Modelling the evolution of luminous matter in the obscured universe

Gian Luigi Granato

Istituto **N**azionale di **A**stro**F**isica

INAF

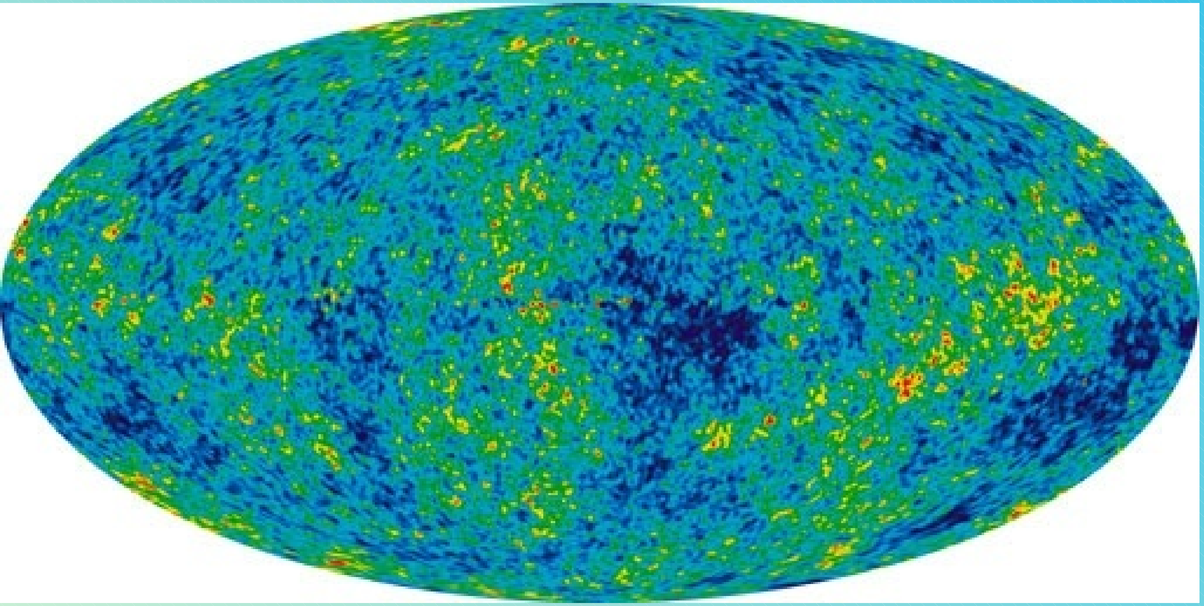
(Osservatorio Astronomico di Trieste)

Plan:

- Basic questions
- Dark matter evolution (N-body)
- Hydro-dynamic simulations
- Semi Analytic Models

A few basic questions and facts to reproduce and understand

Cosmic Microwave Background fluctuations



Galaxies

?

How and when did the galaxies that we see in the present-day Universe form form starting from the primordial density fluctuations ?

Why (at least) 2 very distinct types of Galaxies (or components of), with a clear mass trend (more mass-more spheroid):

Ellipticals (spheroids)
No gas, no star formation
No rotation
$I(r) / \exp[-7.67(r/r_e)^{1/4}]$
L / σ^4 (Faber-Jackson)

