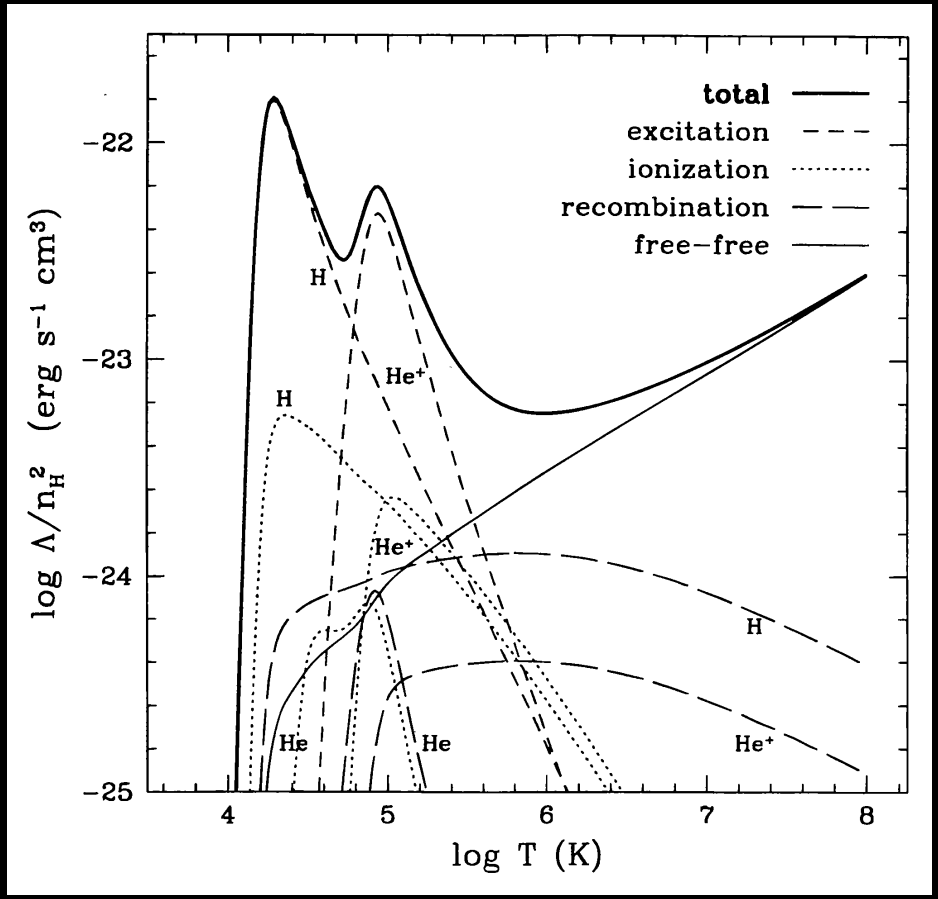
Compare this to the largest DM-only simulations, e.g. the Millennium-XXL with 6720^3, and you can see that hydro simulations are several orders of magnitude smaller in the number of resolution elements they follow.

Cooling and Photo-heating

Baryons radiate energy away by gas cooling. Hence, cooling curves that depend on the atomic physics of electronic interactions.

Gas in the initial conditions of a simulation have a primordial composition of H and He, which agree with observations and Big Bang nucleosynthesis models.

These models have mass fractions of hydrogen (X~0.75) and helium (Y~0.25).

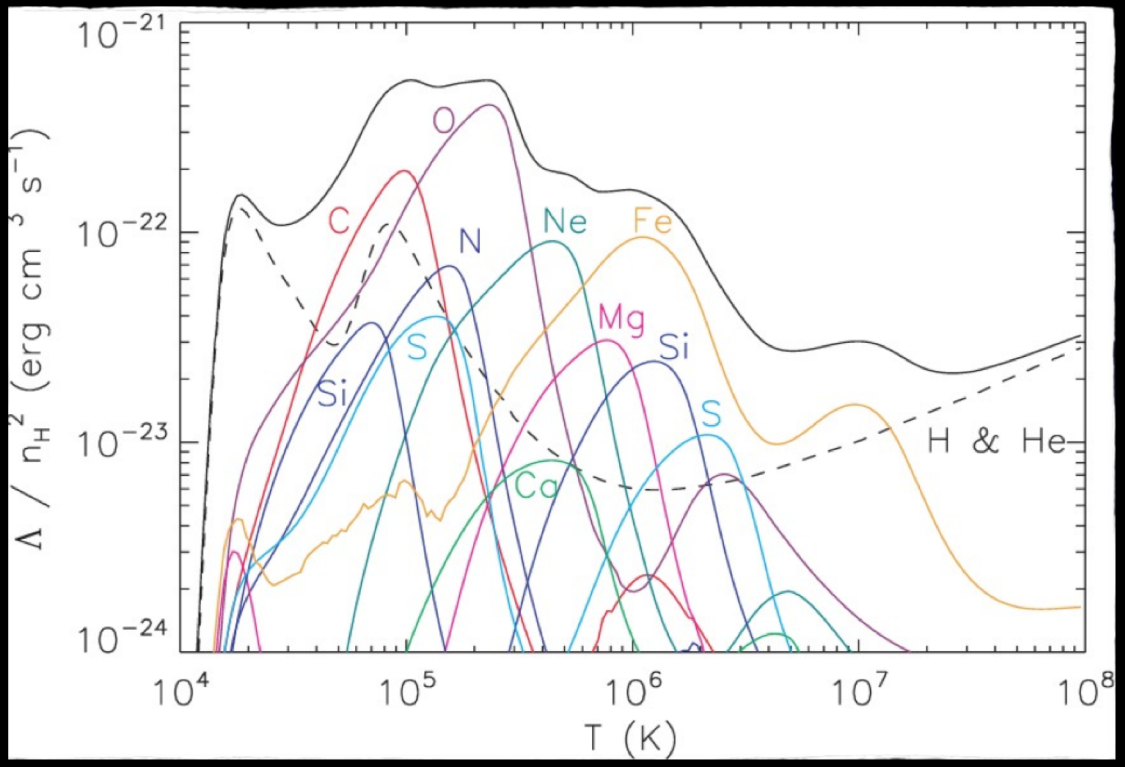


Katz, Weinberg, & Hernquist (1996)

Cooling and Photo-heating

As gas becomes enriched by the nucleosynthetic products of star formation, heavy element (metal) cooling becomes important and dominant at some temperatures (10^5-10^6 K) for gas enriched to solar metallicity (Z~0.012-0.018).

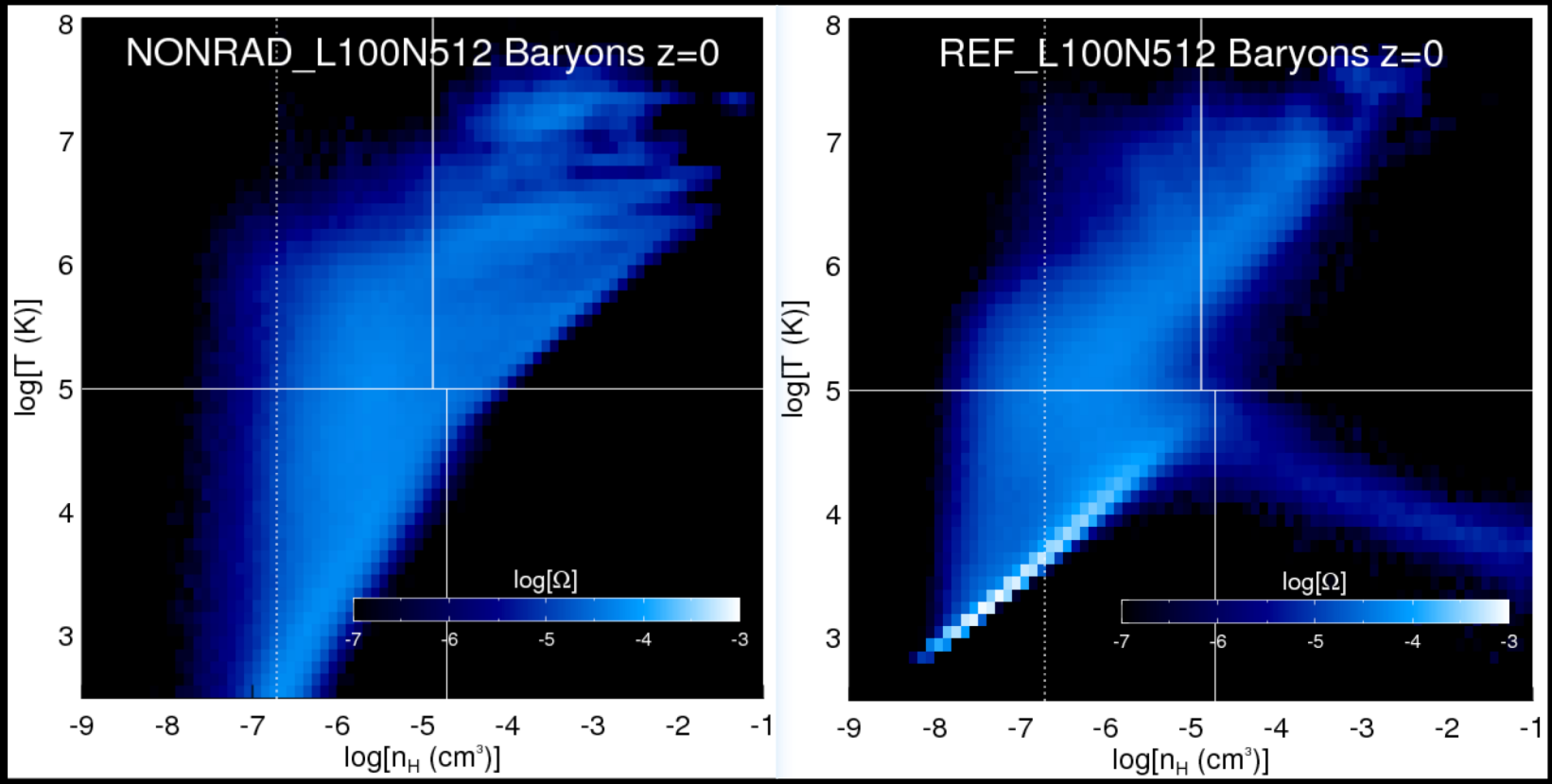
Following the details of gas cooling is critical for galaxy formation, because gas needs to be able to cool into galactic halos to achieve conditions of star formation.



Wiersma, Schaye, & Smith (2009)

Cooling and Photo-heating

The OWLS simulations (PI Schaye) ran a simulation without cooling or photo-heating and gas never cools to densities where star formation can occur.



An adiabatic simulation on the left versus a simulation with cooling + photo-heating on the right. OWLS simulations (PI, J. Schaye).

The Effect of Photo-Heating

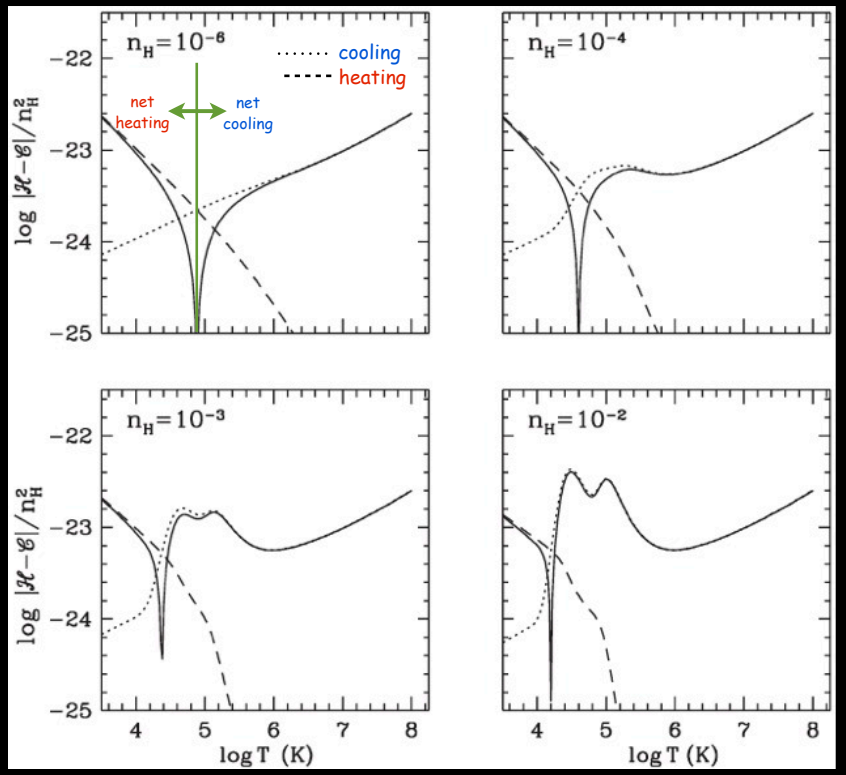
Photo-heating by the meta-galactic ionization background, whose source is UV and X-ray radiation emitted by quasars and star-forming galaxies is a significant source of heating the intergalactic medium (IGM).

Reionization photo-heats all gas (via photo-electric effect) to 10^4 K by z=6, which is normally done in simulations by simply turning on a field J_ν.

$$\epsilon_{x_i} = \int_{\nu_{0,x_i}}^{\infty} \frac{4\pi J_\nu}{h\nu} \sigma_{x_i}(\nu) h(\nu - \nu_{0,x_i}) d\nu,$$

Photo-heating, ε, balances cooling, Λ, to make a next cooling function, Λ_{net}.

$$\Lambda_{\mathrm{net},x_i}(T, z, n_{x_i}, n_e) = \Lambda'_{x_i}(T) n_{x_i} n_e - \epsilon_{x_i}(z) n_{x_i}$$



Adapted from F. van den Bosch notes.

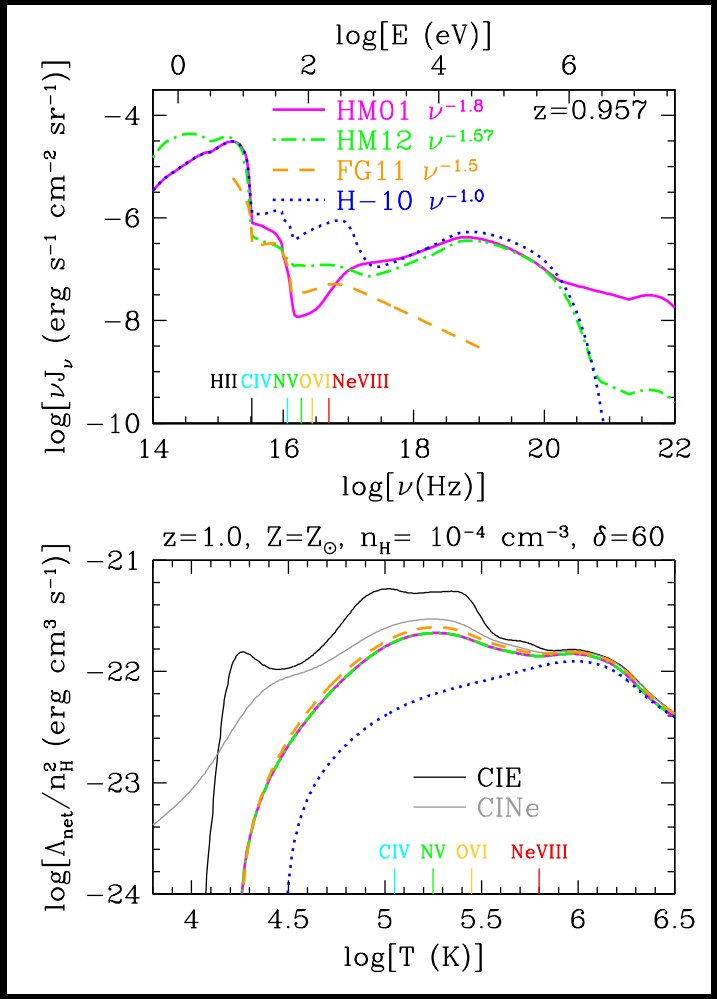
Photo-heating from a uniform background is a heating component that scales with density, while cooling scales with the square of density- creates a density dependence where photo-heating dominates at lower density over cooling (e.g. the IGM).

The Effect of Photo-Heating

The specific shape of the ionization background shapes the net cooling curve, which is critical (as we will see later) for how galaxies form.

Some of my work (Oppenheimer & Schaye 2013a) has been to calculate the net cooling of metal-enriched gases that are out-of-equilibrium (following the individual metal ions and their recombination and ionization timescales).

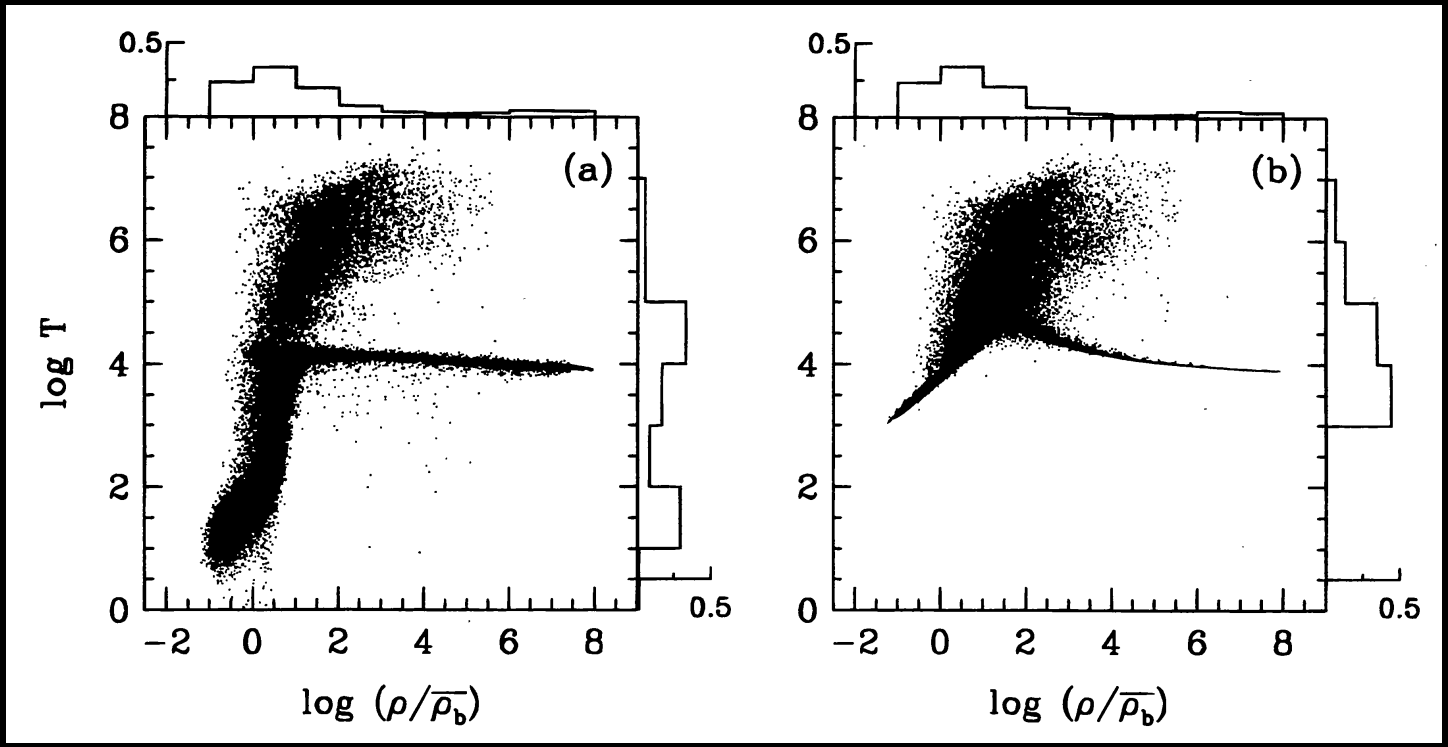
It turns out that non-equilibrium cooling is not very important for the accretion efficiency onto galaxies.



Oppenheimer & Schaye (2013)- photo-heating depends on the shape of the ionization background and non-equilibrium ionization/cooling effects.

The Effect of Photo-Heating

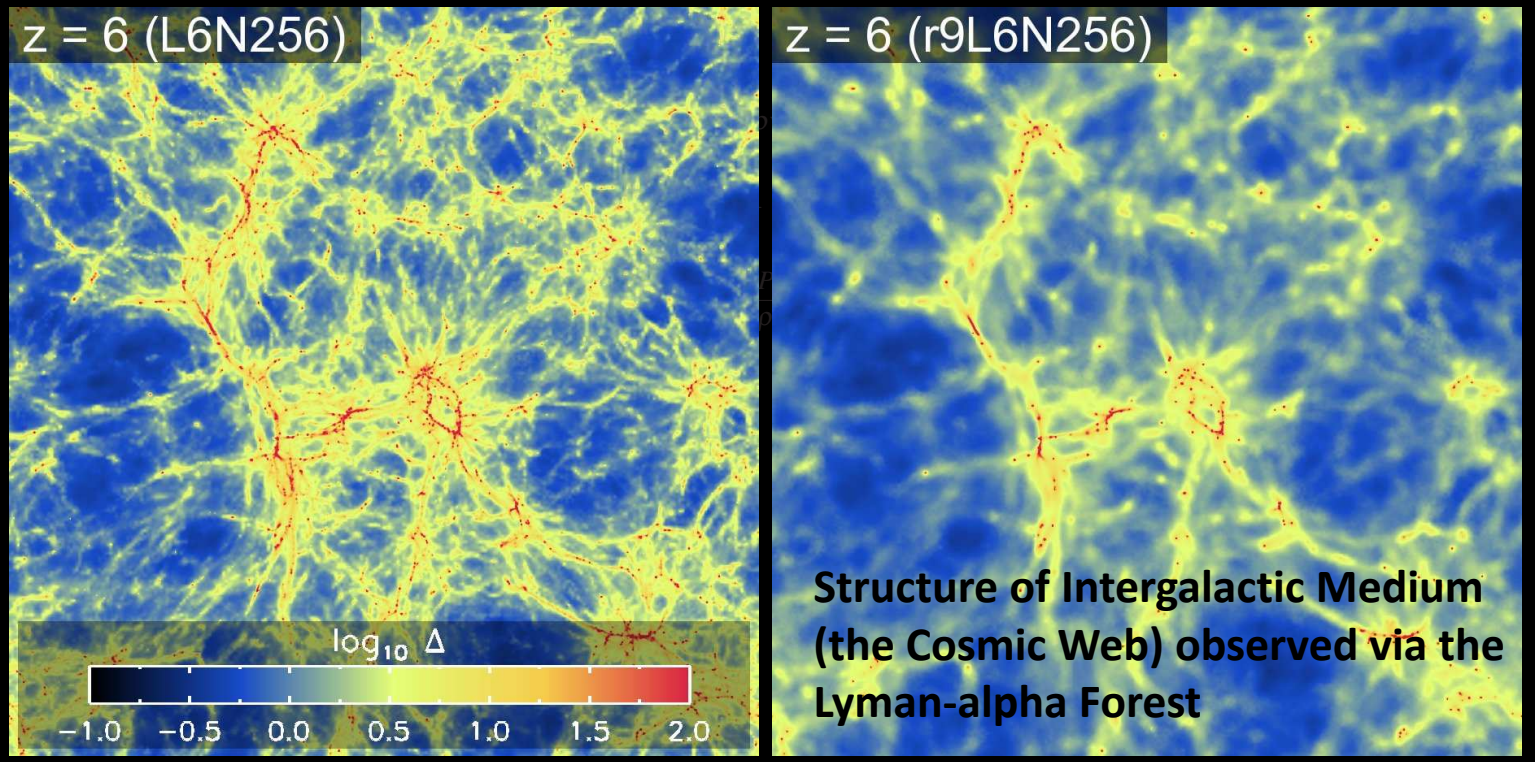
Photo-heating is important for heating the diffuse IGM, as shown by Katz+ (1996) where they include cooling but not photo-heating.



Katz, Weinberg, & Hernquist (1996)- distribution of z=2 gas particles in case without photo-heating (left) and with photo-heating (right).

The Effect of Photo-Heating

Photo-heating smooths out baryons. This has major implications for how easy it is to keep the IGM ionized as well as for the ability for baryons to cool into DM halos.



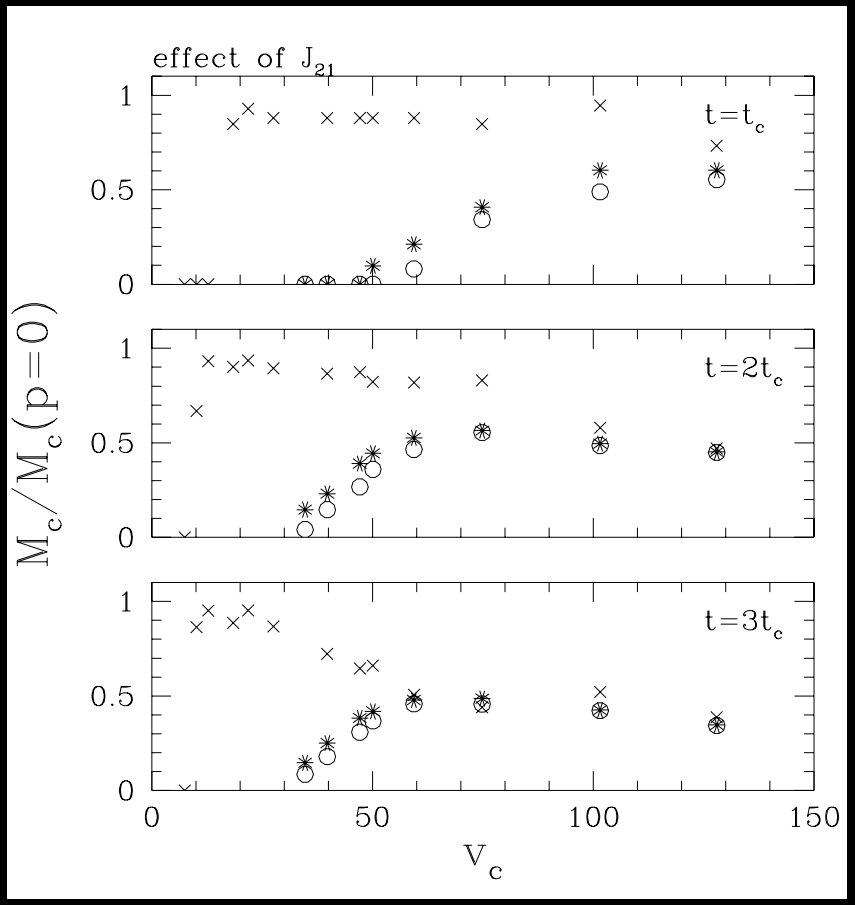
Pawlik, Schaye, & Scherpenzeel (2009)- density distribution of gas at z=6 without and with photo-heating through the process of reionization.