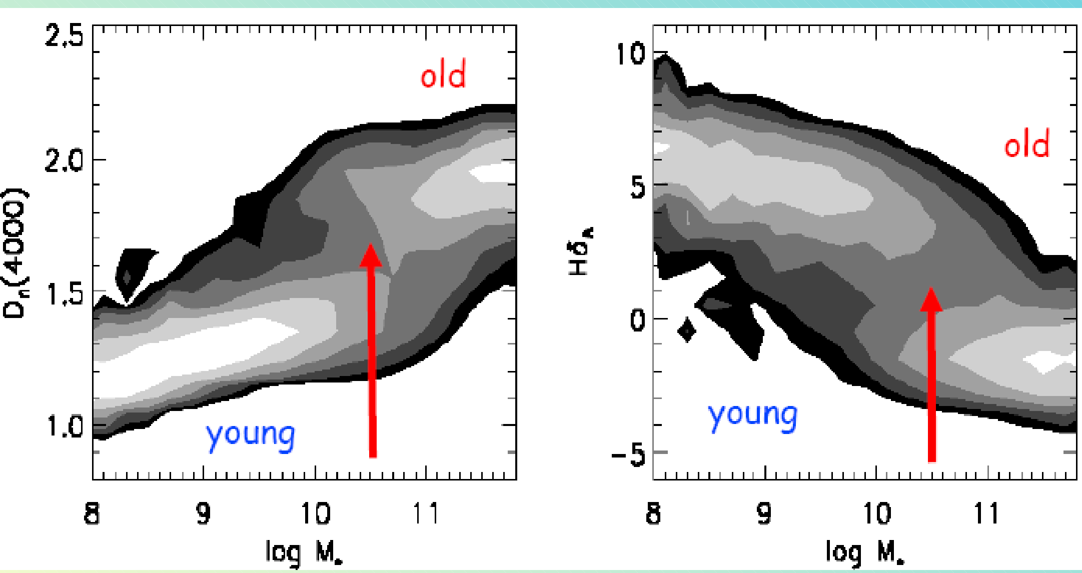
Spirals (disks)
gas, star formation,
rotation
$I(r) / \exp(-r/r_d)$
L / V_c^3 (Tully-Fisher)

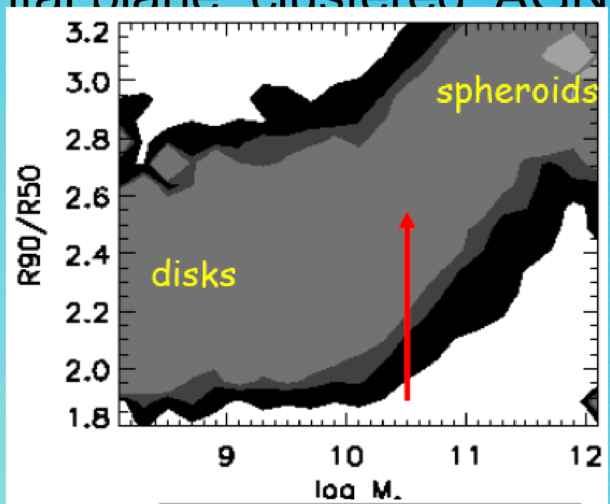
M_{SMBH} / M_{SPH} !! more than 10 orders of magnitude in size

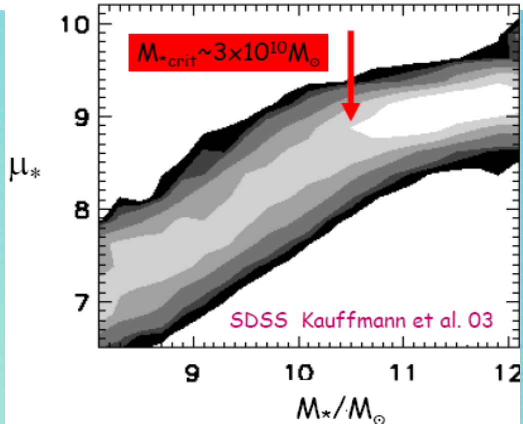
SDSS data suggest a bimodality or dichotomy in the distribution of properties in 10^5 local galaxies (e.g. Kauffmann et al 03).

• for M_* . $3x10^{10}$ M– tend to dominate (concentrations typical of) disk, blue, star-forming, low Z, LSB, M/L/ M, fundamental line, field

• for M_* & $3x10^{10}$ M– tend to dominate (concentrations typical of) spheroids, red old-pop, high Z, HSB, M/L / M^{-1}, fundamental plane, clustered, AGNs





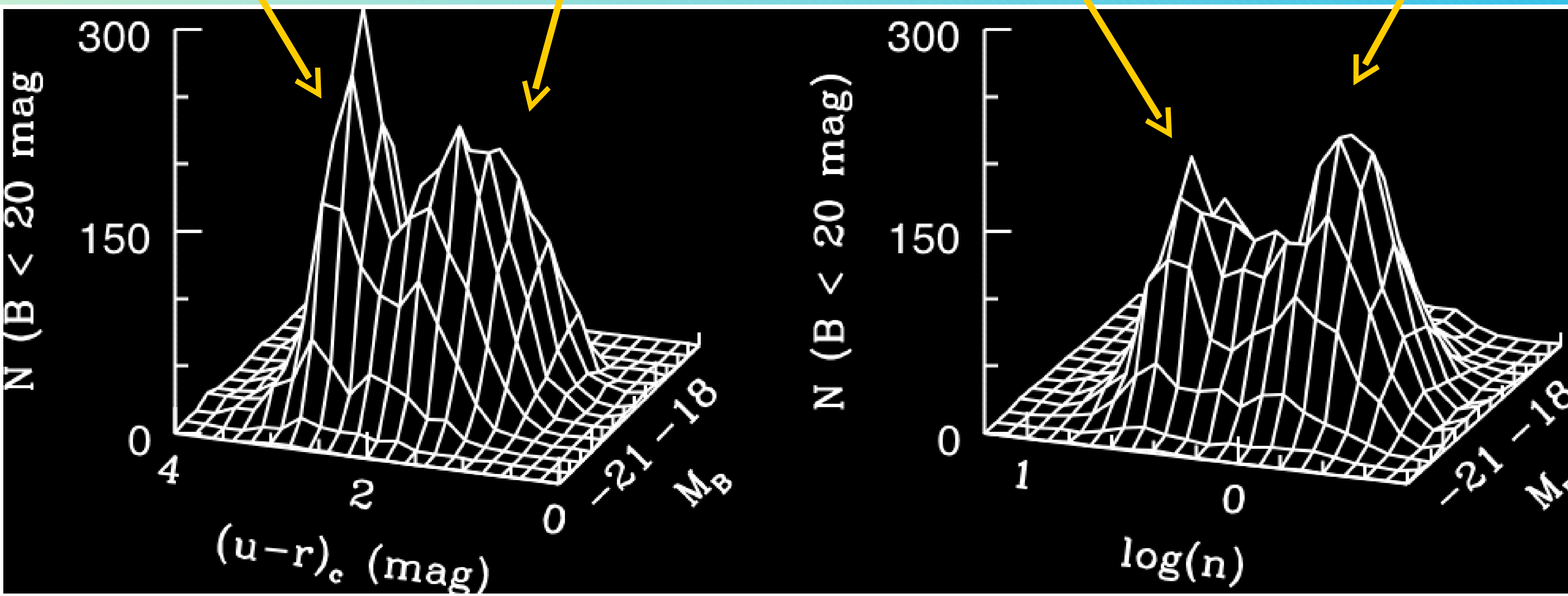


Observed Galaxy Bimodality (Millennium Galaxy catalog)

Observe strong colour *(u-r)* and structural (log*n*) bimodalities (Strateva et al 2001; Baldry et al 2004; Driver et al 2006)