The Effect of Photo-Heating

Photo-heating "boils" baryons out of low-mass halos, where gas photo-heated to ~10^4 K cannot collapse owing to the Jeans criterion of collapse.

Thoul & Weinberg explored this and found that halo below v_{circ}= 30-50 km/s cannot hold onto their baryons.

Mini-halos forming the first stars accrete gas before reionization, so these early halos should be able to host the first stars and galaxies below this v_{circ}~ 30km/s, and agrees with Milky Way dwarf spheroidals that occupy lower mass halos- formed before reionization proceeded.



Thoul & Weinberg (1996)- x's show mass of gas without photo-ionization and o's and *'s show with photo-ionization.

Star Formation

The original star formation criterion in hydrodynamic simulations from the 1990's required gas to be in a convergent flow (div v < 0) and that the gas is Jean's unstable ($t_{sound\text{-}crossing} < t_{dyn}$), according on Katz (1992).

$$\frac{h_i}{c_i} < \frac{1}{\sqrt{4\pi G \rho_i}}$$

A Schmidt law is found to apply when the star formation criterion is achieved such that $\rho_{SF} \sim \rho_{gas}^{\alpha}$ and $\alpha=3/2$ (Silk, 1987). It is implemented in a simulation using variables c_* (a dimensionless parameter, set to 0.1) and t_g, which is the maximum of the gasdynamical time and cooling time.

$$\frac{d \ln \rho_g}{dt} = -\frac{c_*}{t_g},$$

Because of the limited simulation resolution, individual gas particles or cells are treated as individual molecular clouds and discrete star particles form with masses of many times the mass of actual stars ($>10^6$ M_{sol} in a cosmological simulation and $<10^4$ M_{sol} in zoom simulations. In practice a gas particle has a probability of "spawning" a star particle each timestep Δt.

$$p = 1 - e^{-c_* \Delta t / t_g}$$

Star Formation

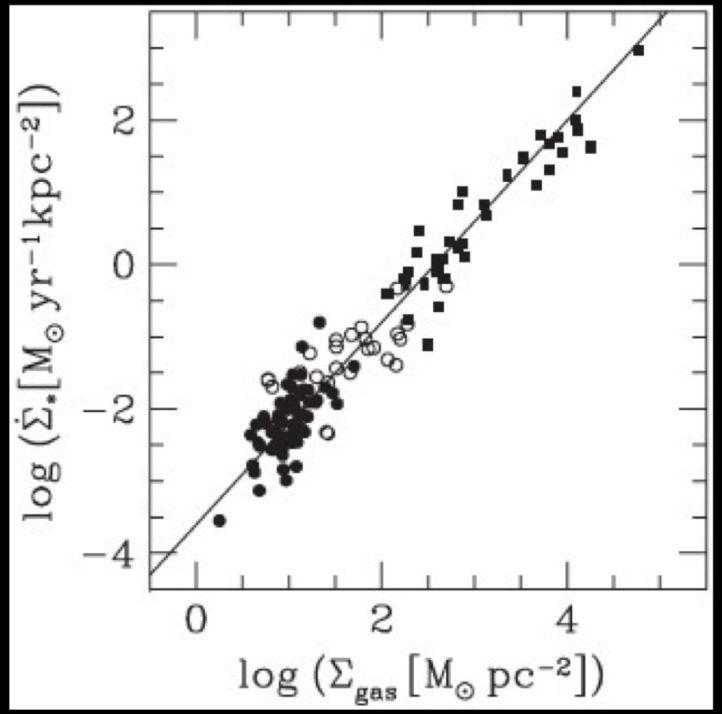
In practice, typical simulations use a simpler criterion that encompasses the pervious physics, scaling to an observationally derived from the Kennicutt (1998) relationship, which empirically relates gas surface density to star formation density, $\Sigma_{SF} \sim \Sigma_{gas}^{\alpha}$. The law is translated into a density law with one free parameter, which is the star formation timescale at the star formation density threshold.

$$\dot{\Sigma}_* \simeq 2.5 \times 10^{-4} \left(\frac{\Sigma_{\rm gas}}{M_\odot \, {\rm pc}^{-2}} \right)^{1.4} M_\odot \, {\rm yr}^{-1} \, {\rm kpc}^{-2}$$

From notes of F. van den Bosch

This is called the Kennicutt-Schmidt Relationship.

The theoretical explanation is that $\Sigma_{SF} \sim \Sigma_{gas}/t_{dyn}$, where $t_{dyn} \sim \Sigma_{gas}^{0.5}$.



Kennicutt (1998)

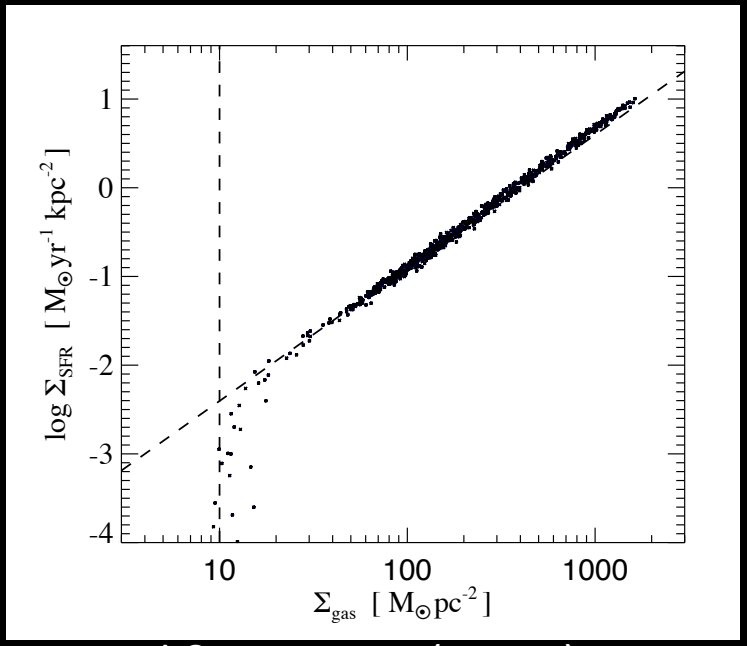
Star Formation

In practice, typical simulations use a simpler criterion that encompasses the pervious physics, scaling to an observationally derived from the Kennicutt (1998) relationship, which empirically relates gas surface density to star formation density, $\Sigma_{SF} \sim \Sigma_{gas}^{\alpha}$. The law is translated into a density law with one free parameter, which is the star formation timescale at the star formation density threshold.

$$t_{\rm SFR} = \frac{\Sigma_{\rm gas}}{\Sigma_{\rm SFR}} = 3.2\,{\rm Gyr}\left(\frac{\Sigma_{\rm gas}}{10\,{\rm M}_\odot\,{\rm pc}^{-2}}\right)^{-0.5}.$$

In the case of Springel & Hernquist (2003a), the density threshold for SF is $n_H = 0.13$ cm^{-3} and τ_{SF}~2-3 Gyr.

The star formation threshold of ~0.1 cm^{-3} is physically motivated by the UV ionizing meta-galactic background being self-shielded and allowing gas to cool and collapse into molecular clouds (e.g. Schaye 2004).

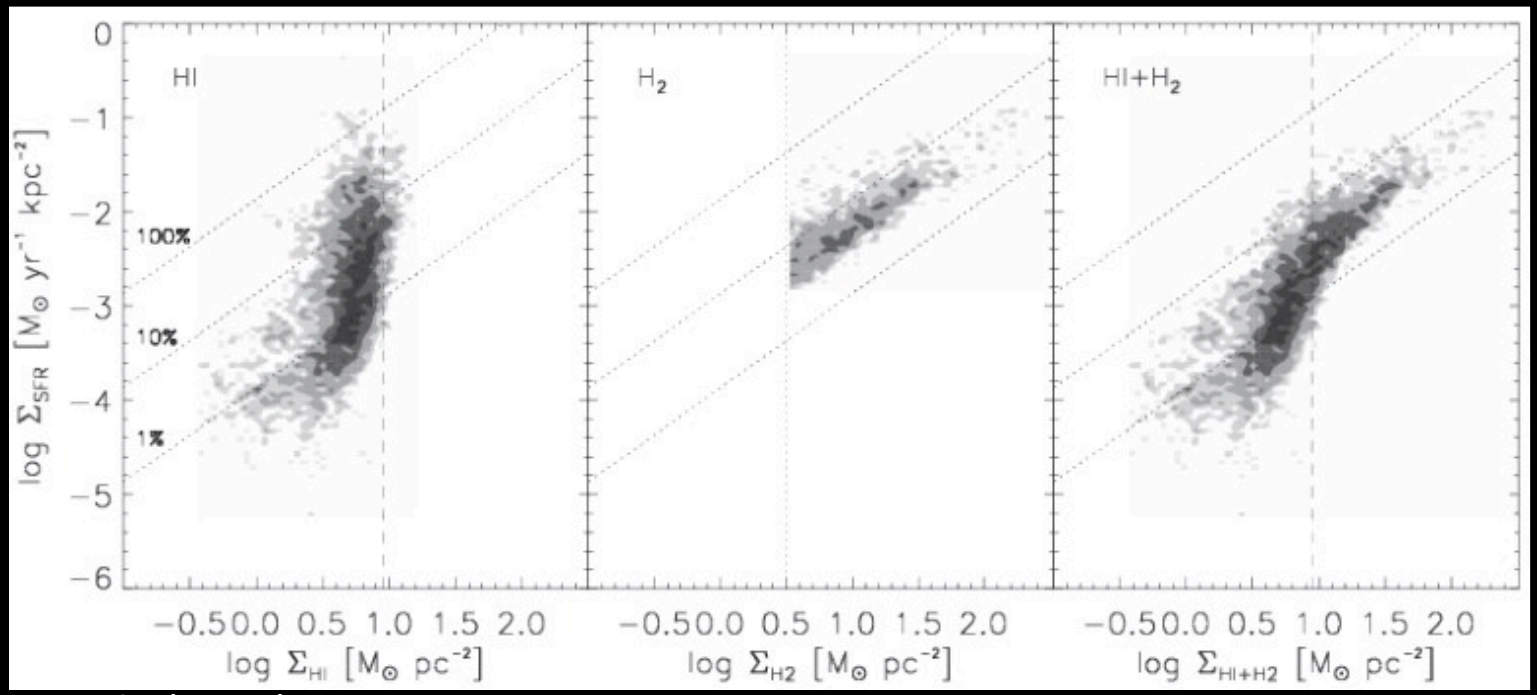


Springel & Hernquist (2003a)

Star Formation

In reality there is multiple phases of gas in the interstellar medium (ISM), including atomic hydrogen (HI, lower density) and molecular hydrogen (H_2, higher density and molecular clouds. More recent findings find a break in the K-S Law and a linear dependence of star formation rate on molecular density.

New simulations (e.g. Christensen 2012?) resolve the ISM and follow the formation of H_2 (by self-shielding from interstellar radiation) and put in a modified SF Law.



Bigiel+ (2008)

Galaxy Formation Overview

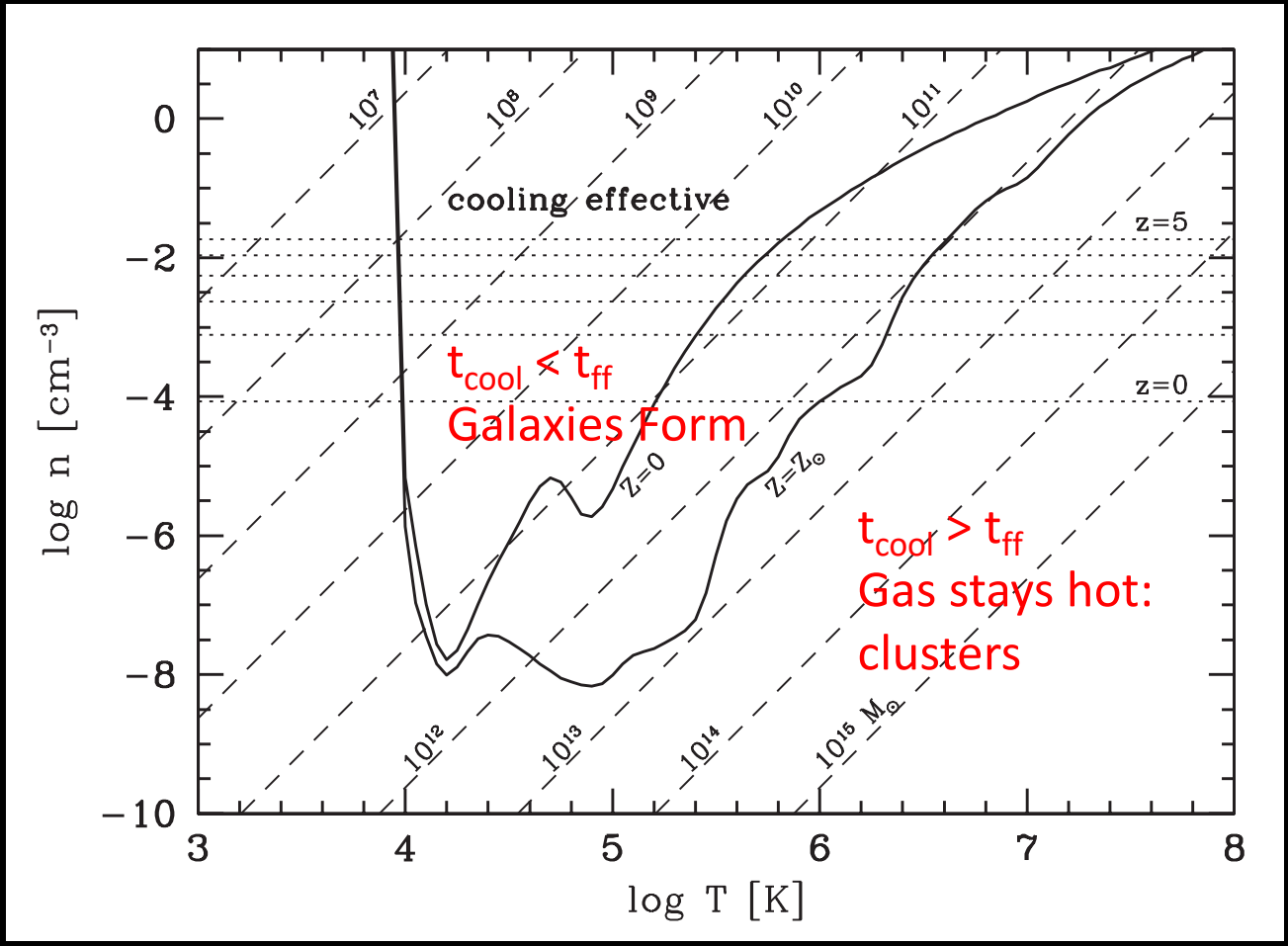
The field of “galaxy formation” encompasses how galaxies form, grow, and evolve from the first stars (at z>10-20) to the present-day Universe (z=0). Why the field is so interesting is because it encompasses so many aspects of astrophysics on all scales:

- Cosmology- cosmological parameters and Hubble expansion are encoded in the initial conditions (ICs).
- “Gastrophysics”- the physics and chemistry of intergalactic gas, which includes atomic physics responsible for cooling and ionization.
- Star formation- conditions for turning gas into stars is at the foundation of galaxy formation (i.e.- what is a galaxy fundamentally– a collection of stars).
- Stellar life and death- feedback from stars as they evolve couples to the interstellar medium and drive outflows out of the galaxy.
- Black hole physics- supermassive black holes live at the centers of galaxies, and create active galactic nuclei (quasars) that also drive outflows out of the galaxy.
- Dark matter- underlies everything in galaxy formation, required for cosmological ICs and slowing of Hubble expansion, needed to gravitationally attract baryons into halos to form stars and grow galaxies.

Galaxy Formation: How Cooling Determines if Star Formation is Achieved

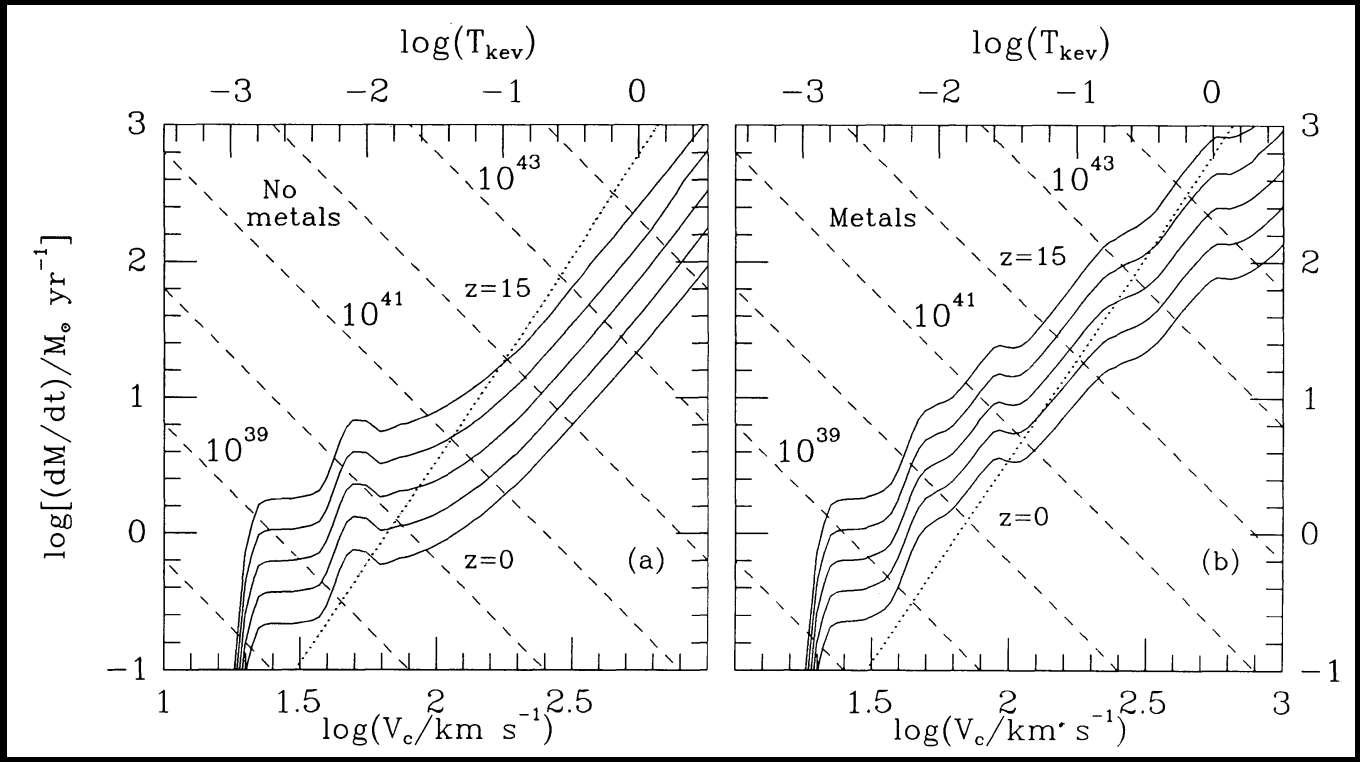
The ability of gas clouds to cool determines if they achieve the density for star formation (Rees & Ostriker 1977, Silk 1977, White & Rees 1978).

Although atomic cooling as a function of density and temperature determines if gas can cool, it is the dark matter that is needed to attract the gas into virialized structures so that they can achieve the ability to cool.



Mo, van den Bosch, White (p. 386)- density (n) and temperature (T) for gas clouds to cool. Masses of virialized clouds, which determines masses of galaxies that form and where galaxies do not form.

Galaxy Formation: How Cooling Determines What Manner Gas Accretes onto Galaxy

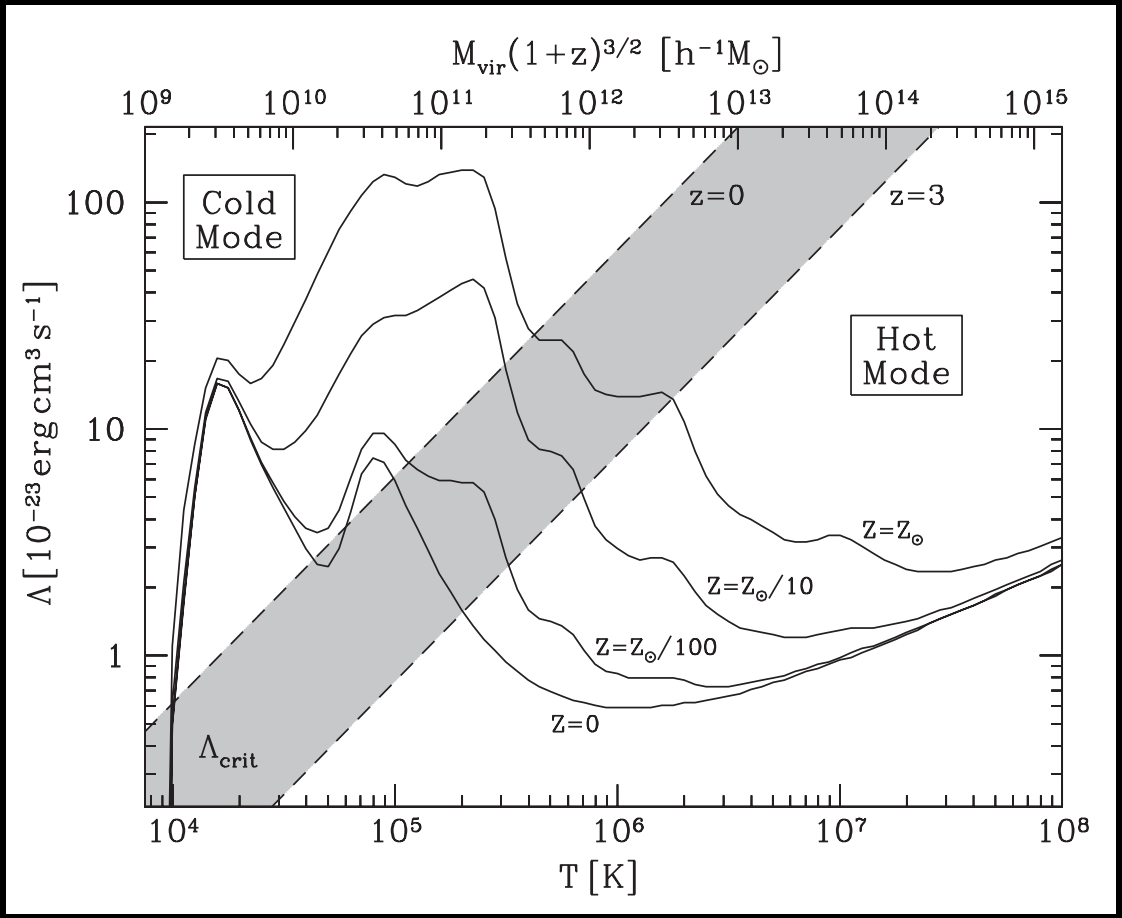


White & Frenk (1991) – The dotted line is redshift-independent gas accretion rate of halos. However gas has to cool as well, and cooling rates are shown as solid lines for 4 redshifts. If the cooling rate is higher than the accretion rate, then galaxy growth is determined by the gas accretion rate (lower velocity dispersion), but if the cooling rate is lower then the cooling accretion rate determine galaxy growth- the non-accreting gas is heated to the virial temperature of the halo instead and forms a hot halo- the intra-cluster medium.

Galaxy Formation: How Cooling Determines What Manner Gas Accretes onto Galaxy

Mass of halo that can cool easily will never shock heat to virial temperature of a halo is small and accrete via "cold mode accretion", while more massive halos shock heat their gas to the virial temperature of the halo, and form a virial shock front through which gas can cool via "hot mode accretion."

Idea of cold/hot dichotomy in how galaxies form around since White & Frenk (1991), and cosmological hydro simulations confirmed this later (Birnboim & Dekel 2003, Keres et al 2005, 2009a).

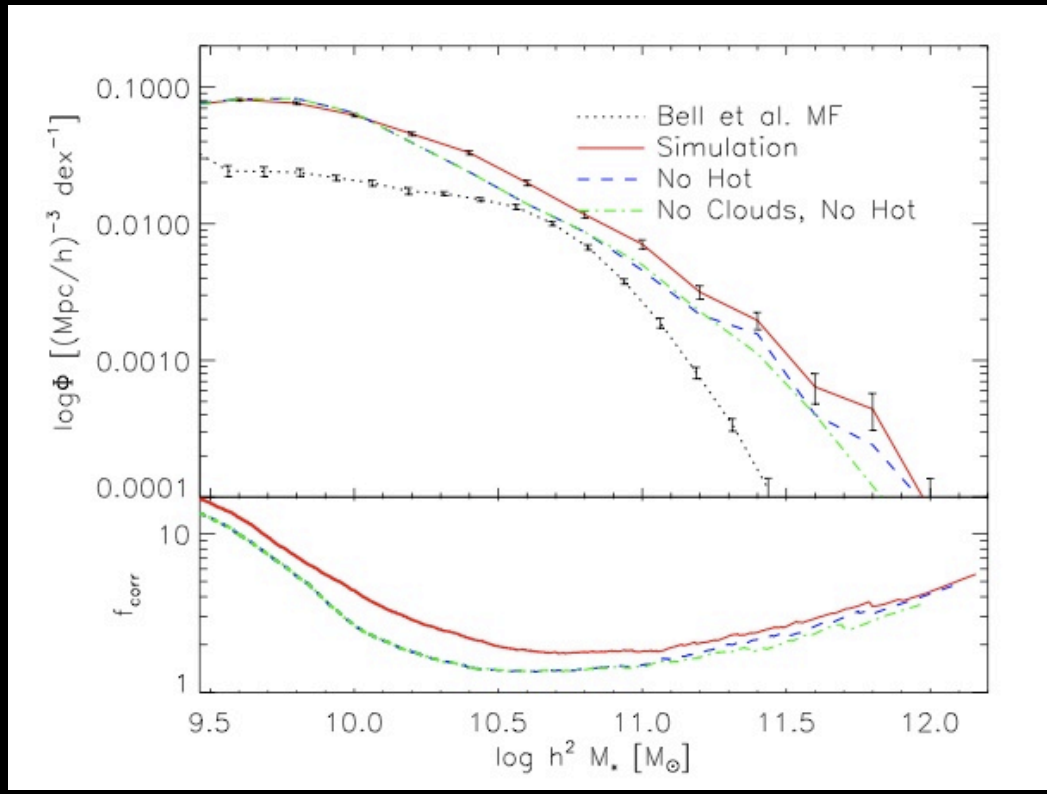


Mo, van den Bosch, White textbook (p. 393)- Gas that can that can cool easily will accrete onto halos directly, gas that can't will go through a virial shock

The Overcooling Problem- A Challenge to the Cold Dark Matter Paradigm?

A simulation with cosmological ICs, gravity for dark matter and baryons, hydrodynamics for baryons, cooling and photo-heating, and a star formation prescription. What galaxies do we get?

Simulated galaxies are too massive!
Star formation is too efficient and too many baryons get converted into stars. Up to 40% of baryons are converted into stars, when the observed mass fraction in galaxies is 4-8%, which is assuming the Ω_{baryon}=0.046, which is constrained via Big Bang Nucleosynethesis and CMB fluctuations (WMAP & Planck). This is the so-called "over-cooling" problem where too many baryons cool into halos (White & Frenk 1991, Balogh+ 2001, Keres+ 2009).



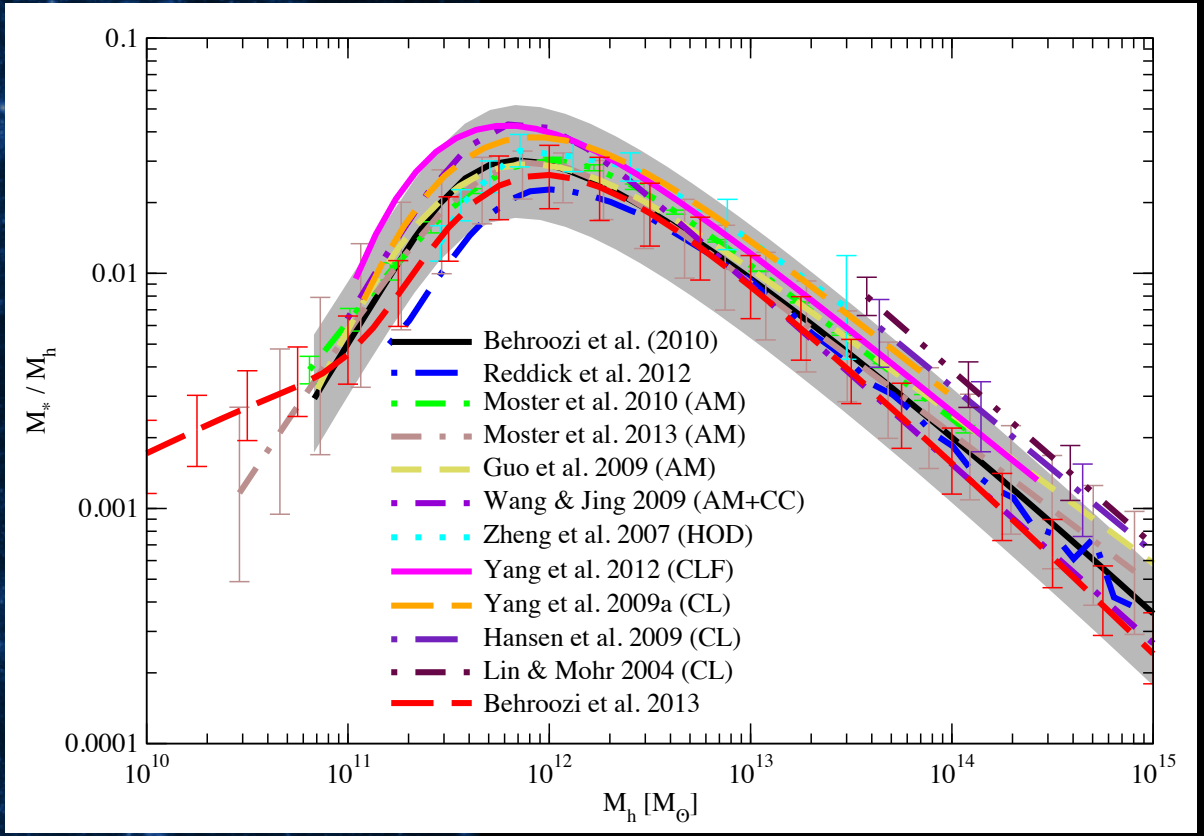
Keres et al. (2009b)- The simulated mass function of galaxies compared to the observed mass function (Bell et al. 2003) at z=0. Galaxies of all masses in the simulations

The Inferred Efficiency of Galaxy Formation

Abundance matching (AM) involves taking the distribution of observed galaxies and matching them up with the theoretical dark matter halo distribution function found in a dark matter simulation.

Abundance matching figures (on the right here) plot the ratio of M_*/M_{halo} as a function of M_{halo} for central galaaxies (we will worry about satellite galaxies later).

It is an efficiency measure of baryon conversion into stars that has a maximum value of Ω_b/Ω_M=0.046/0.28 = 0.16.



Behroozi et al. (2013)- The efficiency of halos turning their baryons into stars (a compilation of results). AM assumes the most massive galaxy occupies the most massive halo and so on toward lower masses (i.e. rank ordering of masses). This assumption is confirmed by how galaxies cluster and distribute themselves in large scale structure, because it matches the clustering of DM halos in a cosmological N-body simulation.

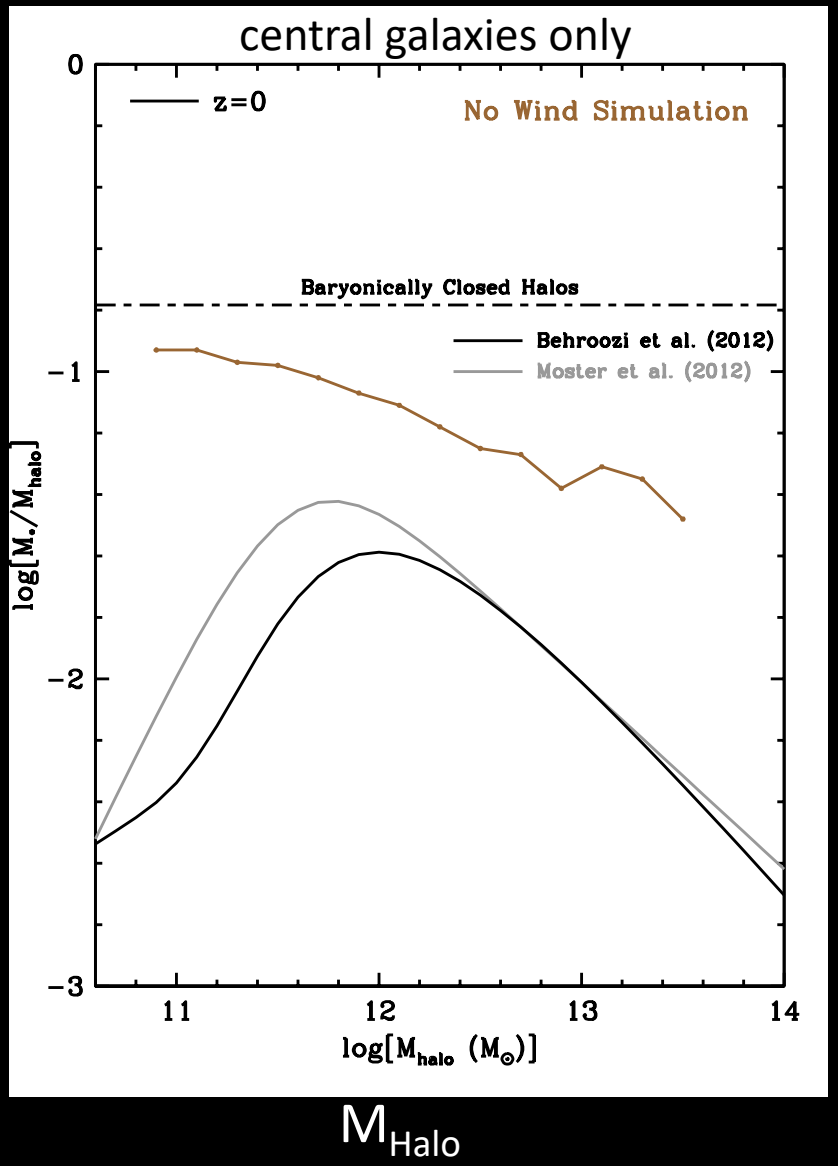
The Overcooling Problem by Halo

Too much gas cold mode accretion cools in the centers of galactic halos, turning into stars without galactic superwinds.

The masses of z=0 central galaxies are much too massive according to abundance matching with halos:

- 10^{13} $M_\odot$ - 5x too massive
- 10^{12} $M_\odot$ - 3x too massive
- 10^{11} $M_\odot$ - 10x too massive

This exemplifies the overcooling or more accurately *the oversupply problem*: *too many baryons form into stars in all halos*.

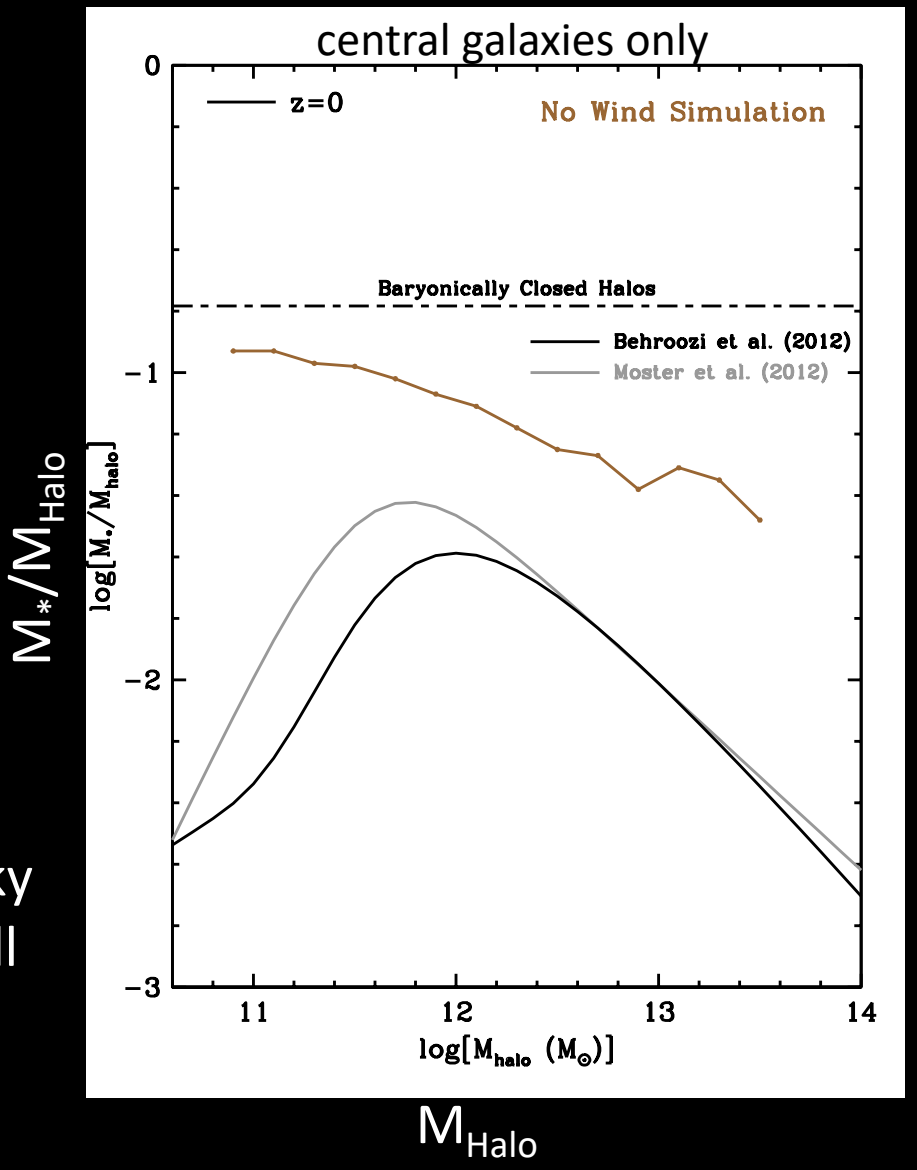


Galaxies are Gas Processing Factories

One can think of galaxies as gas processing factories (Dave, Finlator, Opp. 2012). With no galactic superwinds, galaxies make one product: stars.

No Galactic Superwinds

In fact, accretion in halos below Milky Way mass (~10^{12} M_{sol}) turn almost all of their baryons into stars. ~30% of baryons are converted to stars.

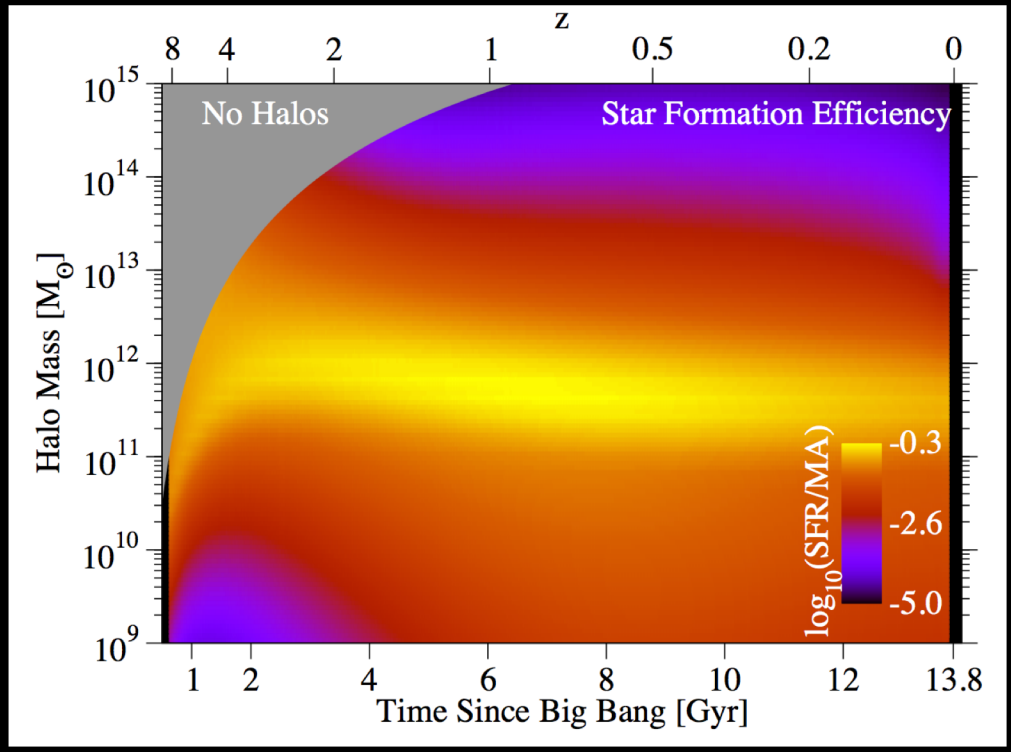


Galaxies are Gas Processing Factories

The conversion rate of baryons into stars by star formation (SFR) in the galaxy can be compared with the growth rate of the dark matter halo. The ratio that is considered is the SFR over BHAR, the baryonic halo accretion rate, which is the DM halo growth rate multiplied by $f_b=\Omega_b/\Omega_M$.

$BHAR = dM_{halo}/dt \times f_b$

SFR approaches BHAR at $M_H \sim 10^{11.5}$-$10^{12.0}$ $M_\odot$, which means SF is efficient here, however becomes very inefficient at higher and lower masses. At higher masses (>10^{12} $M_\odot$), the gas cannot cool onto galaxies, but at low masses (<10^{11} $M_\odot$) it was predicted that SF should be more efficient as gas can effectively cool.



Behroozi+ (2013b)- The efficiency of star formation relative to baryonic halo accretion rates onto halos calculated from a ΛCDM N-body simulation. The efficiency of baryon conversion into stars is a very peaked function, which is puzzling.

Galactic Superwind Feedback

Feedback in the form of energy, mass, and heavy metals nucleosynthesized in stellar evolution are returned from star formation to the interstellar medium (ISM) as can super-massive black hole feedback from active galactic nuclei (AGN).

1) Star formation-driven feedback

- Supernova energy (10^{51} ergs/SN) is thermalized and converted into kinetic outflows leaving the disk of a galaxy (Chevalier & Clegg, 1985).
- Radiation pressure- UV photons from OB stars create radiation pressure that couples to dust and drives gas out of a galaxy (Murray, Quaertart, & Thompson 2005). These are considered “momentum-conserving winds.
- Also, stellar winds,l photo-ionizational heating, and cosmic rays from supernovae, can also couple to the ISM and drive galactic superwinds.

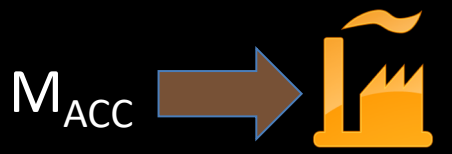
2) Black hole (BH) feedback- a small fraction of the rest-mass energy from accretion onto a central black hole can be re-radiated, which is a significant amount of energy

- Radiation pressure from an AGN (a quasar) couples to gas and drives a galactic scale wind quenching energy
- Jets emanating from the BH extend Mpcs and transfer energy to the interglactic medium.
- “Radio”-mode feedback of continued radio-active AGN (evolved ellipticals in clusters/groups) couples to the surrounding hot halo gas and prevents later inflow.

The Overcooling Problem:
Solved by Superwind Feedback?!

Star formation-driven galactic superwind (GSW) feedback is implemented in our simulations, usually by giving a kicking wind particles at a rate proportional to the star formation rate.

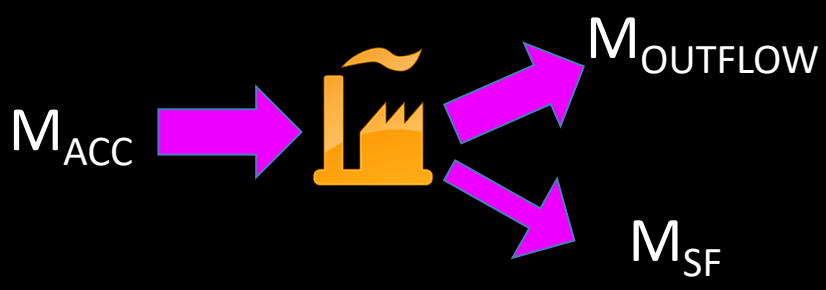
No Galactic Superwinds



Two basic parameters to apply feedback:

Mass loading where: $M_{wind} = \eta \times M_{SFR}$
Wind velocity where: v_{wind}

$\eta \sim \sigma^{-1}$ Mass-loading



Momentum-conserved feedback (vzw)
where σ is the galaxy velocity dispersion
$v_{wind} \sim \sigma$, $\eta \sim \sigma^{-1}$ where $M_{wind} = \eta \times M_{SFR.}$

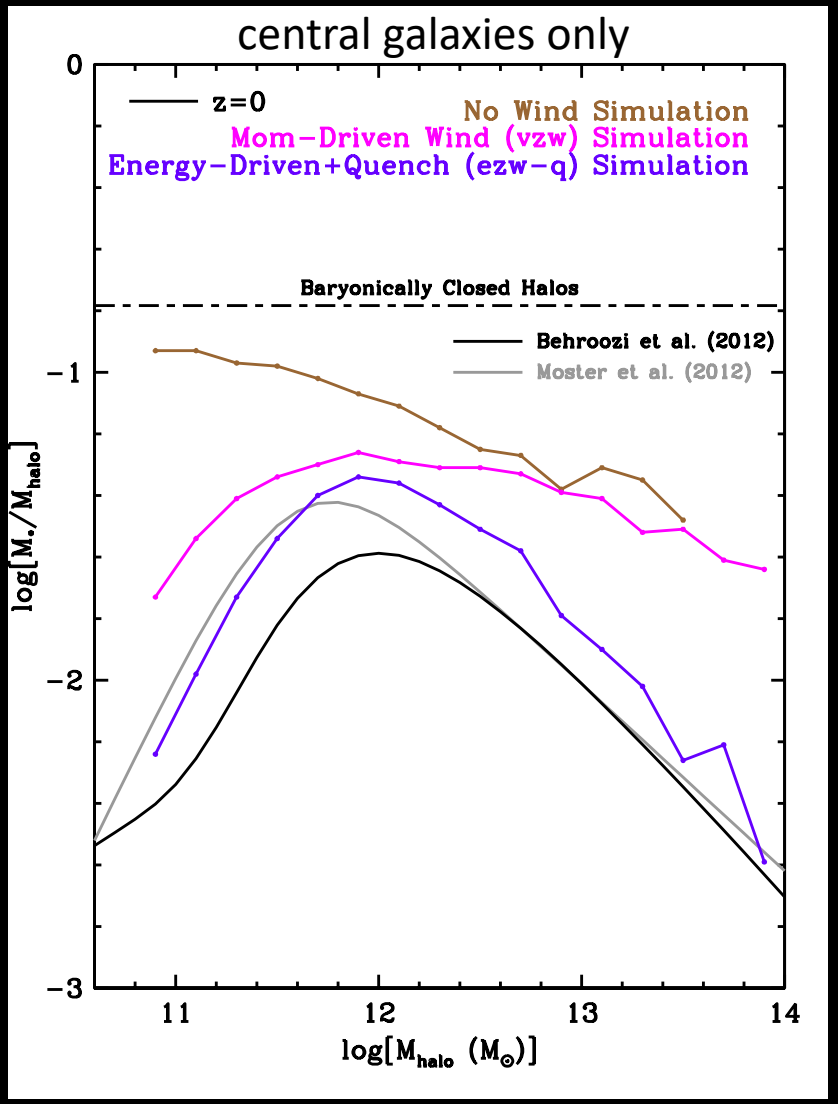
Dave, Finlator, Opp. 2012

The Overcooling Problem:
Solved by Superwind Feedback?

Galaxy and star formation is suppressed by kicking

Momentum-conserved feedback (vzw) where σ is the galaxy velocity dispersion $v_{wind} \sim \sigma$, $\eta \sim \sigma^{-1}$ where $M_{wind} = \eta \times M_{SFR}$ ~50% SN energy into kinetic winds.

Also, I've done energy-conserved feedback (ezw) where $v_{wind} \sim \sigma$, $\eta \sim \sigma^{-2}$ where low-mass galaxies are suppressed more (and quenching at high masses).



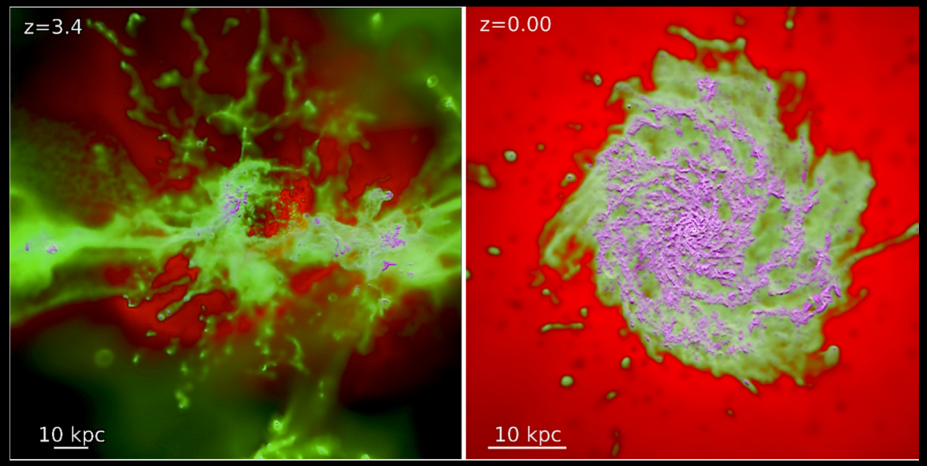
M_*/M_{Halo}

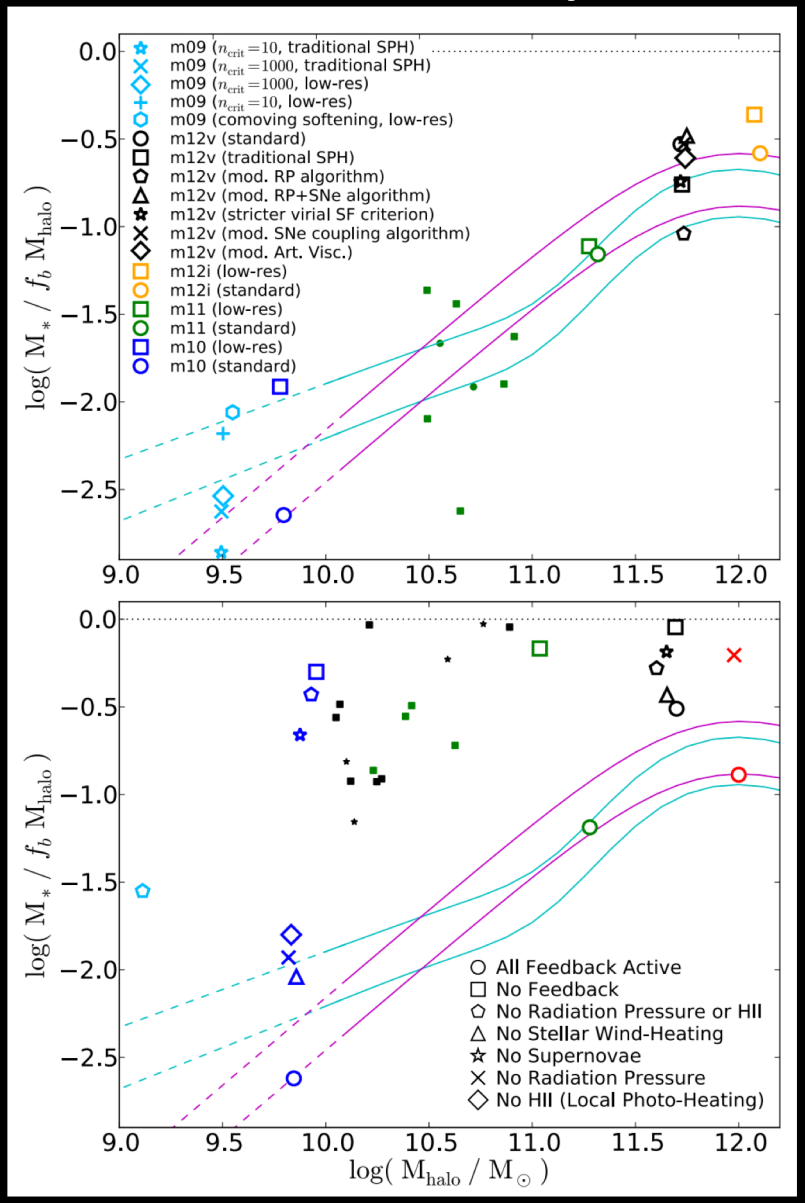
M_{Halo}

Zoom simulations can resolve how stellar feedback severely reduces dwarf satellite efficiency

Hopkins+ (2014) run cosmological zoom simulations and resolve galaxies living in halos down to ~10^9 M_{sol}. Can resolve the ISM, and include multiple forms of feedback from supernovae and stellar winds as well as radiative transfer for radiation pressure (momentum-driven winds), and photo-ionization.