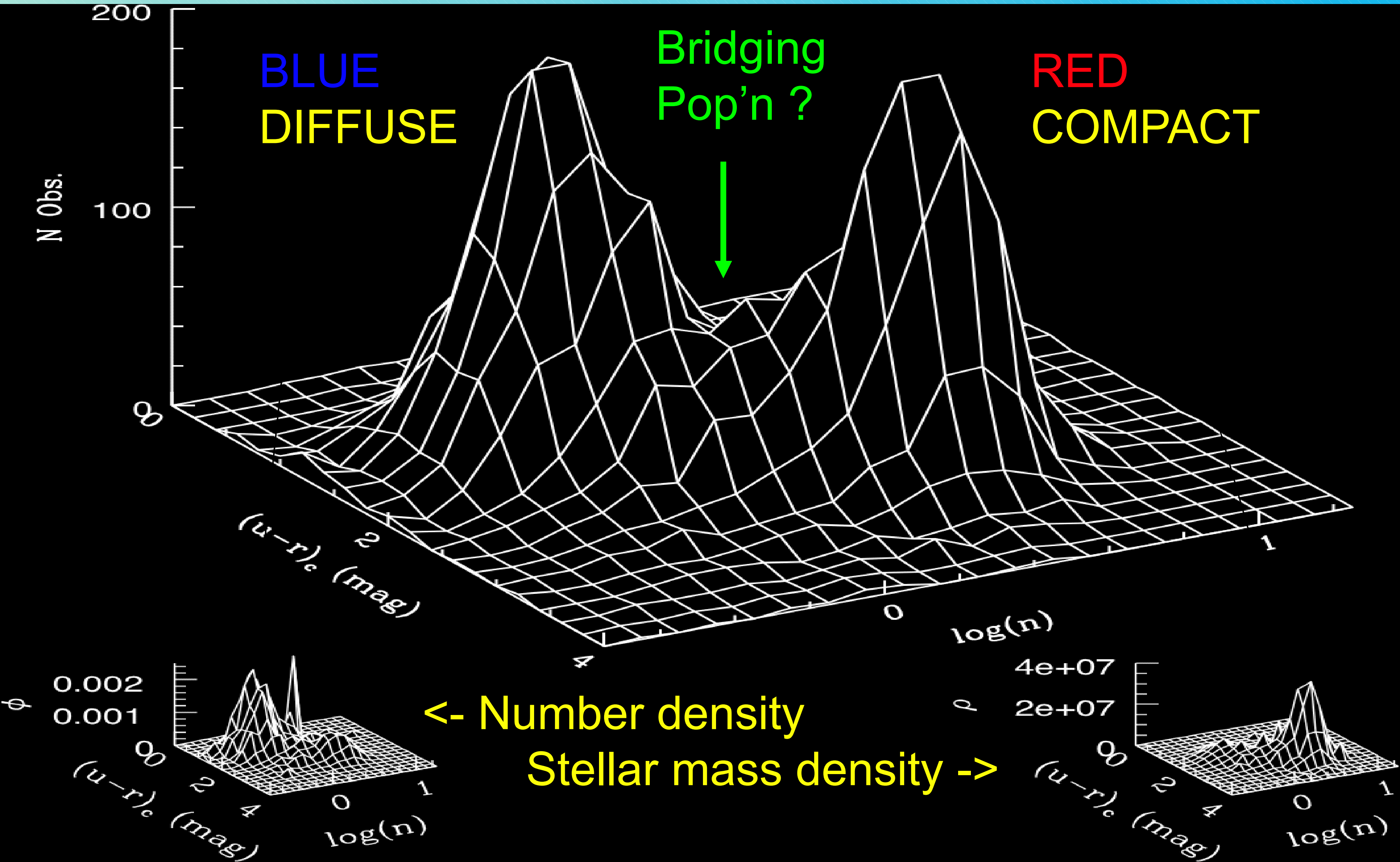
OBSERVED DISTRIBUTIONS (M_B < -16)