The combined effect of these multiple forms of stellar feedback resonate with each other, and reduce setllar masses by a factor of 100x in lower mass galaxies.



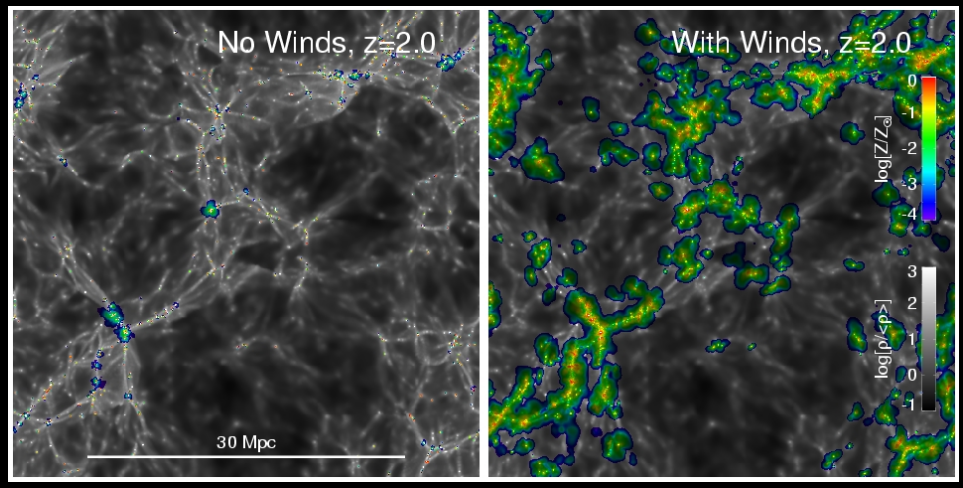


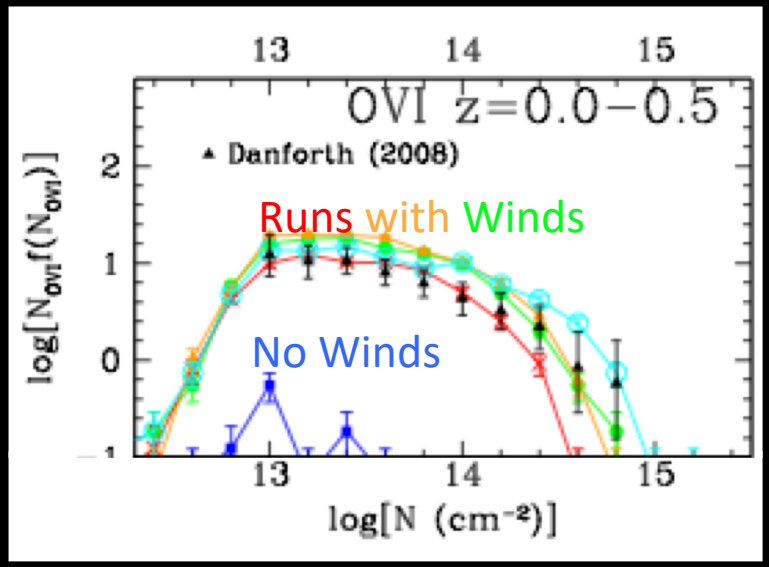
Another Effect of Feedback is to Chemically Enrich the Gas Outside Galaxies- the Intergalactic Medium

The diffuse gas in the intergalactic medium (IGM) is too underdense to form stars, and it is expected to have primordial composition (76% Hydrogen, 24% Helium).

Instead there is a significant amount of "heavy metals" (e.g. carbon, oxygen, silicon, iron) in the IGM, and only simulations with galactic superwinds can transport the material from galaxies where these metals are nucleosynthesized and ejected from stars to the IGM.

Hence, in addition to feedback solving the overcooling problem by reducing the stellar content of galaxies, the signatures of feedback are seen in the pollution of the cosmic web. We will also see that feedback can alter dark matter halo profiles as well.





Oppenheimer, Davé+ (2006, 2008, 2009ab, 2012)- frequency of heavy element absorption in simulated quasar sight lines.

Challenges to ΛCDM Cosmology

Cold dark matter has had many key successes including: rotation curves of galaxies, large scale clustering of galaxies, the Lyman-α forest power spectrum, weak lensing tracing dark matter; not to mention the CMB anisotropies, Big Bang nucleosynthesis, and baryonic acoustic oscillations.

The most notable possible tensions under ΛCDM are:

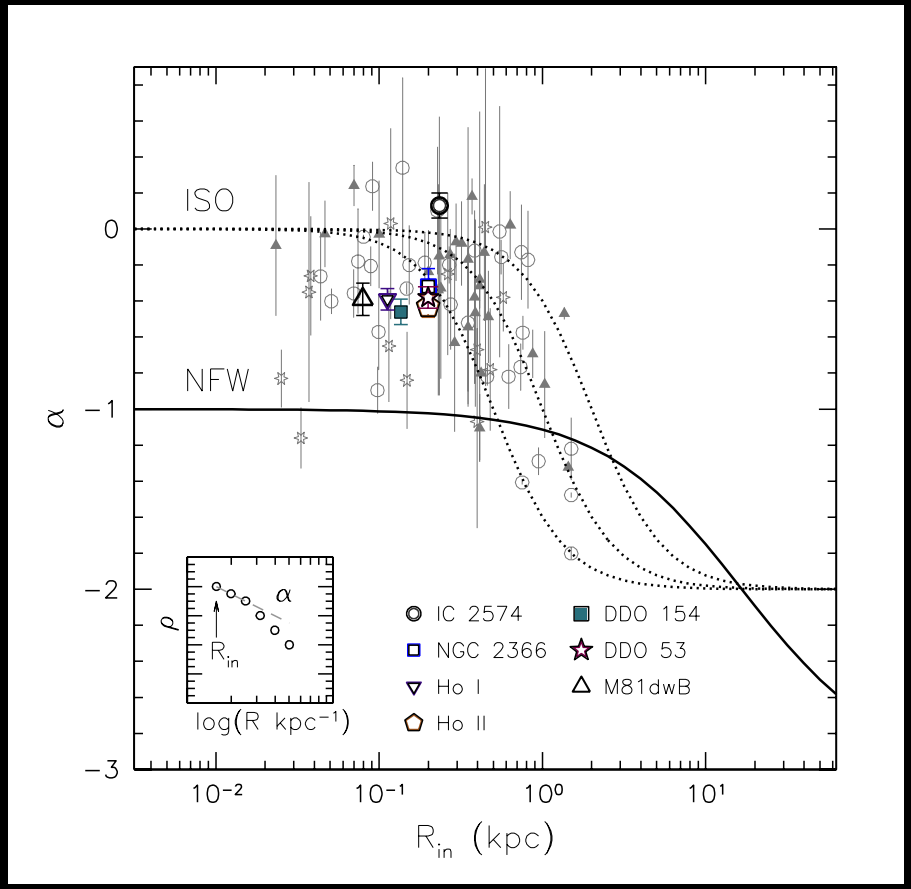
1) The existence of core-like profiles of galactic cores: the core-cusp problem.
2) The predicted excess of the number of Milky Way satellite halos compared to observed galaxies: the missing satellites problem.
3) The low velocity dispersions of Milky Way satellite: Too big to fail.
4) The properties of galaxies filling voids

Dark Matter Halo Profiles: Deviations from NFW Owing to Baryons

NFW profiles are expected to have "cusp-like" inner DM profiles such that $\rho \sim r^{-1}$, because $\rho =\sim 1/(r(1+r^2))$.

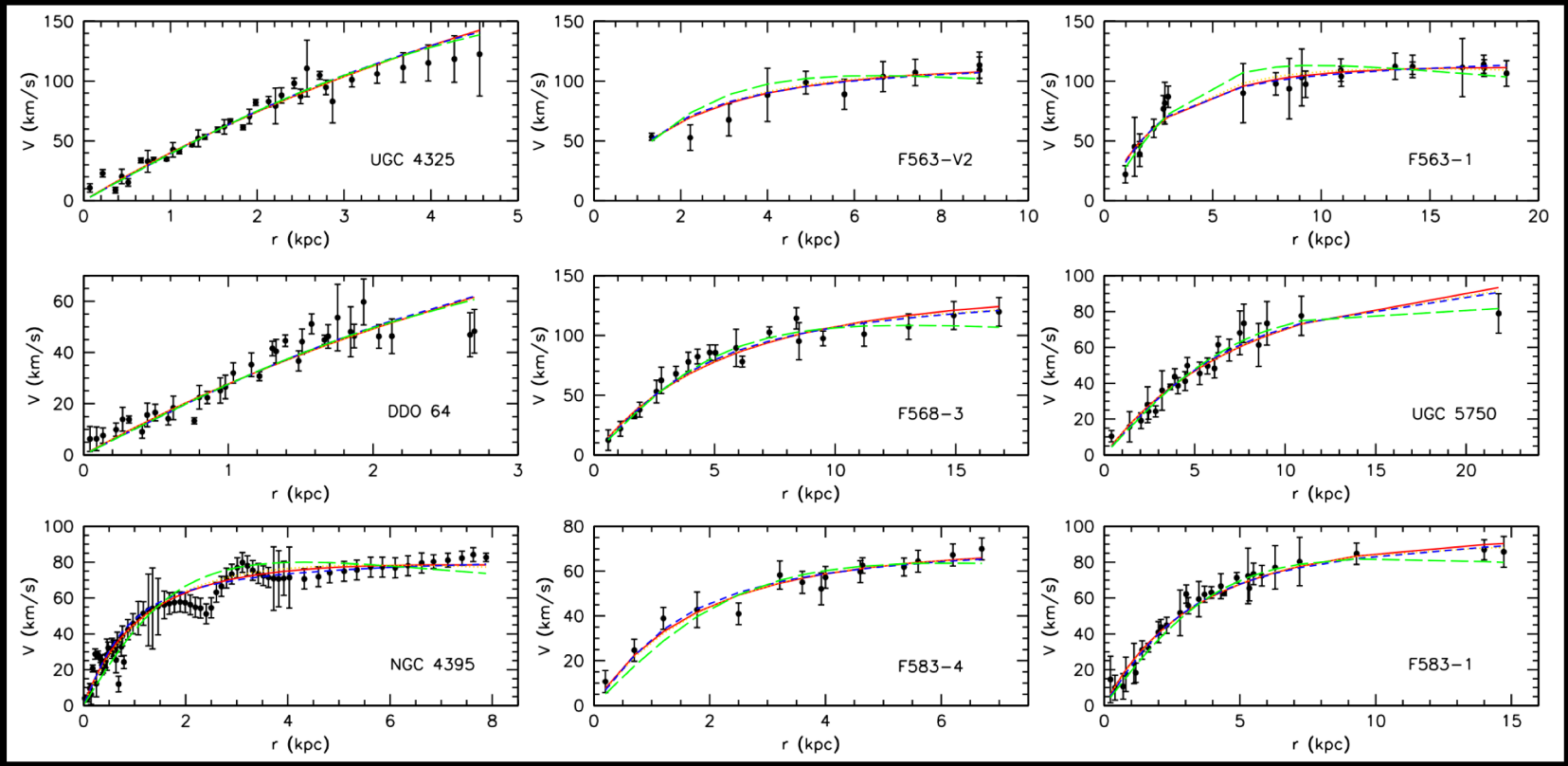
However, observations suggest an even shallower profile where $\rho \sim r^{\alpha}$, and $\alpha > -1$. These flatter profiles are "core-like" with some profiles approaching a flat inner distribution $\alpha = 0$.

Observations of dwarf galaxies, Milky Way satellites, and even large clusters indicate that the inner dark matter profiles are more core-like than cusp-like as expected from NFW. ***This is referred to as the core-cusp problem.***



Oh et al. (2011)- Inner slope of DM profile inferred from gas in dwarf galaxies. $\alpha = -1$ is a cusp-like NFW profile, and $\alpha > -1$ is more core-like.

Observations of Non-NFW Cored Profiles

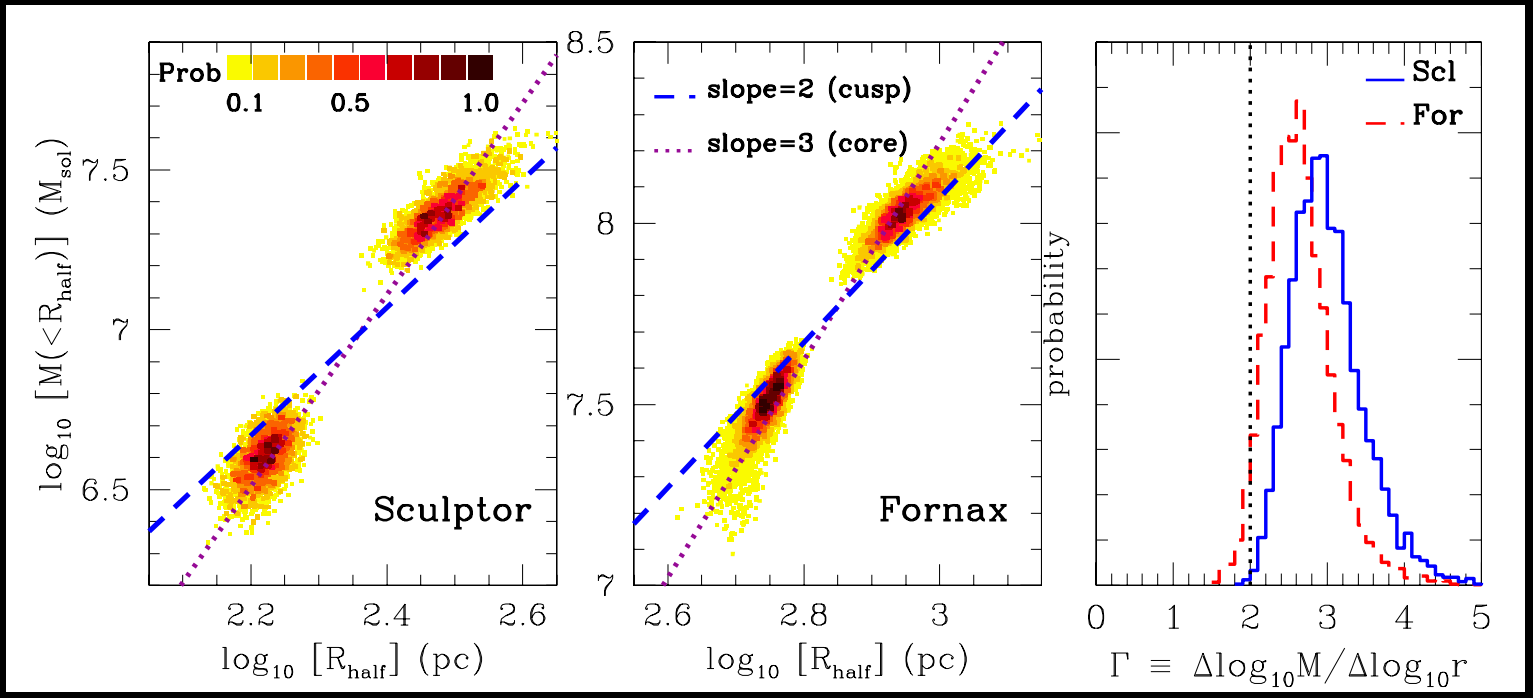


Kuzio de Naray et al. (2010)- Slowly rising rotation curves observed in low-surface brightness galaxies.

For $\rho \sim r^{\alpha}$, the circular velocity, $v_{circ} \sim r^{1-\alpha/2}$

Hence, for a cored profile, $\alpha=0 \rightarrow v_{circ} \sim r$ and for a cusped profile $\alpha=1 \rightarrow v_{circ} \sim r^{0.5}$ and a maximum of v_{circ} occurs outside of where $\alpha=2$.

Observations of Non-NFW Cored Profiles

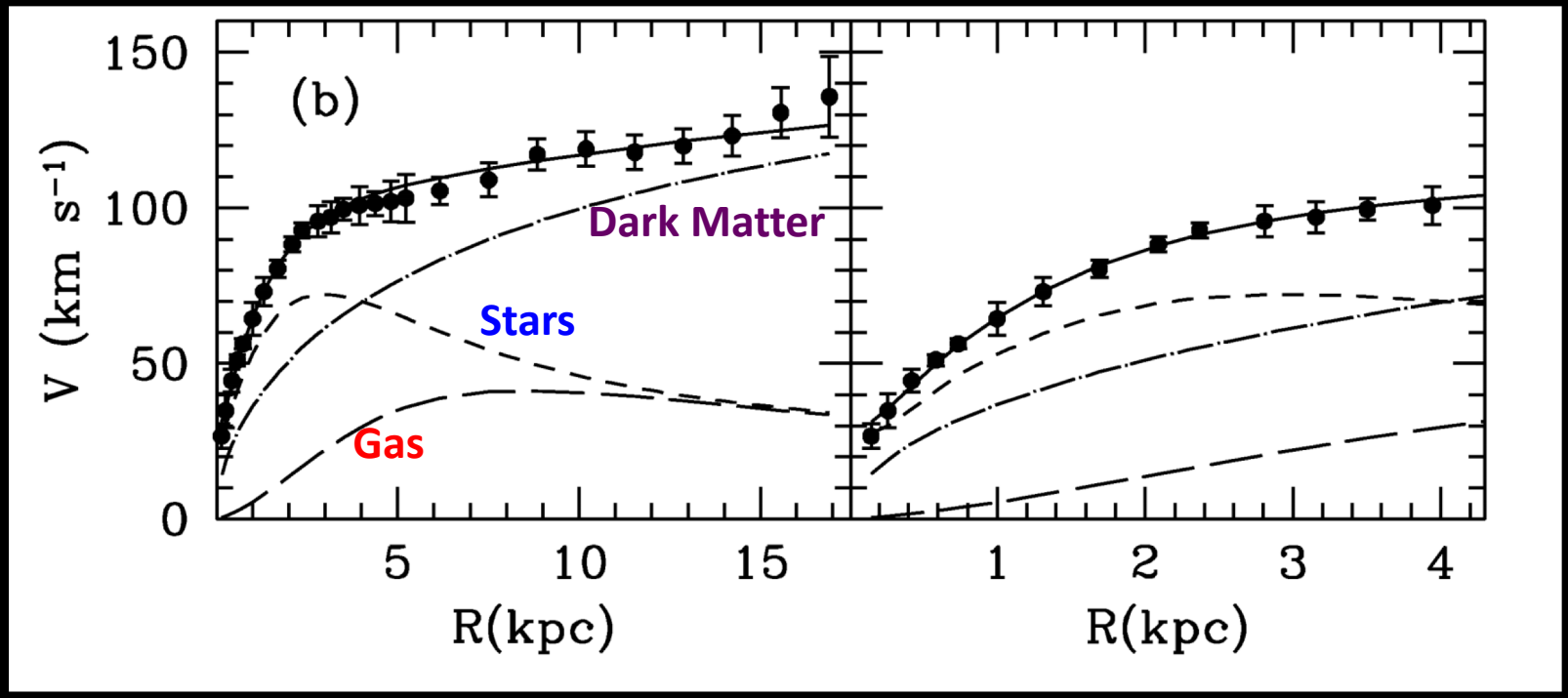


Walker & Penarrubia (2011) observe the distributions of two different stellar subcomponents in local Milky Way dwarf spheroidals to constrain the mass distribution as a function of radius

$$\Gamma \equiv \frac{\Delta \log M}{\Delta \log r}$$

Γ=2 for NFW cusp, Γ=3 for cored profile. Dwarf spheroidals (DSs) have a divergent core from a NFW cusp. DSs more core like.

Observations in the Universe: not just Dark Matter, but Stars and Gas



Corbelli et al. (2003) observe the rotation curve of M33 (local Group, small spiral). They observe the circular velocity as a function of radius, as well as the radial mass distribution of stars and gas directly, and infer "missing mass" is dark matter. Technique pioneered by V. Rubin (1970's papers).

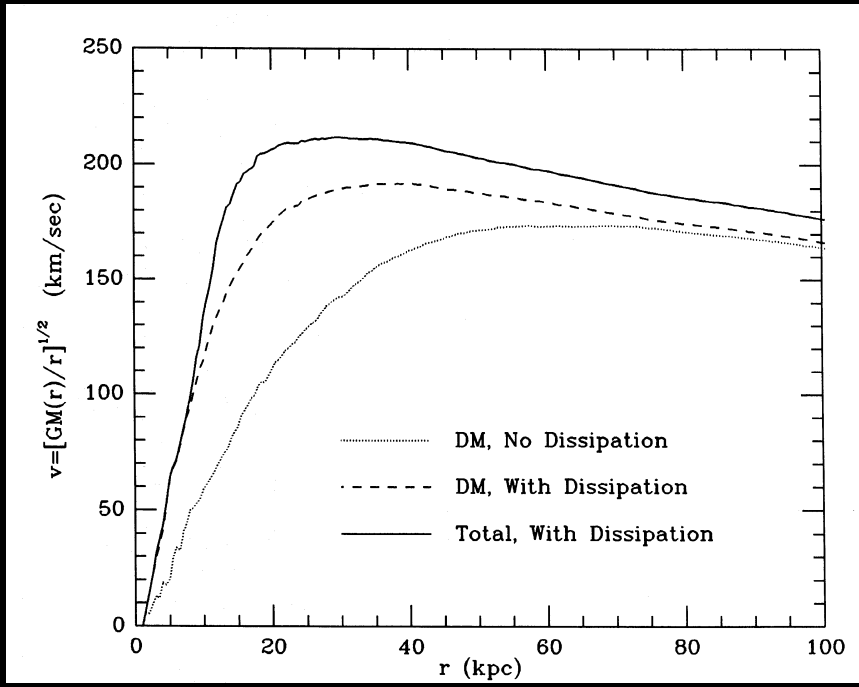
Need to consider dissipative baryons that form stars and how they affect dark matter.

Adiabatic Contraction: The Cooling of Baryons Slowly Reshaping a DM Profile

Blumenthal et al. (1986) considered the dissipational effects of baryons causing a contraction of the core. The core contracts slowly as the timescale for infall is longer than the dynamical timescale- a "squeezable" core results that they argued was reversible.

The result was the formation of a flatter rotation curve as a function of distance, and a more cuspy density profile in excess of the a DM-only simulation. ***Adiabatic contraction goes in the opposite direction of making DM profiles core-like, and exaggerates the core-cusp problem.***

Semi-analytic models (SAMs) that track halo growth in DM-only simulations use the adiabatic contraction formula from Blumenthal et al. (1986), however we will see that this is not the right picture when we consider the effect of outflows, which rapidly change the profile.



Blumenthal et al. (1986)- the effect of dissipational baryons steepening the rotation curve of a galaxy. Forming a peak and then a decline.

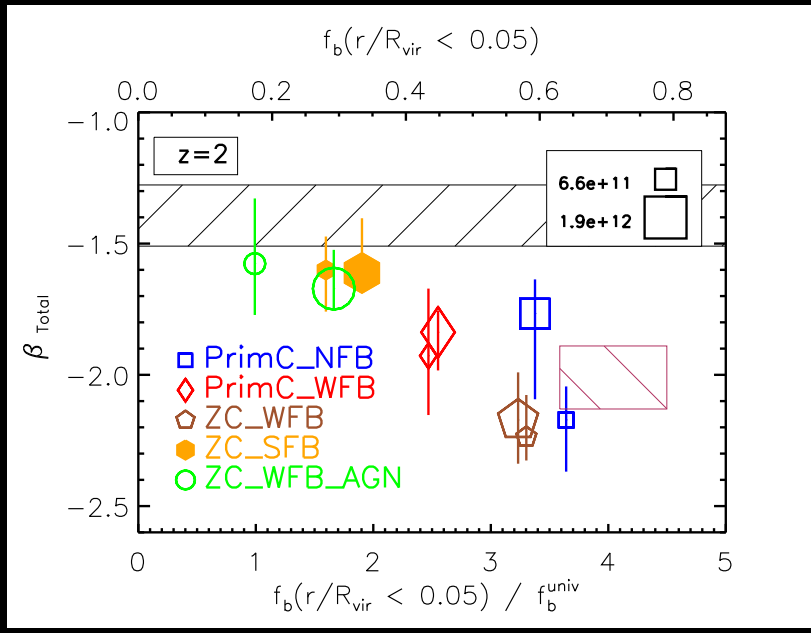
The Back-reaction of Baryons on DM Halo Profiles: Radiative Cooling

DM-only simulations give "cuspy" profiles at the center as the halo profile approaches $\rho \sim r^{-1}$.

However, the baryonic effect is great in the center, owing to radiative cooling allowing gas to lose its thermal energy and settle in the center. What results is a greater concentration of baryons, where baryonic actually exceed the dark matter mass in the inner few kiloparsecs of a L* galaxy (i.e. $M_H \sim 10^{12}\ M_{sol}$).

A halo profile results that is even more concentrated in the center than a NFW profile, and in fact very nearly the radial slope of an isothermal sphere, $\rho \sim r^{-2}$ (Koopman+ 2006).

In baryonic simulations without strong feedback, the DM is shaped into a very cuspy, isothermal-like inner profile.



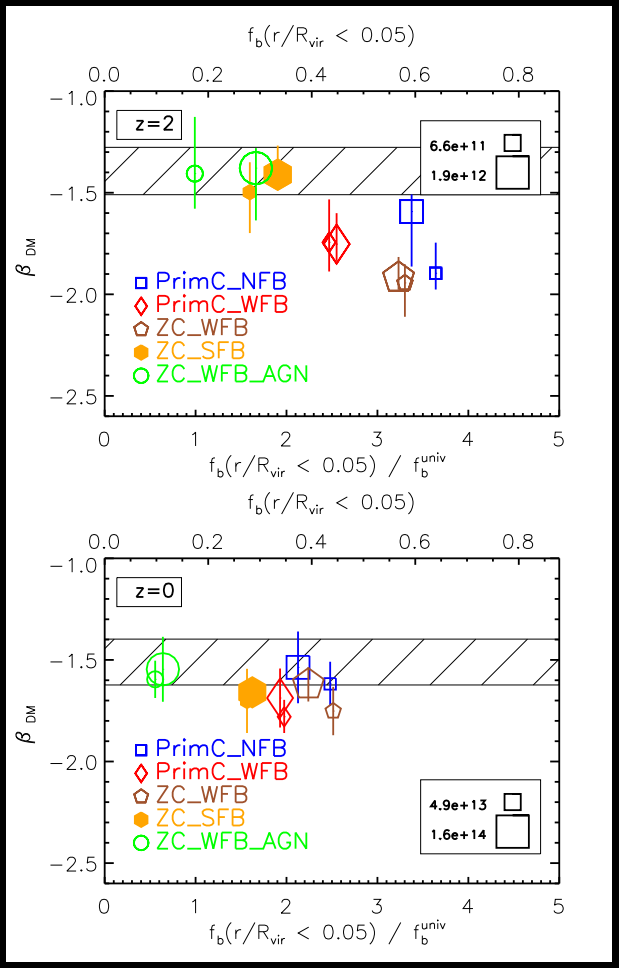
Duffy et al. (2010)- observations from Koopman et al. (2006) in hashed brown region, colors are simulations, and hashed black region is DM-only profile (self-similar). Inner halo profile measured from r=0.025-0.05 R_{vir}.

The Back-reaction of Baryons on DM Halo Profiles: Superwind Feedback

The effect of feedback is the opposite of radiative cooling. Baryons are kicked out of the central regions in galactic superwinds driven by star-formation or active galactic nuclei (AGN), which operate in the centers of galaxies and can significantly reduce the slope of the inner profile as well as the baryon fraction.

Surprisingly, the AGN feedback prescription in Duffy+ (2010) counter-balances cooling and make a halo profile much more similar to the DM-only profile.

Feedback makes inner profiles more "core"-like, i.e. flatter profiles. Feedback is less effective in larger halos, because gravitational binding energy grows faster with halo mass than the feedback energy available to unbind the profile, which is proportional to the galaxy stellar mass and/or black

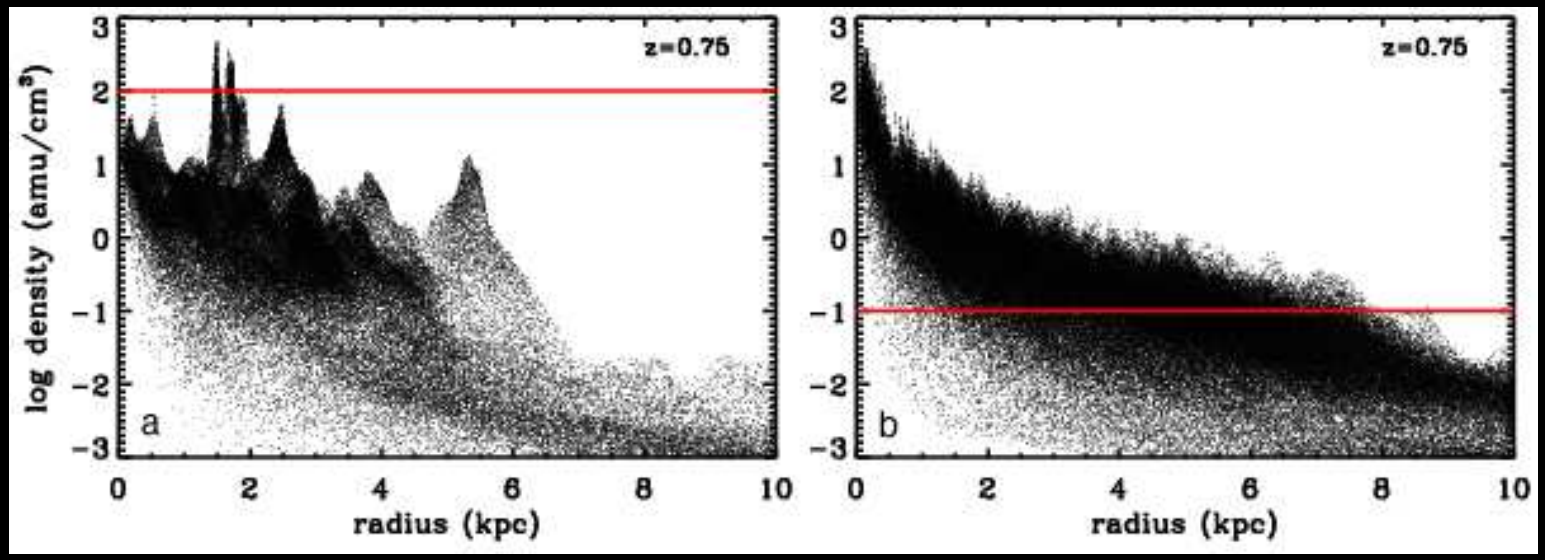


Duffy et al. (2010)- Galactic-scale halos on top at z=2, Cluster/group-scale halos on bottom at z=0.

The Effect on Feedback of Changing Star Formation Density Threshold

Using zoom simulations, Governato+ (2010) set a higher star-formation density threshold. This allows the creation of higher density peaks where feedback can be imparted to lower density where the thermal energy cannot is not radiated away.

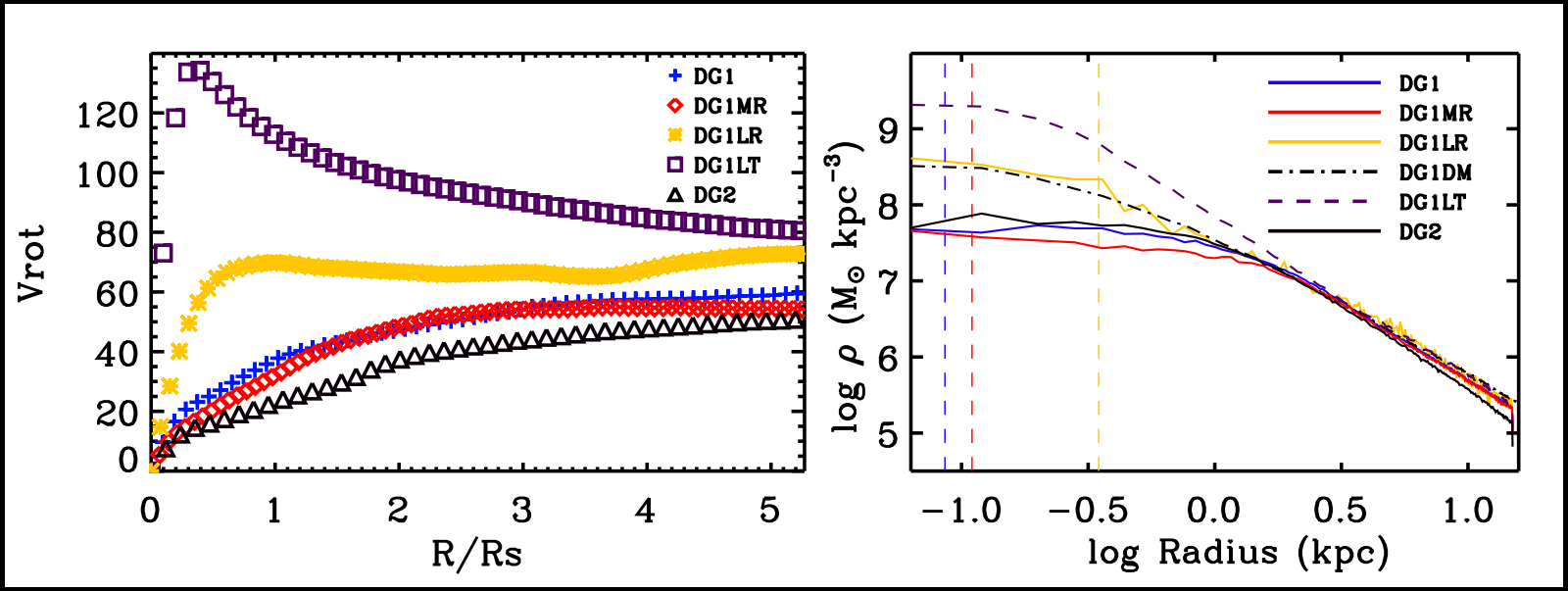
Resolving the substructure of the ISM is critical for creating more efficient feedback.



Governato+ (2010)- Simulations of gas density as a function of radius in 2 simulations. The red line shows the star-formation density criterion.

The Effect on Feedback of Changing Star Formation Density Threshold

With a higher star formation threshold, the distribution of dark matter and baryons turned from cusp-like to core-like.

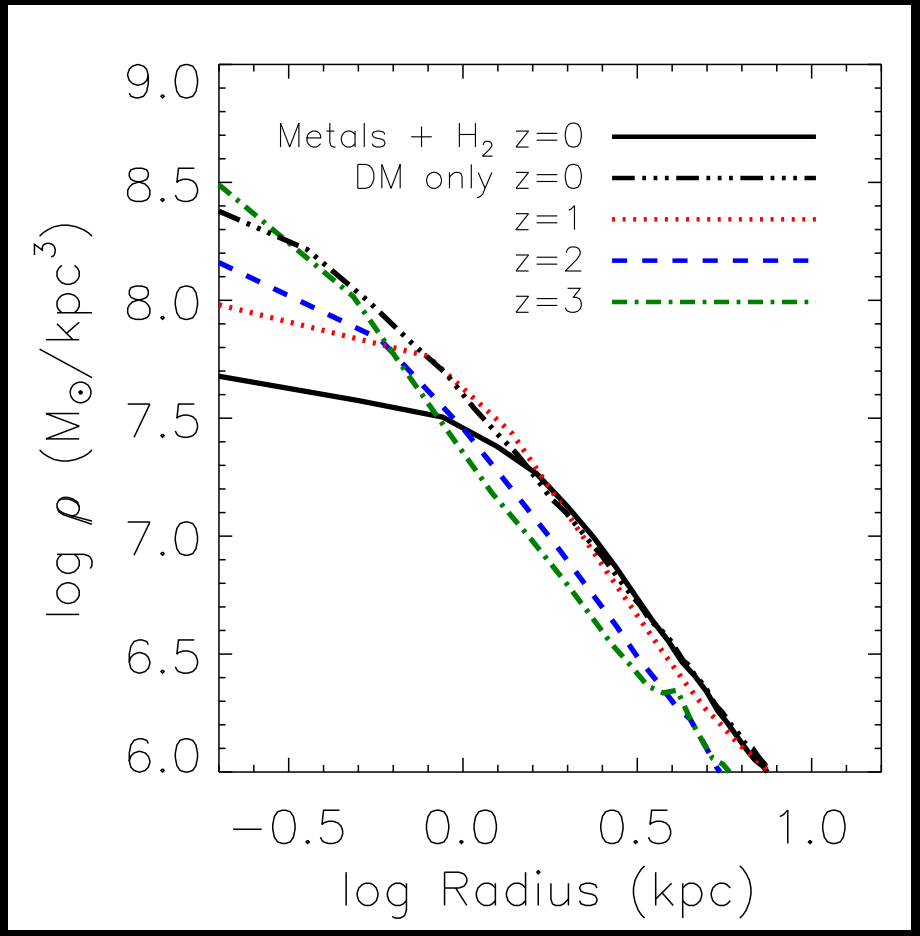


Governato+ (2010)- The rotational velocity curve is not peaked in the inner scale radius with the ISM resolved, but instead it is slowly rising (left), and the matter density distribution as a function of radius show an inner core, while a gas simulation with ineffective resolution makes an even more cusp-like core than a dark matter only simulation (right).

Feedback gradually flattens DM profiles in low-mass centrals

In simulations of dwarf galaxies with supernova thermal feedback, the inner profile evolves to become more core-like from z=3->0, while the DM-only profile maintains an NFW profile throughout.

The key is gas is continually removed from the inner profile by winds, and the profile flattens over a Hubble time.

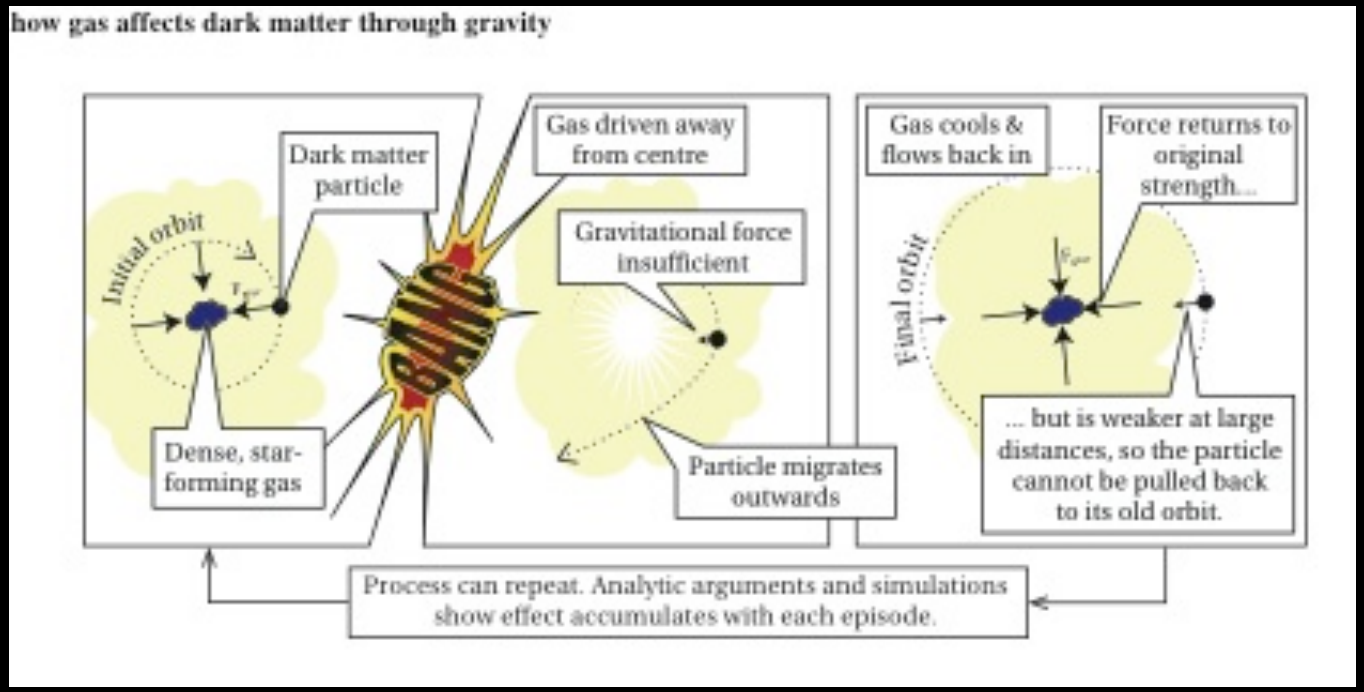


Governato+ (2012)

Feedback gradually flattens DM profiles in low-mass centrals

Removal of baryons from a DM halo is a fast process that does not follow the slow nature of adiabatic contraction. Circular orbits of DM in the center widen as baryons are removed by feedback, and do not re-contract when baryons re-accrete in the center.

A cusp-like NFW profile $\rho \sim r^{-1}$ becomes a core-like profile, $\rho \sim r^{0}$.

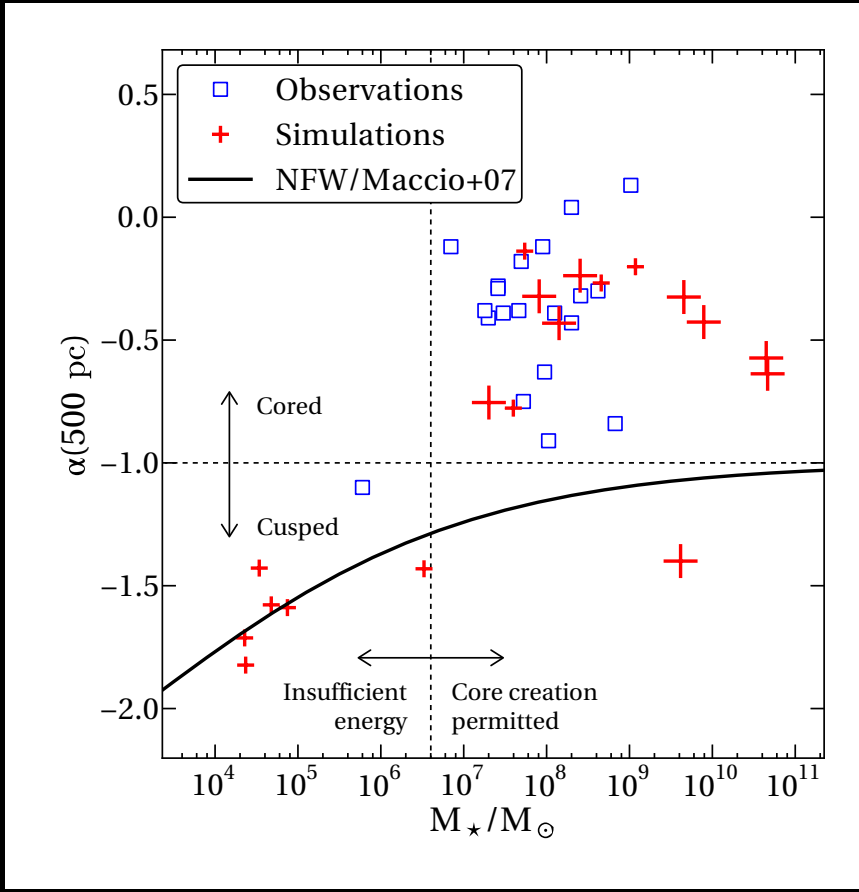


Pontzen & Governato (2014)

Cusped versus Cored Profiles in Dwarf Galaxies

Dwarf galaxies in the Milky Way halo and local dwarf irregulars can be used to measure the radial profile in the inner 500 pc.

More massive dwarfs, $M_* > 10^7\ M_{sol}$ have core-like profiles, but simulations suggest steeper profiles for $M_* < 10^7\ M_{sol}$ because star formation is so inefficient, that the feedback mechanism cannot remove baryons. Hence these galaxies are predicted to have cusp-like, NFW profiles. Observing these profiles is very hard to do, owing to the lack of tracers (stars) in them.



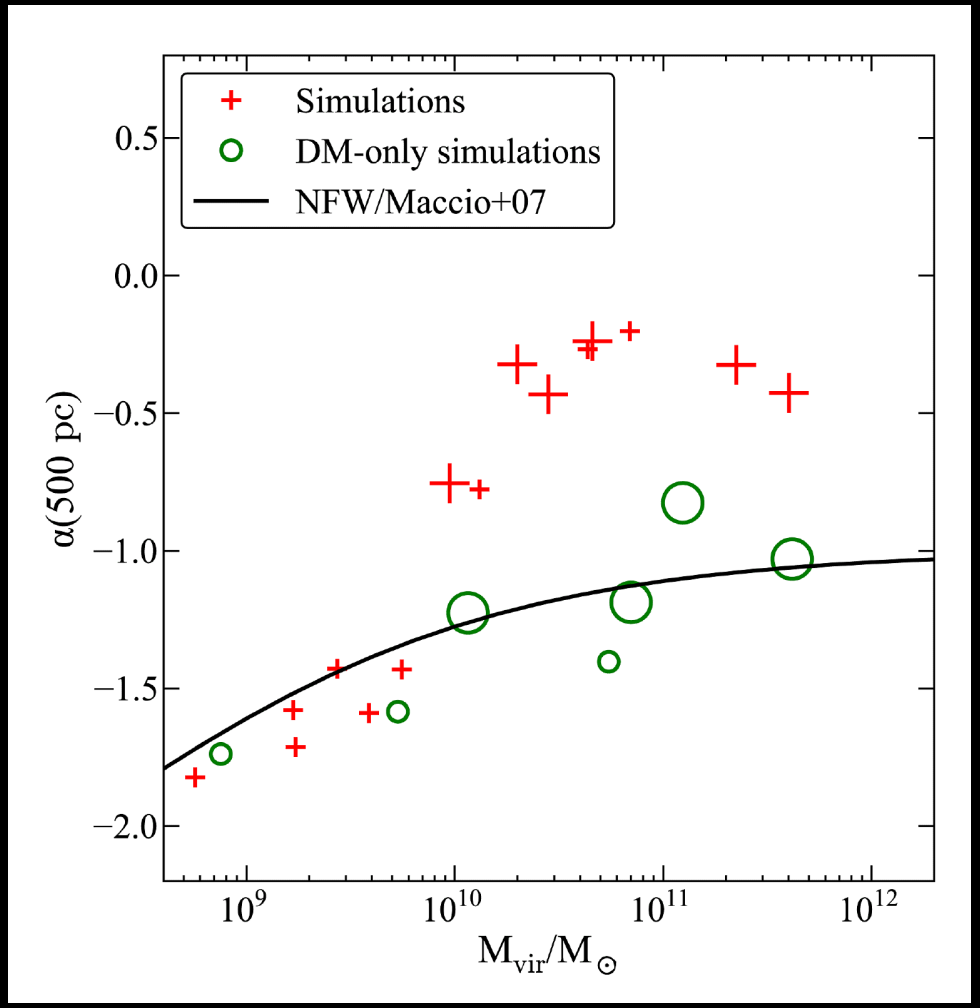
Governato+ (2012)- Inner slope of dark matter within 500 pcs inferred from dwarf galaxies (blue) and measured in simulations (red).

Cusped versus Cored Profiles in Dwarf Galaxies

Slopes of dark matter within 500 pc now shown as a function of halo mass.

The cutoff for cored profiles at $M_* < 10^7\ M_{sol}$ translates to a $M_{DM} < 10^{10}\ M_{sol}$. No observations are shown, but DM-only simulations show the relation expected fro NFW profiles (they are self-similar, but the divergence is because 500 pcs is larger fraction of the scale radius for a smaller halo).

The hydro simulations show that around $M_{DM} \sim 10^{10}\ M_{sol}$ baryons can shape a dark matter halo.



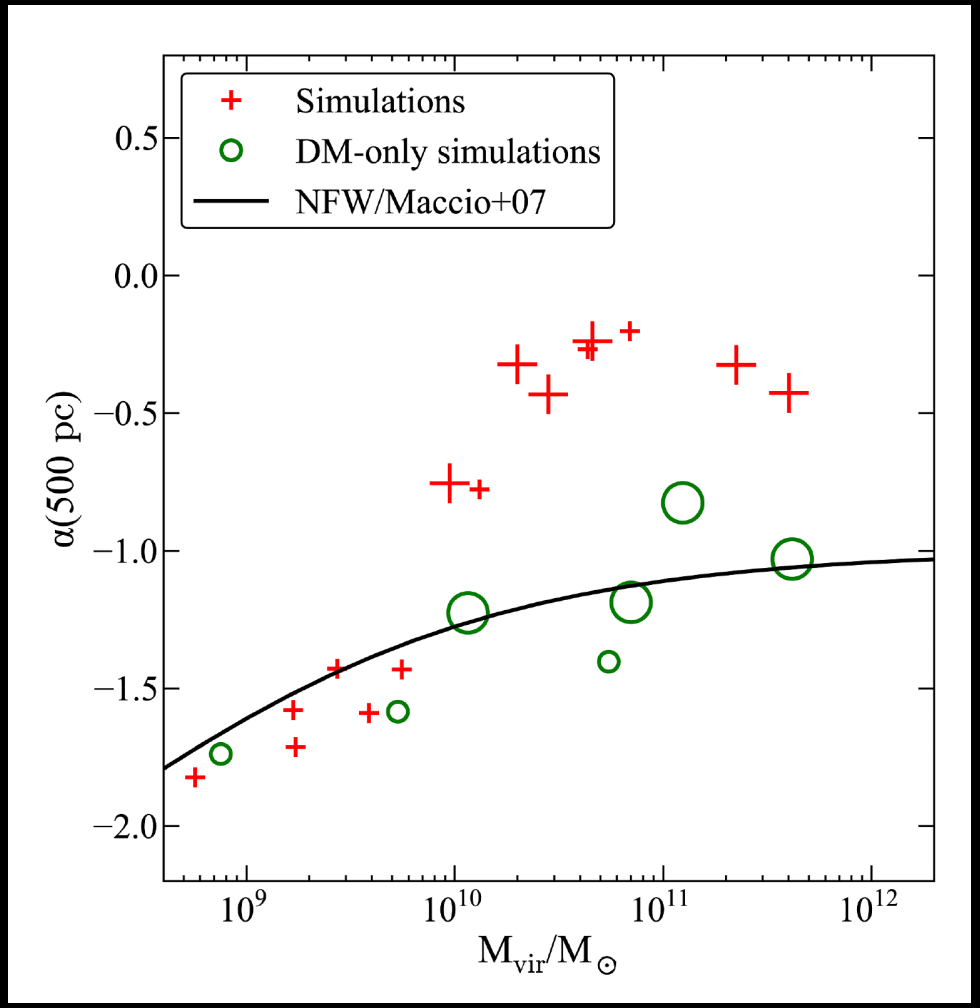
Governato+ (2012)- Inner slope of dark matter with and without hydro.