RED BLUE COMPACT DIFFUSE

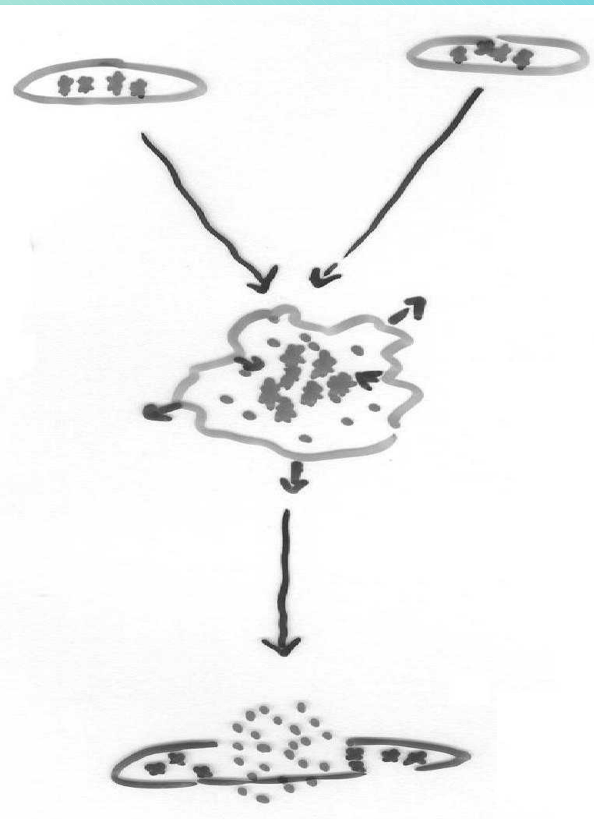


Galaxy bimodality in (u-r)-log(n)

Driver et al, 2006



SCENARIO A: All galaxies are born as disks. Dichotomy mainly induced by galaxy merging. Suggested by simulations of galaxy encounters and by the hierarchical build up of dark matter structures.