Cusped versus Cored Profiles in Dwarf Galaxies

Slopes of dark matter within 500 pc now shown as a function of halo mass.

The cutoff for cored profiles at $M_* < 10^7\ M_{sol}$ translates to a $M_{DM} < 10^{10}\ M_{sol}$. No observations are shown, but DM-only simulations show the relation expected fro NFW profiles (they are self-similar, but the divergence is because 500 pcs is larger fraction of the scale radius for a smaller halo).

The hydro simulations show that around $M_{DM} \sim 10^{10}\ M_{sol}$ baryons can shape a dark matter halo.

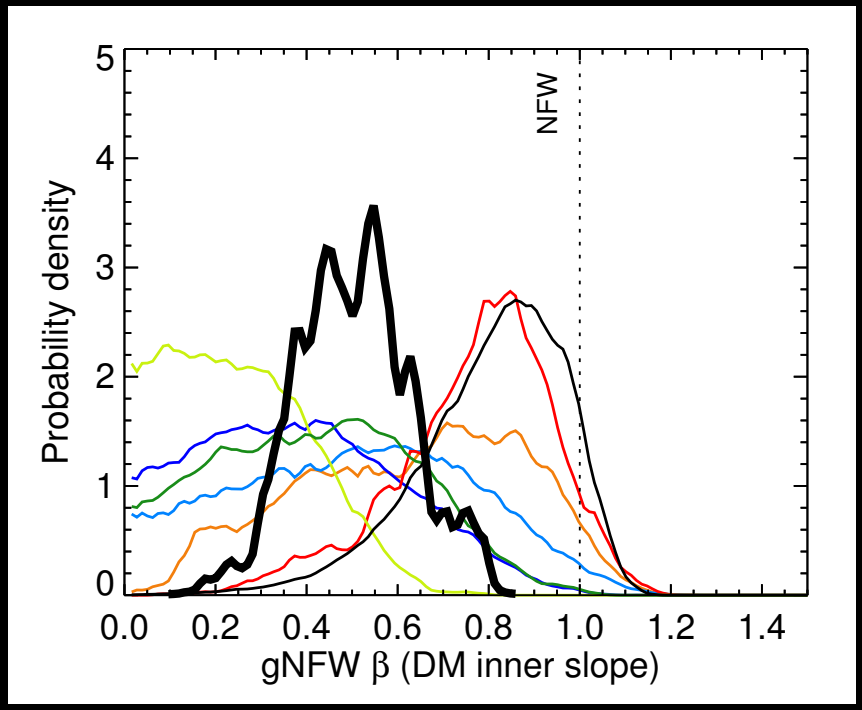


Governato+ (2012)- Inner slope of dark matter with and without hydro.

Cored Profiles in Brightest Cluster Galaxies

Clusters, where $M_{DM}>10^{14}$ M_{sol} are valuable laboratories for understanding dark matter profiles, because stars and hot diffuse gas (the intracluster medium) can be traced by observations. The profiles are consistent with an NFW profile of low concentration as expected in a cluster.

However, the inner profile traced by the brightest cluster galaxy is observed to be core-like on scales of ~10 kpc. Dynamical interactions, including merging galaxies with binary black holes have been invoked to flatten the inner mass profile.

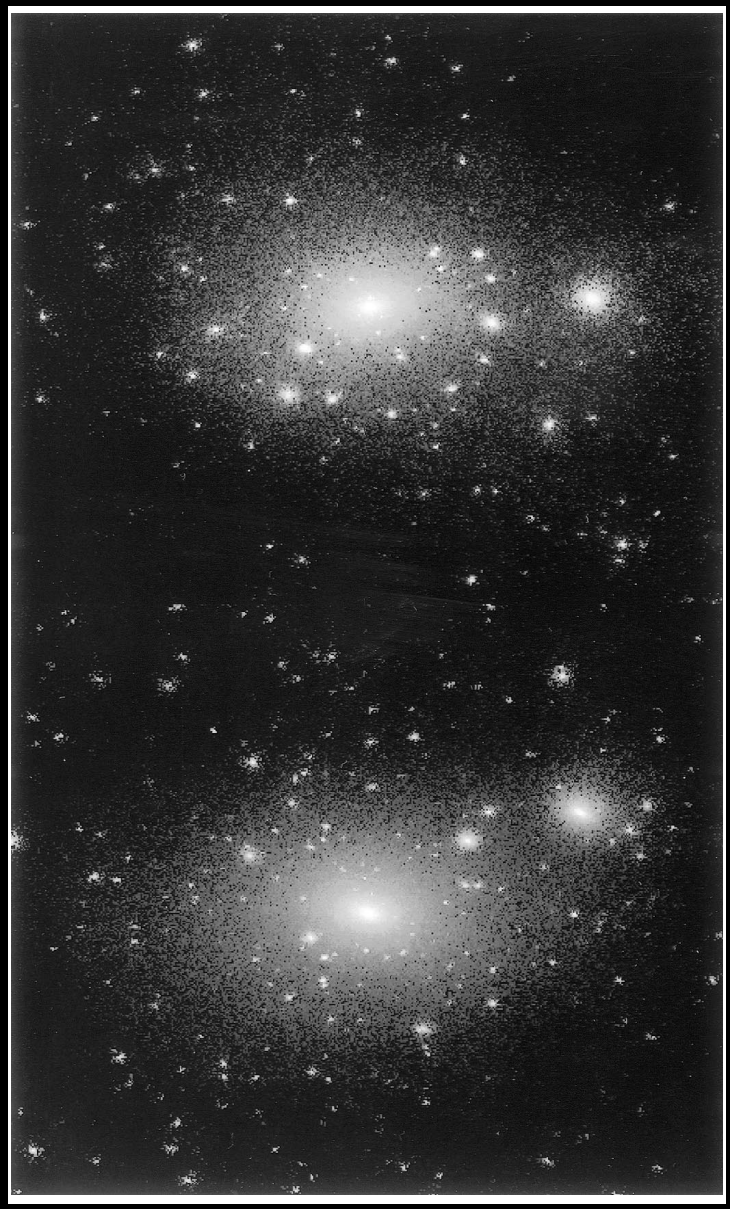


XXX…

The Missing Satellite Problem

Our own Milky Way halo is expected to have a bunch of DM subhalos that should be capable of hosting galaxies.

Moore+ (1999) pointed out that DM halos should be very nearly self-similar, i.e. a cluster-sized halo like the Virgo cluster (top) has the same substructures as a galaxy-sized halo like the Milky Way (bottom), but at a different mass scale (250x different) and a size scale (6x different).



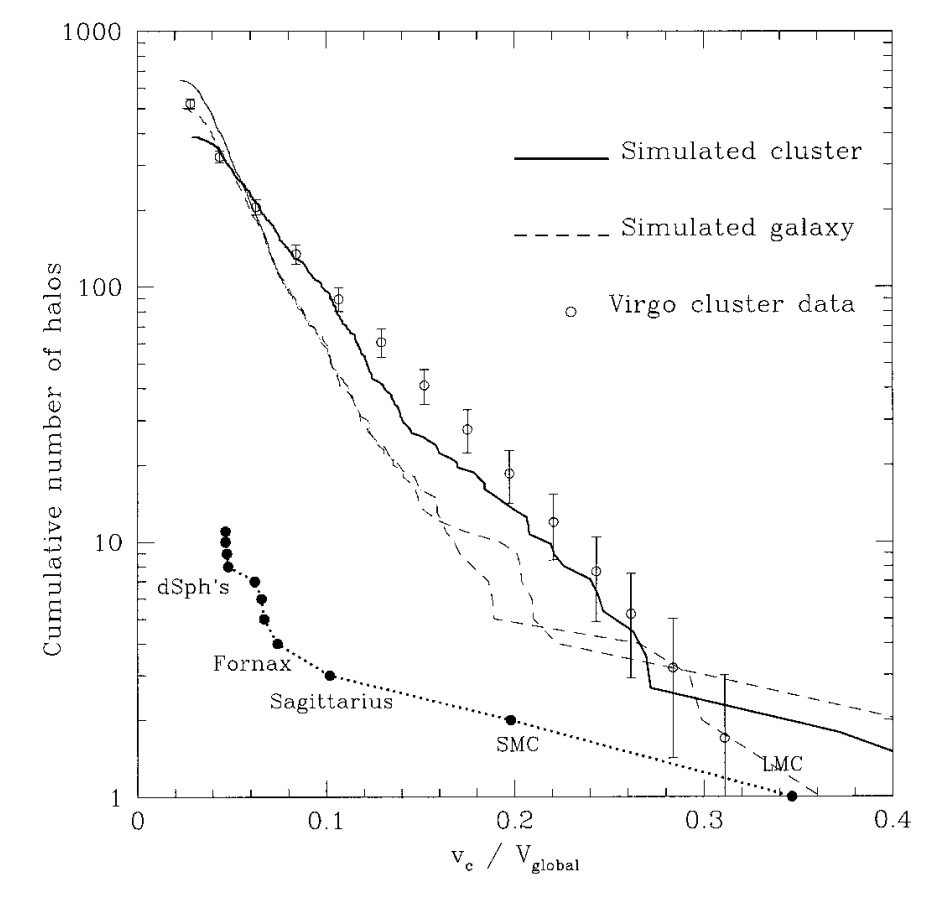
Moore+ (1999)- DM-only simulation of 5×10^{14} M_{sol} cluster and 2×10^{12} M_{sol} MW-sized halo

The Missing Satellite Problem

Instead, the Milky Way halo has many fewer DM halos with obvious galaxies in them compared to simulations. However, the cluster satellite mass function matches simulations.

There should be many more satellites observed in the Milky Way with higher velocity dispersions than observed. So it isn't just that there are fewer than expected satellites, but the observed satellites have lower velocity dispersions.

Does this suggest that there is a mass threshold below which DM halos cannot form? Or is it more complicated?



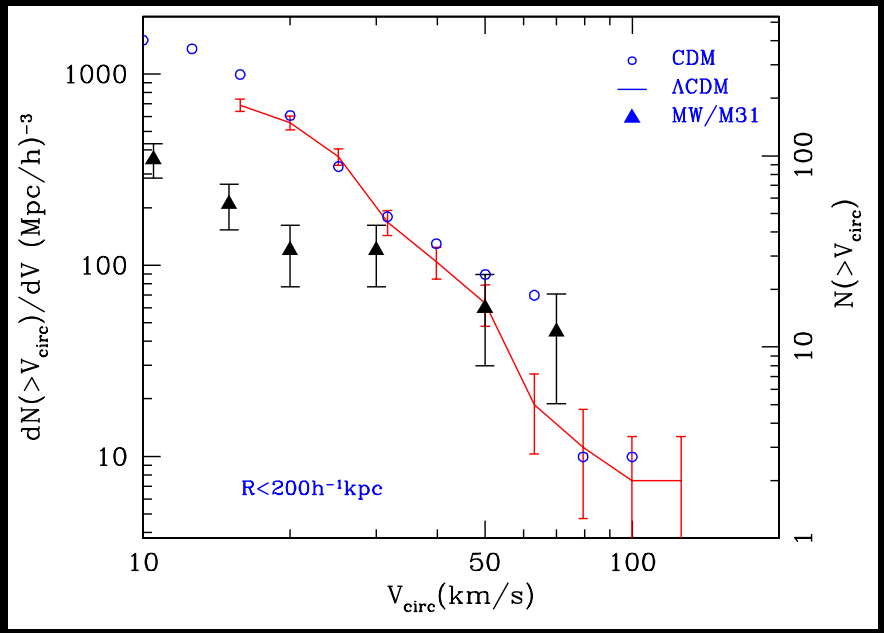
Moore+ (1999)- Cumulative number counts of satellite galaxies (observations as data points) and simulated DM halos (lines).

The Missing Satellite Problem

Klypin+ (1999) identified the same missing satellite problem at the same time.

They showed that while the number density of galaxies with v_{circ} >=50 km/s agrees with ΛCDM, there are up to 10x too few galaxies below according to DM-only simulations.

Klypin also considered the Andromeda halo, and showed that this problem was not unique to our MW halo only. The same missing satellite problem was observed there.



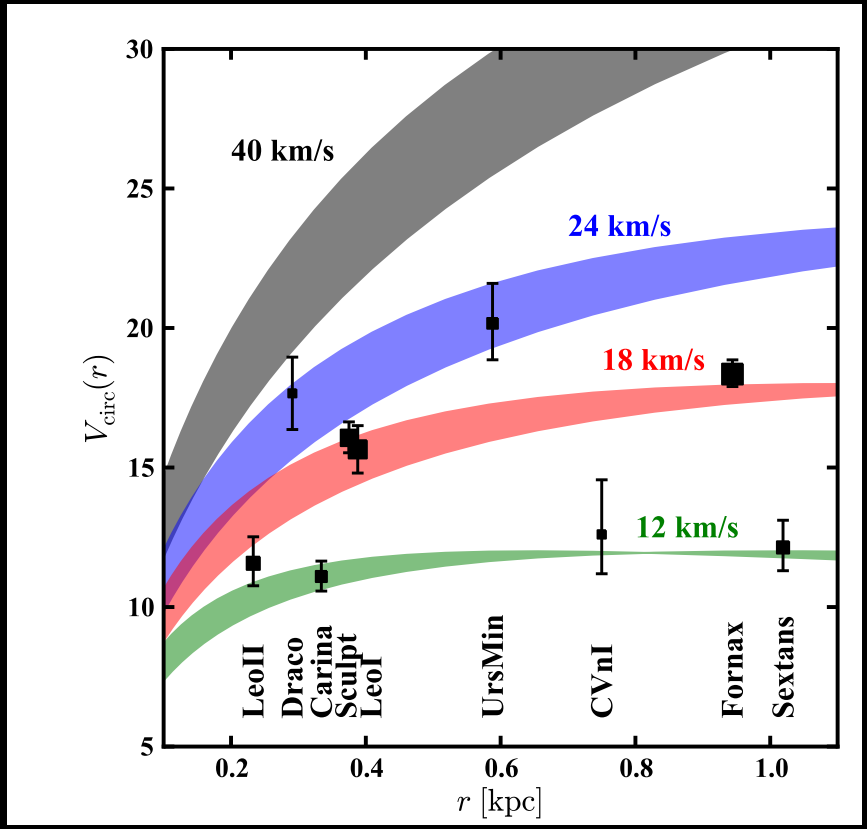
Klypin+ (1999)- simulations (in blue and red) compared to observed satellites (black triangles).

Too Big To Fail- velocity dispersions of observed Milky Way satellites appear too low.

Boylan-Kolchin+ (2011,2012) showed the derived velocity dispersion of Milky Way dwarf spheroidal galaxies are v_{max}= 12-25 km/s, but ΛCDM cosmology predicts >10 subhalos with v_{max}>25 km/s. While a solution to this problem is to say that subhalos with v_{max}>25 km/s don't form stars, they are "too big to fail." Why would these massive subhalos fail, but the lower mass halos still form stars?

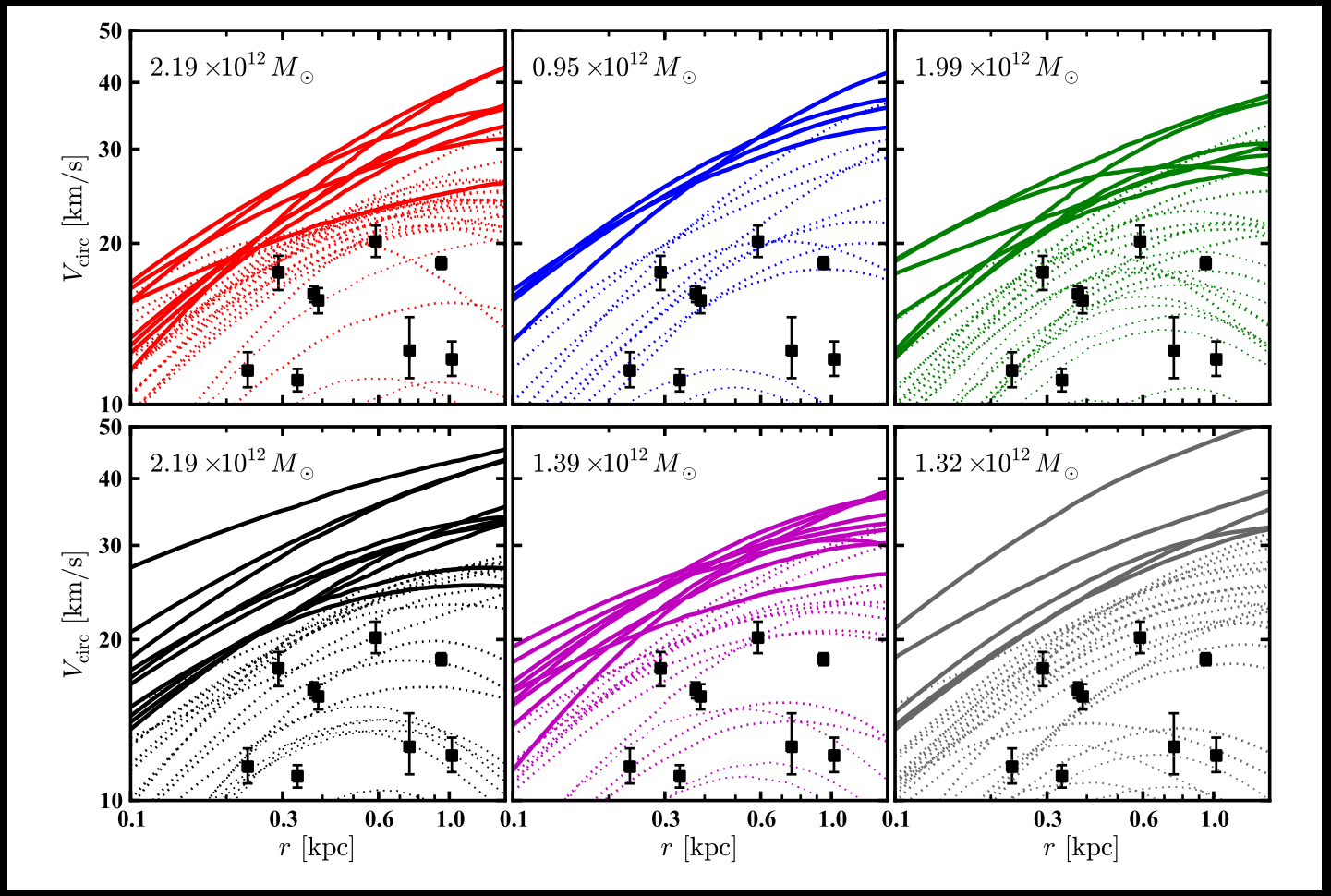
This research uses the DM-only Aquarius simulation with up to 120 million DM particles per MW-mass halo (1-2x10^{12} M_{sol}) with resolution 2-20x$10^3 M_{sol}$ and a softening length of 66 pc.

Can baryonic physics change velocity dispersions of DM halos like this? Boylan-Kolchin (2012) argue no.



Boylan-Kolchin+ (2012)- Observed v_{max} shown ias black data points, simulated velocity profiles shown as bands.

Too Big To Fail

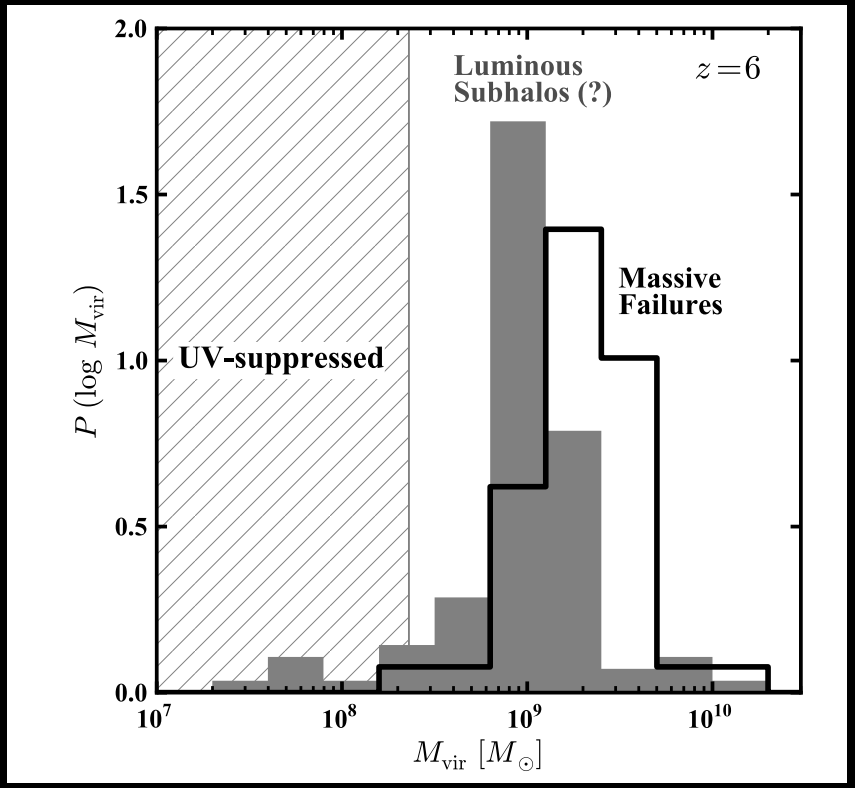


Boylan-Kolchin+ (2012)- Even the lowest mass Aquarius simulations (0.95x10^{12} M_{sol}) have velocity profiles as a function of radius that do not agree with the observations.

Too Big To Fail

One idea is that UV photo-heating boiling baryons out of halos can alter these dwarf subhalos, but UV suppression occurs at a too low of a subhalo mass to solve the too big to fail problem.

Even at z=6, the progenitor subhalos have grown larger and have higher velocity dispersions than the halos that are too big to fail. So this problem is in place during the reionization era (z=6-15), and cannot be solved with the subsequent 12-13 Gyrs of galaxy formation, so argues Boylan-Kolchin.

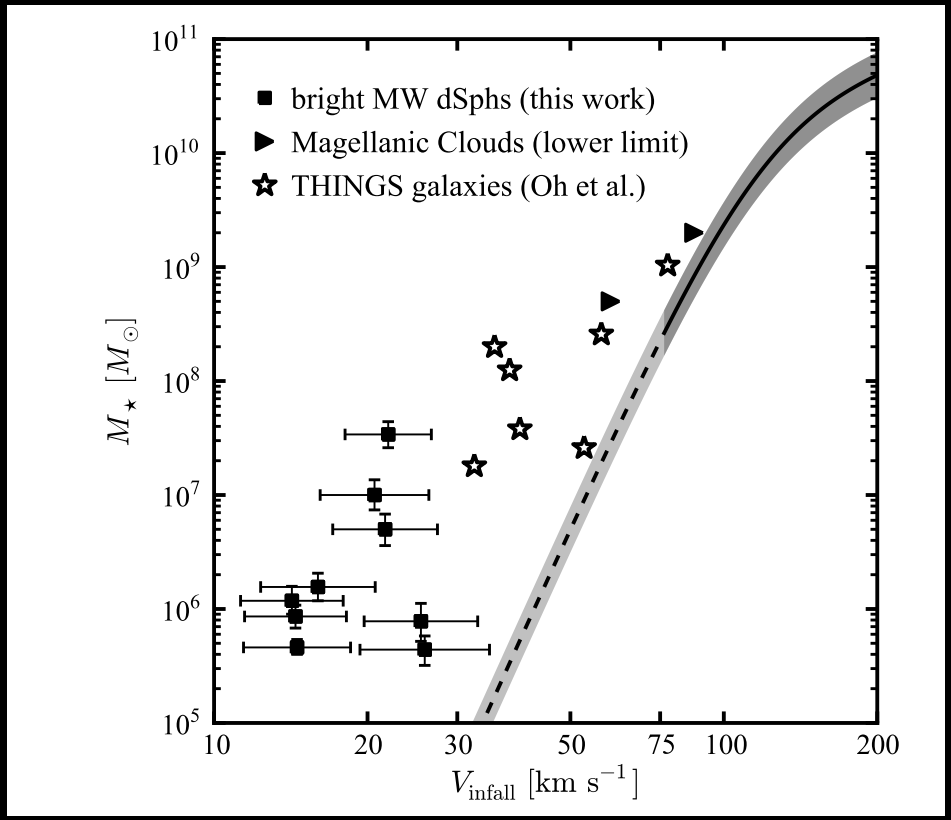


Boylan-Kolchin+ (2012)- Simulated subhalos even at z=6 have formed already to be too massive and have too high of velocity dispersions compared to MW dwarf galaxies observed today.

Too Big To Fail

Abundance matching constraints also show that MW dwarfs and other dwarf galaxies (THINGS survey by Oh et al. 2011) are too massive given their derived velocity dispersions.

The abundance matching constraints are from DM-only simulations, so this suggests that something else needs to change the the velocity dispersions of dark matter, whether it be baryonic physics or a different type of dark matter than cold dark matter.



Boylan-Kolchin+ (2012)- Mass and v_{max} of observed galaxies compared to abundance matching constraints (Behroozi+ 2010) extrapolated to lower mass halos.

Cold, Warm, & Hot Dark Matter

The definitions of dark matter are based on the masses of dark matter particles, which is unknown. The smaller the mass, given in terms of energies, i.e. m/c^2, then the hotter it is, because the particles have a larger velocity an can "free stream" longer distances.

When particles free stream further, they wipe out primordial density fluctuations of the power spectrum of the Universe on small scales.

-**Hot dark matter (HDM)**- free streaming much larger than an early proto-galaxy. Could be relativistic neutrinos with masses 10-100 eV.