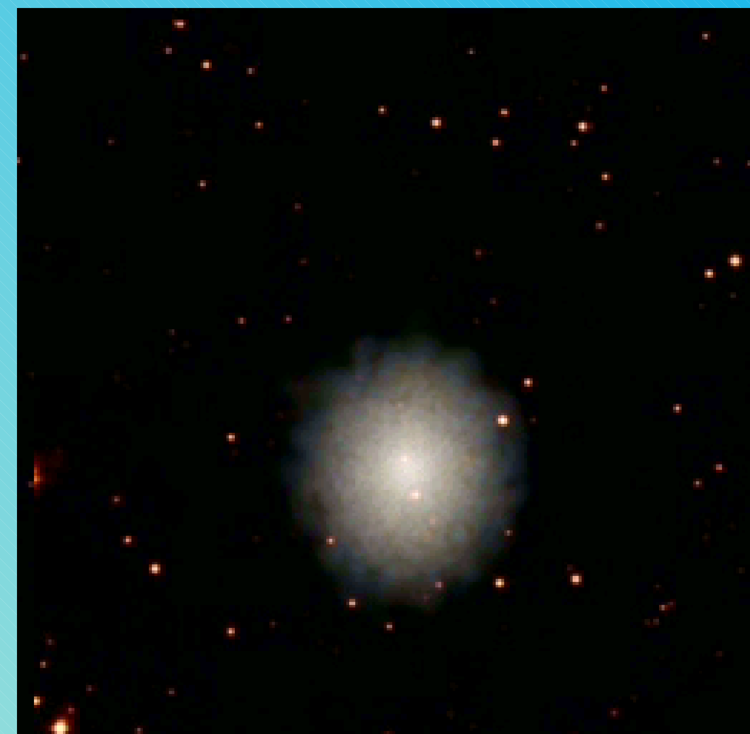


clouds of gas settle on disks with moderate ongoing star formation

Merger promote fast star formation and redistribute existing stars into spheroid

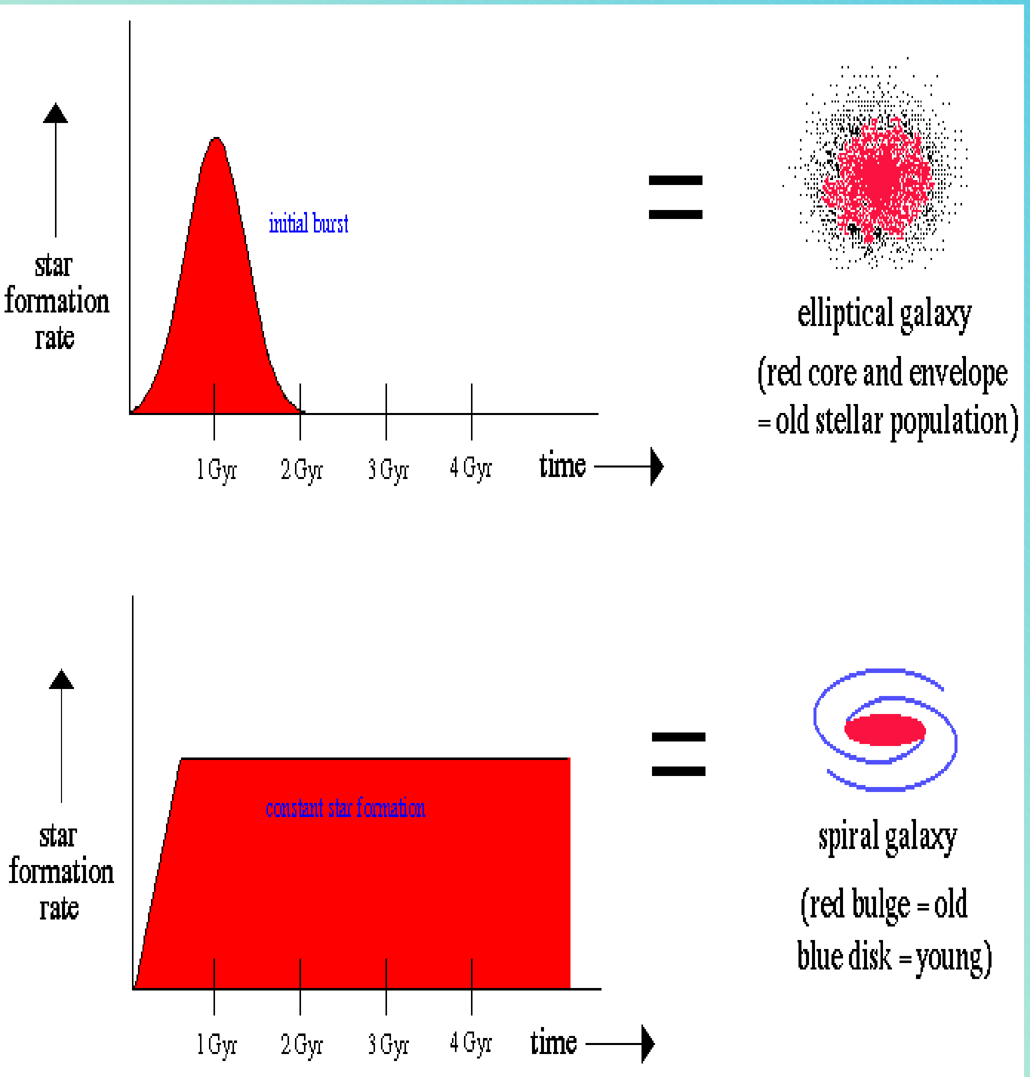
Over time, more gas is accreted in a new disc

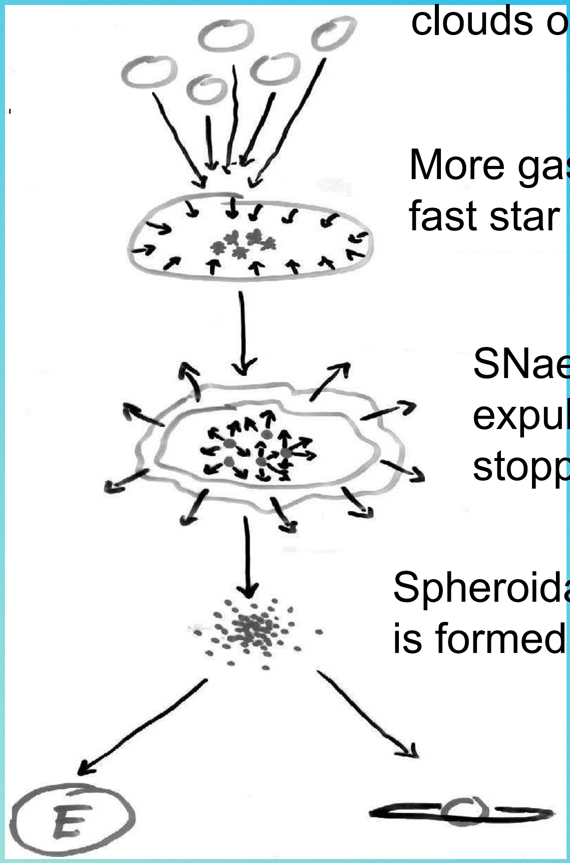
Collision between two disks, producing something similar to a spheroidal galaxy