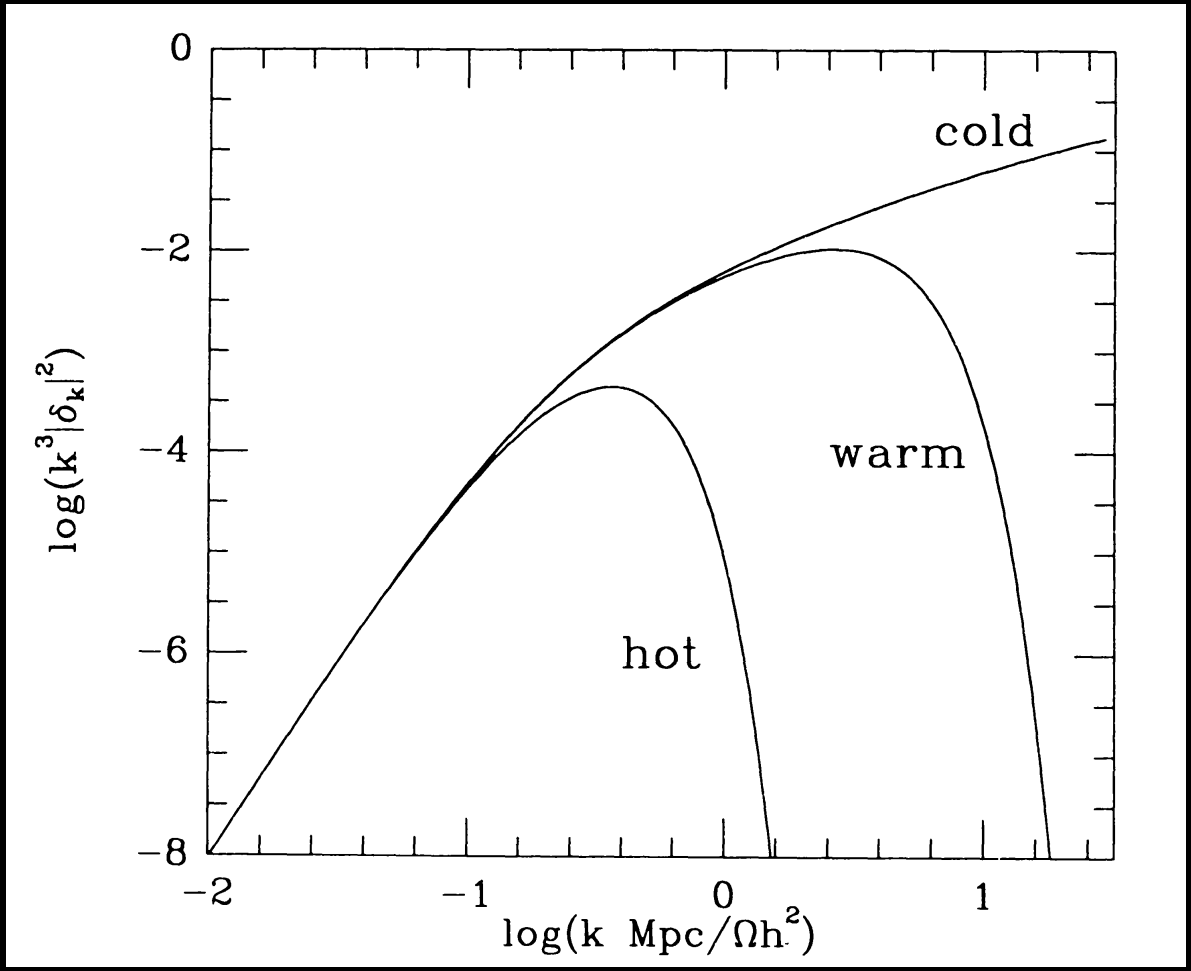
-**Warm dark matter (WDM)**- free streaming is smaller than HDM, but on the length scale of a dwarf galaxy, so that it could wipe out the fluctuations responsible for forming dwarf galaxies (i.e. the fact that few dwarf galaxies are observed around the Milky Way). Masses 0.3-3.0 keV are standard.

-**Cold dark matter (CDM)**- free streaming length is smaller than perturbations responsible for the formation of even dwarf galaxies. A heavier particle is required that is never relativistic. Masses easily exceed a few keV and could be much, much larger- an area of active research in particle physics and super-symmetric string theory.

Cold, Warm, & Hot Dark Matter

The effect of non-cold dark matter is to cut off the power spectrum at small scale (large k's).

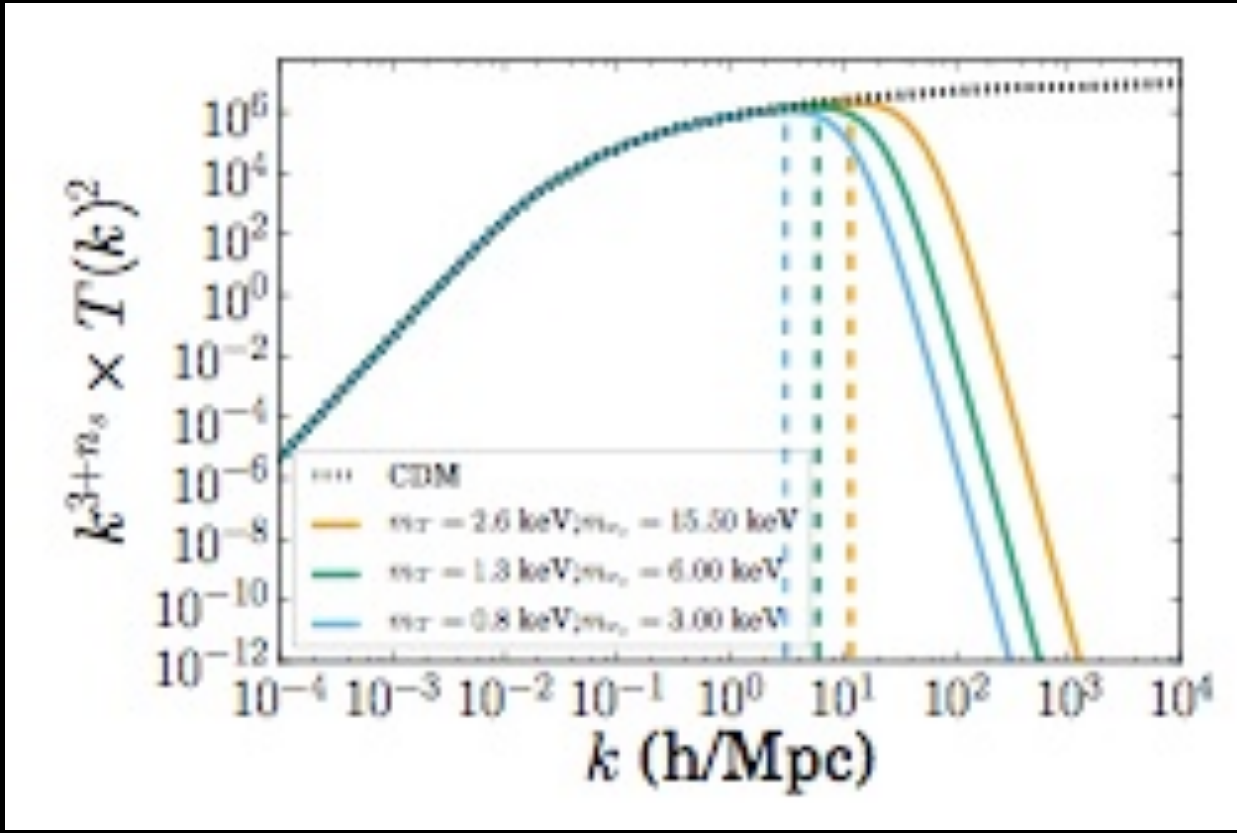
The nature of warm or hot dark matter is to severely depress fluctuations at small scales according to the mass-energy of the dark matter particle.



Power spectra for different types of dark matter.

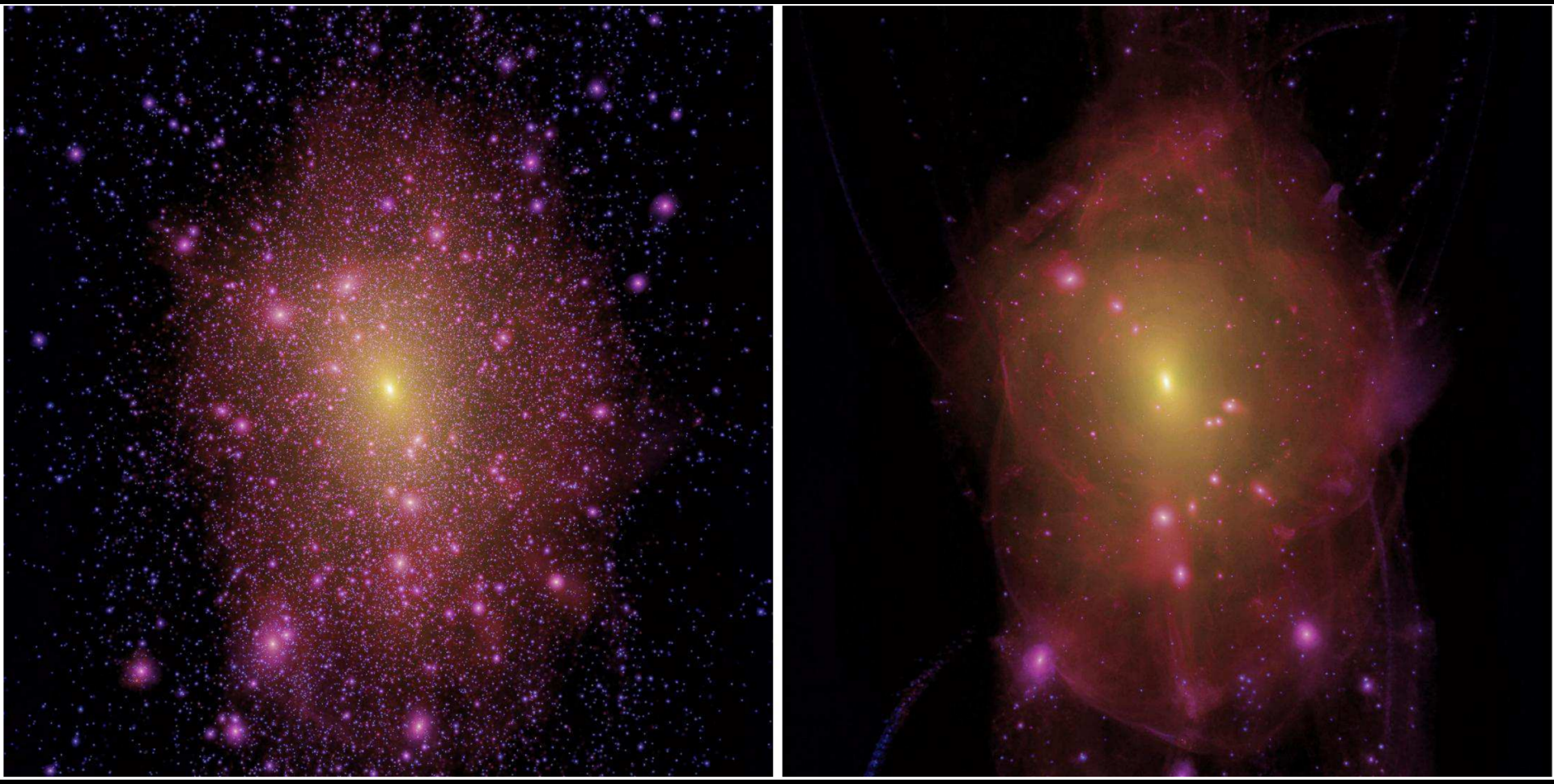
Warm Dark Matter

Typical energies of WDM explored, 0.5 keV to 3 keV are in the allowed range where larger galaxies can form, but change the formation of low mass galaxies.



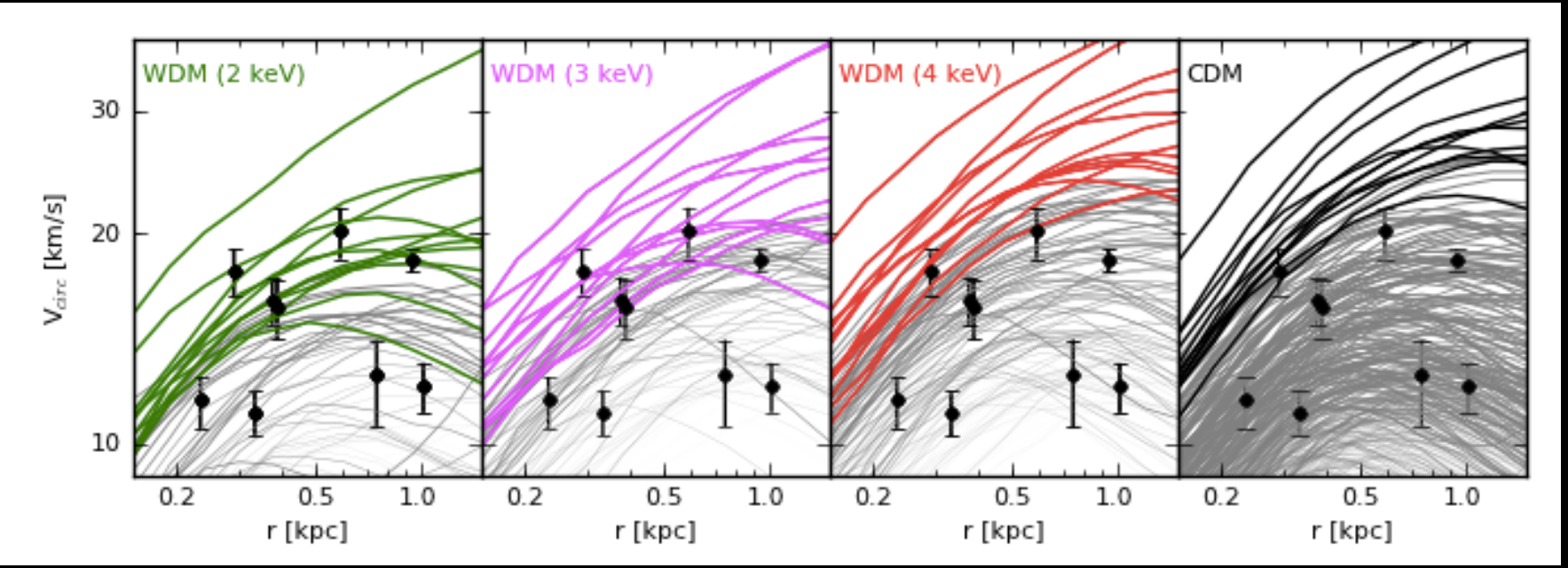
Schultz+ (2014) Power spectra from the initial conditions of a simulation. The thermal energies are listed, but this paper also quotes sterile neutrino masses, which have different energies, but behave the same way.

Warm Dark Matter



Lovell+ (2012) simulation with CDM (left) and WDM (right- 2 keV sterile neutrino). Small halos disappear almost completely.

WDM and Too Big to Fail



Schneider+ (2014)- Can too big to fail be solved by warm dark matter? They indicate no, and that for even 2 keV WDM models, there exist too many subhalos with v_{circ}=12 km/s, although not as many as in CDM.

High-z Galaxy Counts- Another work (Schultz+ 2014) looked at high-z galaxy counts in simulations and showed that reionization-era galaxy counts (z>=6) can constrain the WDM energy, because the small halos hosting early galaxies will be suppressed by WDM. Should provide great constraints with James Webb Space Telescope.

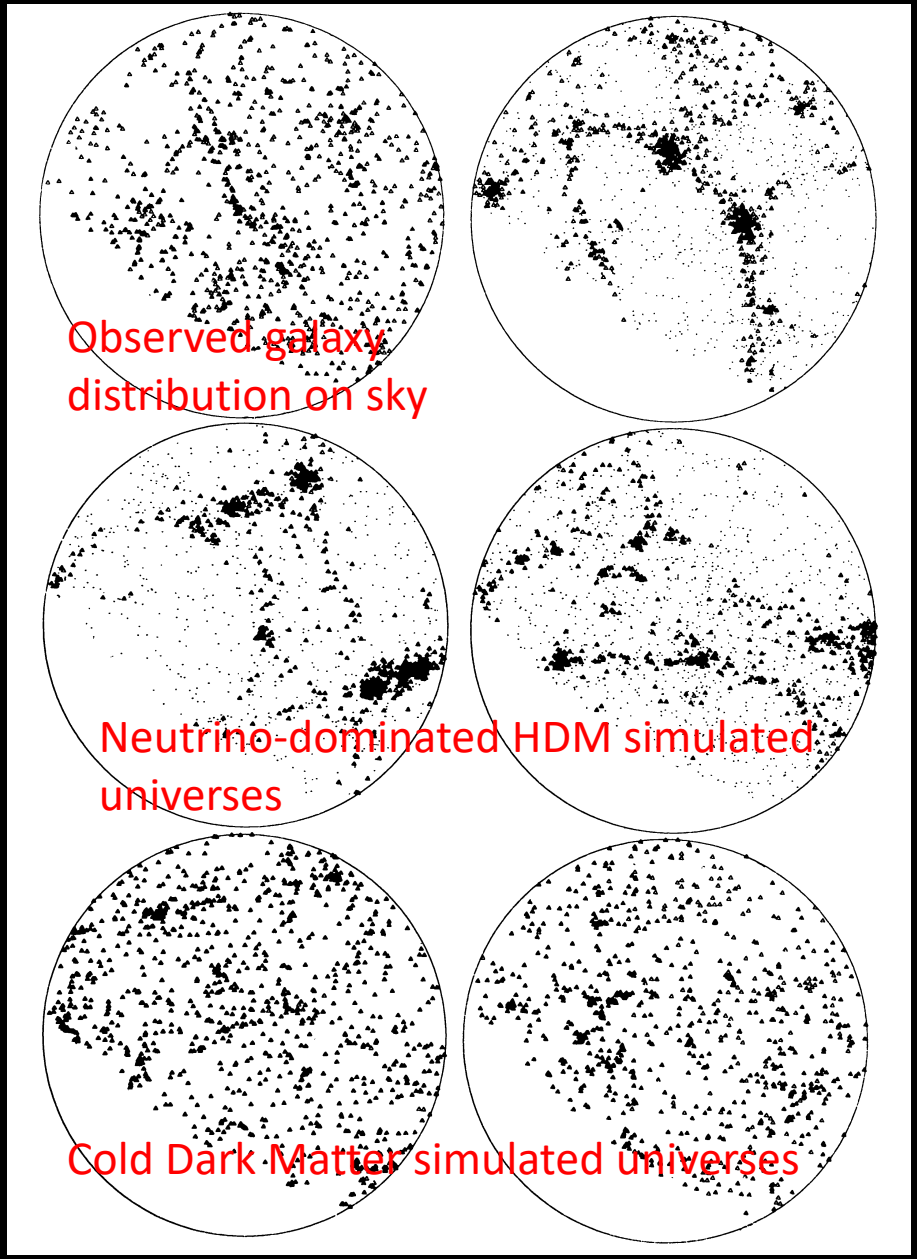
Hot Dark Matter is Dead

Rudimentary simulations (by today's standards) showed the distribution of galaxies would be very different in a HDM Universe where fast-streaming neutrinos were dark matter.

Structures on scales larger than a galaxy would be wiped out in the early Universe, and galaxies could only form in very large collapsed structures, which disagree with observations of the first modern galaxy surveys done in the 1980's (the CfA survey).

Cold Dark Matter- hierarchical structure formation, small things form first, and then grow into big things.

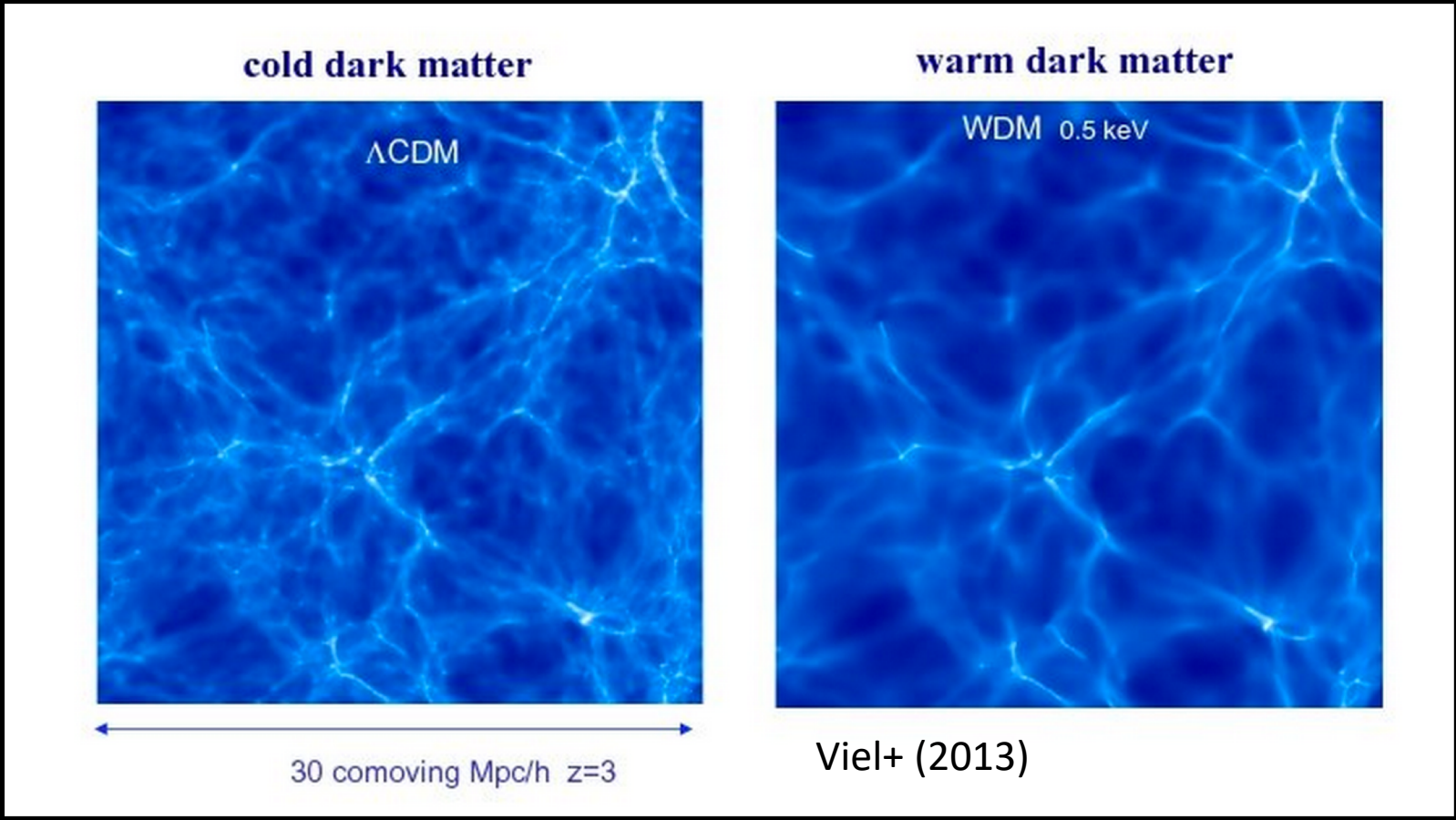
Hot Dark Matter- large structures have to collapse first and then galaxies form.



White (1987)

Warm Dark Matter- The Lyman-α Forest

The Lyman-alpha forest, which is traced by the cosmic web, appears significantly smoothed out if WDM of energy 0.5 keV is used. This is observable by using the power spectrum of Lyman-α Absorbers.

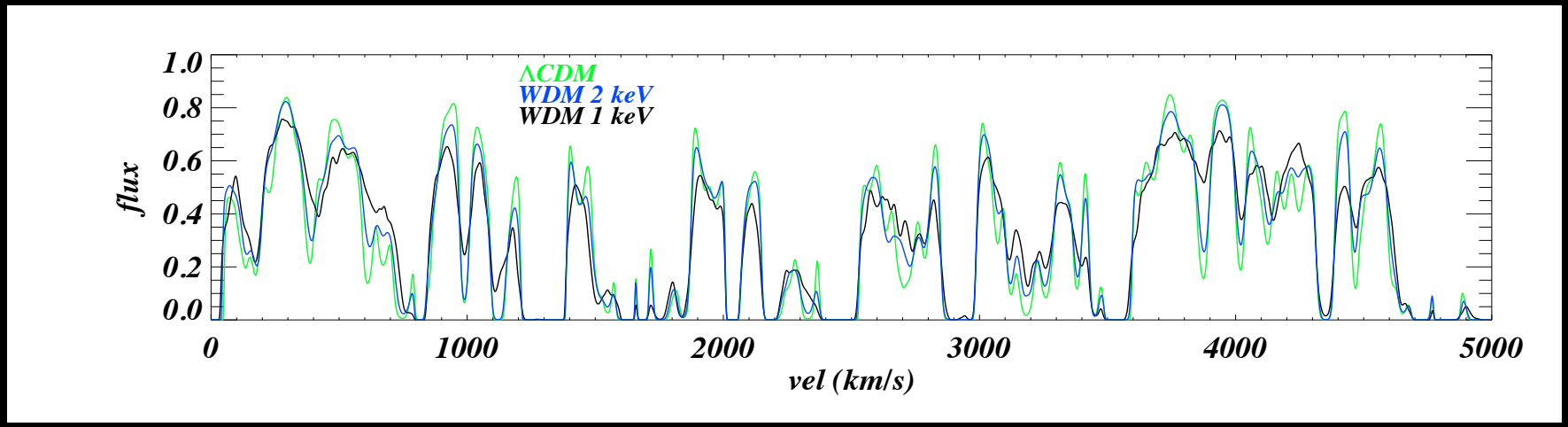


Viel+ (2013)

Warm Dark Matter- The Lyman-α Forest

Using a quasar as a background source, absorption line spectra show Lyman-α forest absorption tracing the cosmic web structure- HI absorption lines correspond to filaments in the large scale structure, while voids have less absorption. This is the best observable of the nature of the intergalactic medium, and its clustering at z>4 is sensitive to the smoothing owing to warm dark matter.

Cosmological hydrodynamic simulations are very good at simulating the details of the Lyman-α forest .



Viel+ (2013)- Simulated z=4.6 Lyman-α forest spectra for CDM and two WDM temperatures, which smooth out the observed forest.

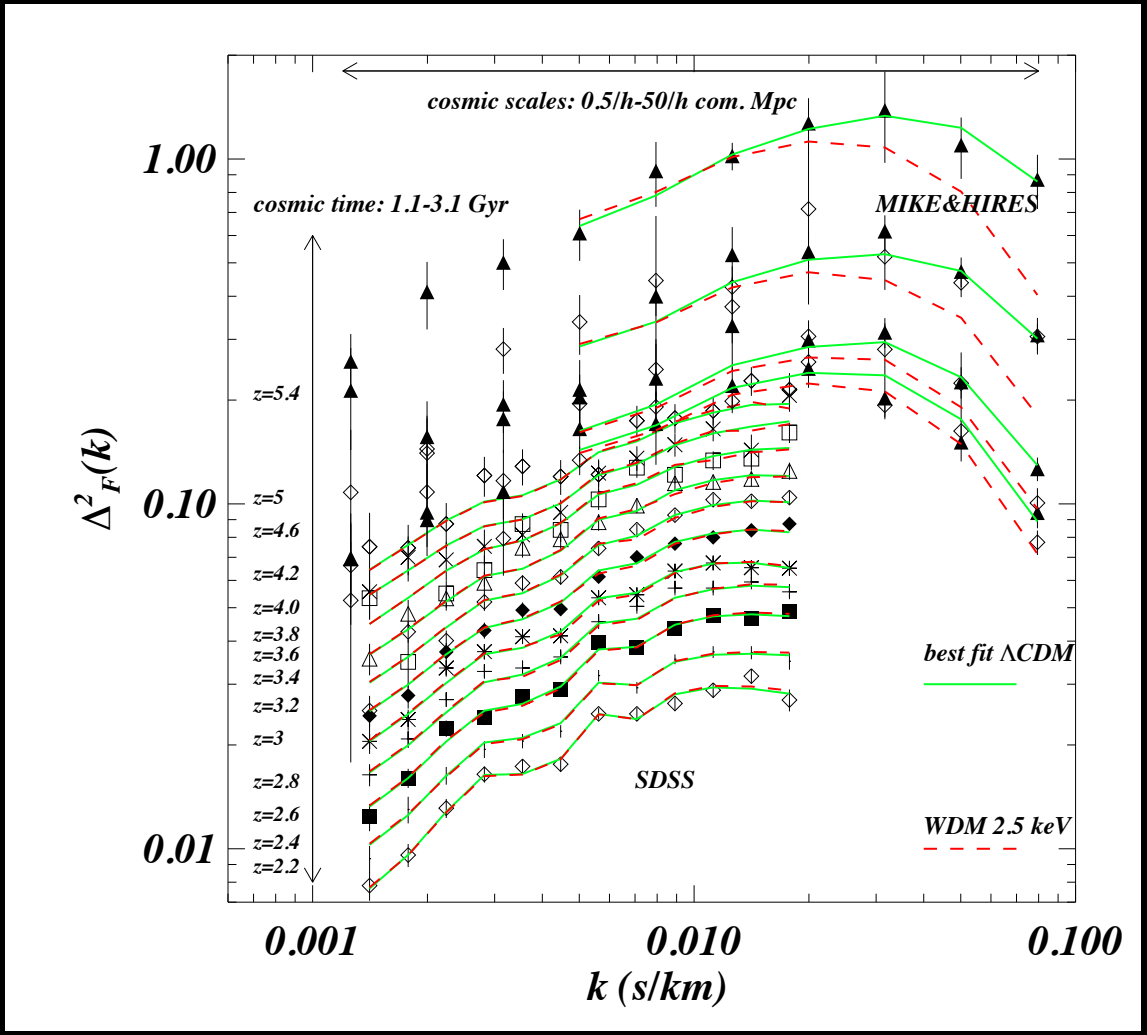
Warm Dark Matter- The Lyman-α Forest

The Lyman-α power spectrum is a powerful probe of the effect of WDM.

High-resolution spectra capable of resolving <10 km/s are required at z>4 where WDM can affect the distribution of the observed gas.

In Viel+ (2013), WDM with particle mass of 2.5 keV can easily be ruled out.

This is a strong argument against WDM, but it can be complicated by the nature of the Lyman-α forest (e.g. temperature broadening, Hubble broadening, the strength of the ionizing background).



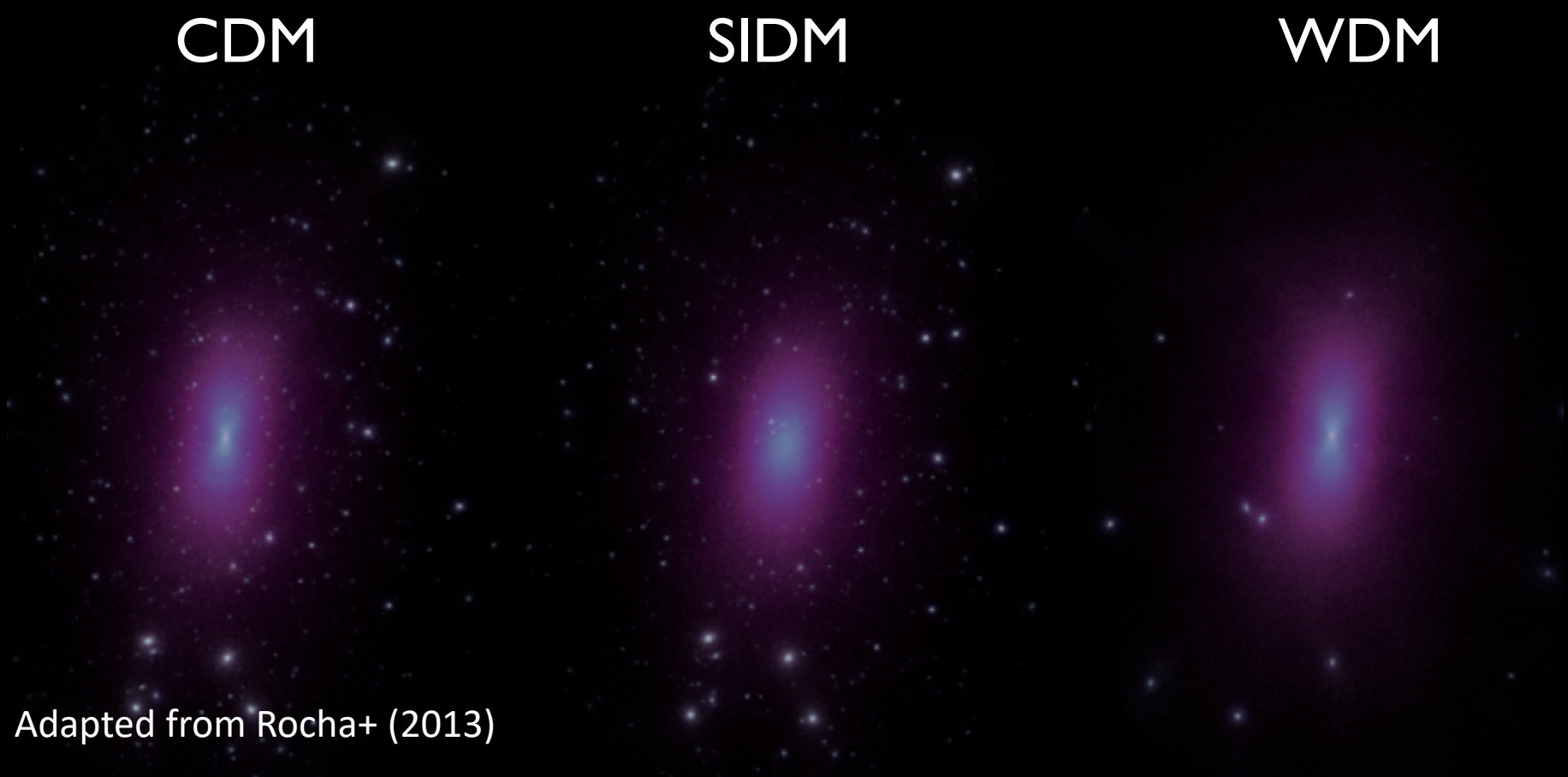
Viel+ (2013)

Self-Interacting Dark Matter

Spergel & Steinhardt (2000) proposed that dark matter that interacts with only itself (and not baryons) as a way to make cored profiles.

CDM vs. WDM vs. SIDM

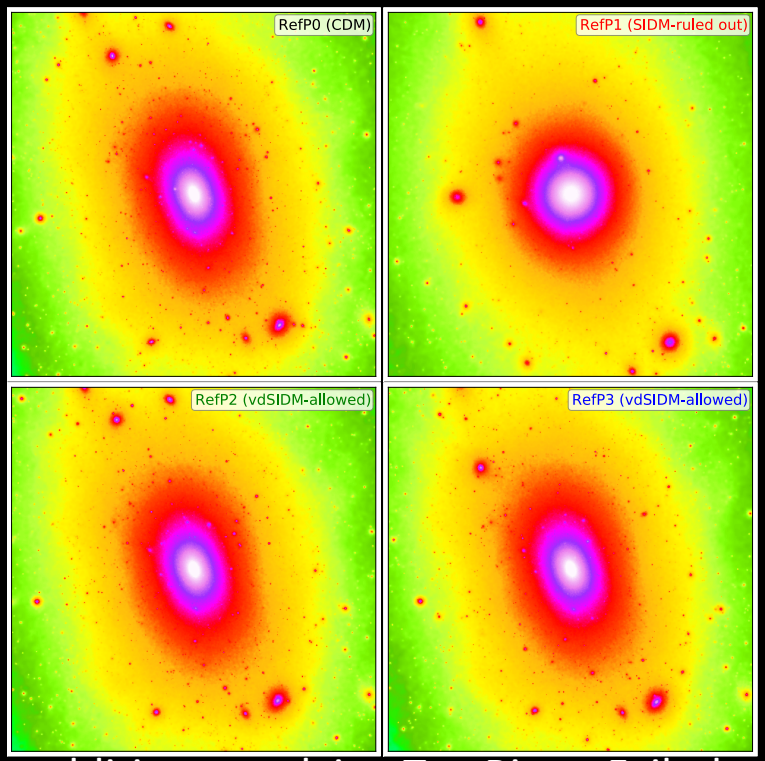
Unlike WDM, SIDM preserves subhalo number counts, but create cores in subhalos, because the DM collisions (several per particle over a Hubble time in cores) isotropize orbits and create cores that tend toward sphericity as opposed to triaxiality.



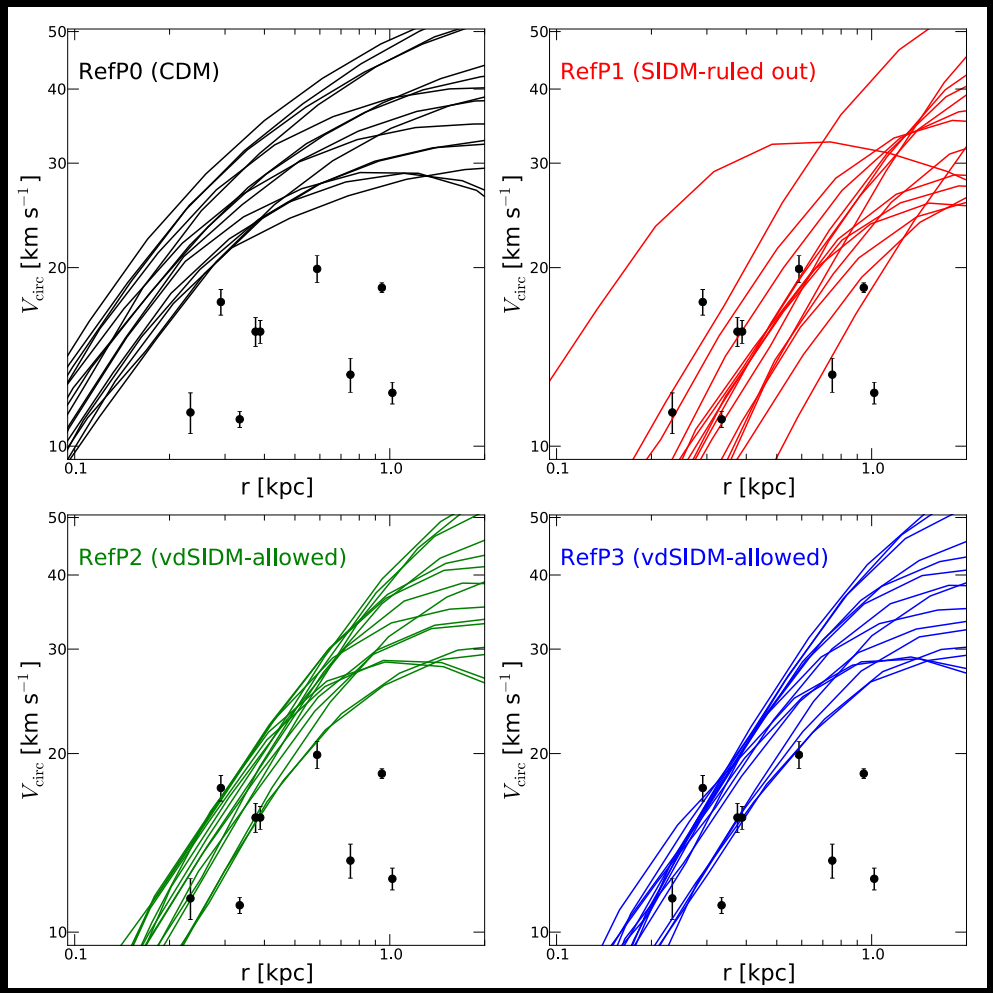
Adapted from Rocha+ (2013)

Too Big to Fail and SIDM

Aquarius simulations (Vogelsberger+ 2012) run with SIDM are able to reduce the velocity dispersions of the most massive observed subhalos.



In addition to solving Too Big to Fail, the central DM halo profile can maintain some ellipticity, which is required by observations of clusters.



Vogelsberger+ (2012)- SIDM simulations are able to match the velocity dispersions of the biggest subhalos.

The Void Phenomenon as a Challenge to ΛCDM Cosmology

Peebles (2001) put forth the challenge that voids of the scale 15 h^{-1} Mpc and larger should be a natural place for low-mass galaxies to form. ΛCDM he argued expects that voids should preferentially hold lower-mass halos that form low-mass, dwarf galaxies.

The challenge to this picture is that there is an abrupt transition from filaments to voids, where galaxy density falls off dramatically, and the rare galaxies that are observed in voids are a mix of L* and dwarf galaxies, not preferentially dwarf galaxies.

The void phenomenon was termed a crisis for ΛCDM, and a great test because statistics of void galaxies and shapes of voids themselves are sensitive to Ω_M, dark energy, modified gravity, and the nature of DM.

The trouble with simulations however is that they have to be large enough to form voids (a large box) with enough resolution to form galaxies in low-mass halos. This in fact is very expensive and usually requires DM-only simulations with halo occupation statistics (i.e. galaxies "painted" onto halos.

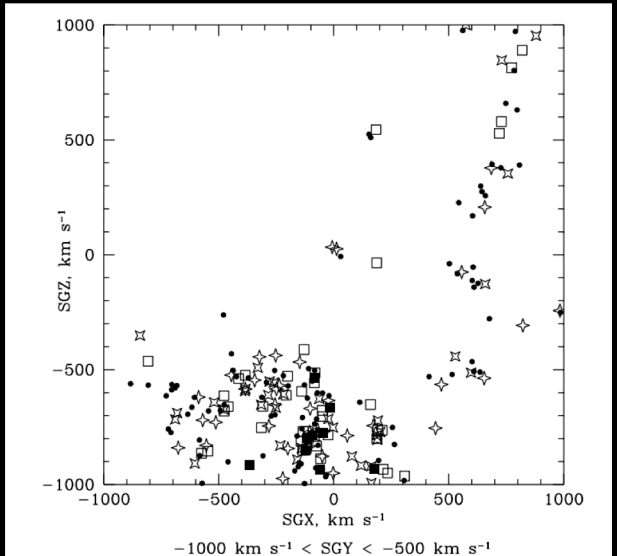
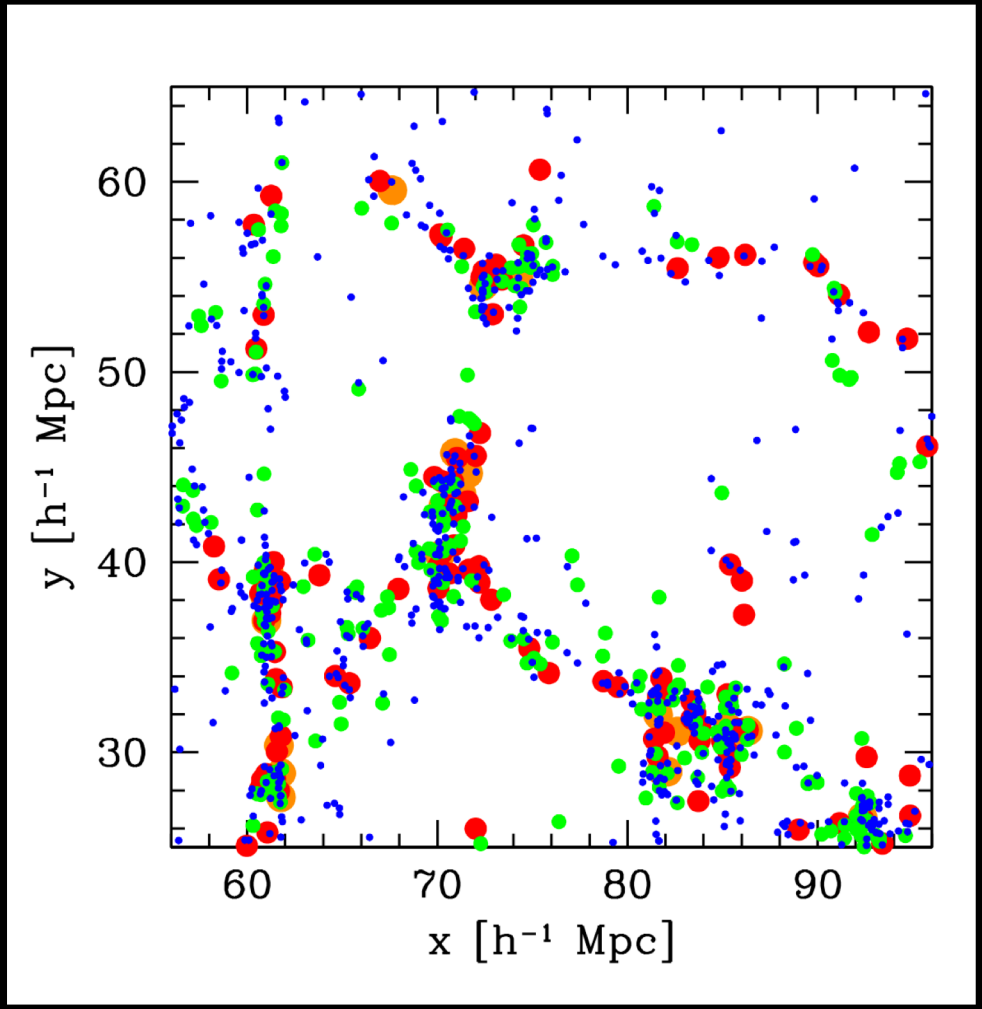


FIG. 2.—Map of ORS galaxies in a slice in redshift space. The normal to the slice points in the direction of the Virgo Cluster, from the opposite side of the Milky Way. Filled squares are elliptical galaxies, open squares are S0 galaxies, crosses are Sa–Sc galaxies, plus signs are later spiral types, and filled circles are dwarfs and irregulars.

Peebles (2001)- galaxies observed around local void

The Void Phenomenon as a Challenge to ΛCDM Cosmology

Tinker & Conroy (2009) used a halo occupancy distribution (HOD) based only on halo mass to populate halos in a large ΛCDM N-body simulation, and were able to reproduce the key observations of the void phenomenon. Low-mass halos exist in voids, but the luminosity of galaxies in the low-mass halos declines so quickly that you don't see a larger population of dwarfs. Galaxies populate filamentary walls at the edges of voids with all types of galaxies.



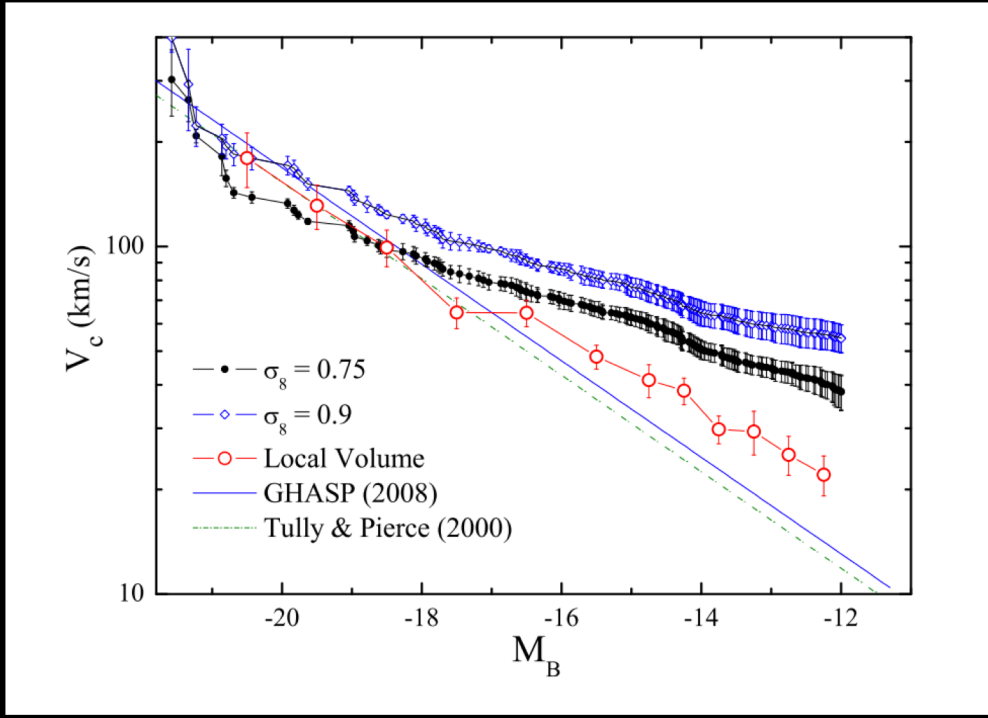
Tinker+ (2009)- simulated galaxy map

The Void Phenomenon as a Challenge to ΛCDM Cosmology

Tikhonov & Klypin (2009) look at the void problem in the local volume (<8 Mpc away) and search for very faint dwarfs tracing halos that should form in small local voids, and find ΛCDM over-predicts 10-fold the number of observed faint dwarfs.

This relies on studying ultra-faint galaxies in the local group, and finding velocity dispersions as a function of magnitude down to M_B=-12. Observations find that the velocity dispersion is too low for a given luminosity, and that these ultra-faint galaxies should live in more massive halos.

Are halos below v<35 km/s not forming stars in voids by some other process (photo-heating)? Or is there a process reducing the velocity dispersion of void halos, similar to Local Group halos in the Too Big to Fail problem? See also Gottlober+ (2003).



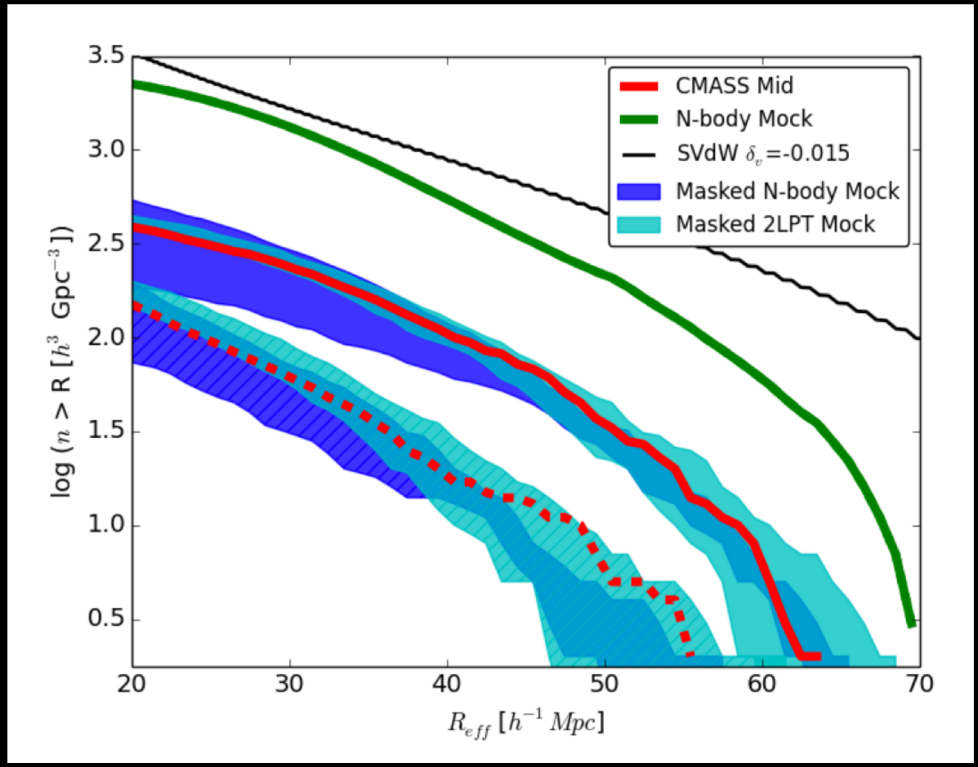
Tikhonov+ (2009)- ΛCDM sims. Are blue and black. Observations are red.

Void Properties as Tests of ΛCDM

The number density, sizes, and even ellipticities of voids are predictions of N-body simulations that can be tested against observations. P.M. Sutter et al. (2014) is one example of making a void catalogue from the Sloan Digital Sky Survey and comparing to simulated catalogues from N-body simulations.

Recent work shows there is apparently good agreement with ΛCDM observations.

See also Sheth & van de Weygaert (2004) for how voids sizes evolve and create a size distribution peaked at a characteristic scale- large voids grow hierarchically, while small voids become swallowed leading to a preferred void size.



Sutter+ (2014)- # density of observed voids (red) by radius versus simulations processed to match data (blue, cyan).