NB: no build in link between morphology and star formation history. Also simulated mergers have too low phase space density when too little gas, or reform disk when too much gas.

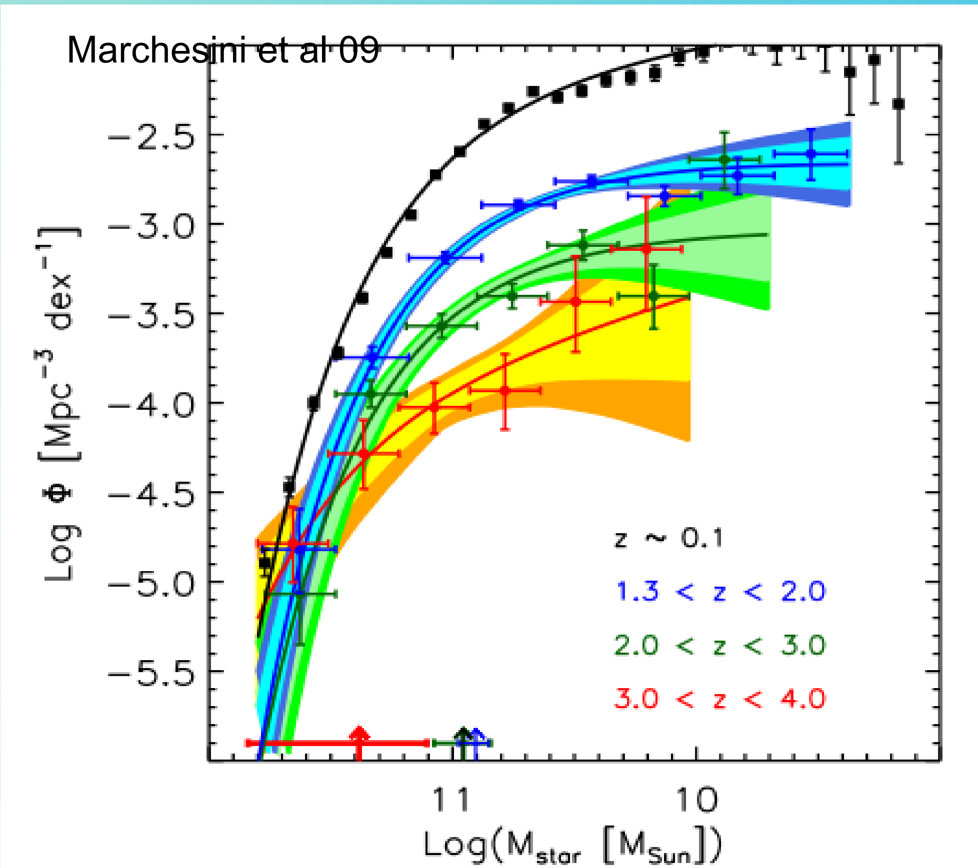
SCENARIO B: quasi-monolithic fast collapse of a gas at high z to form a spheroid, in suitable conditions followed by slow accretion to add a disk. Classically suggested by detailed chemical evolution models (“archeological downsizing”). Now supported by evidences of “downsizing in time”

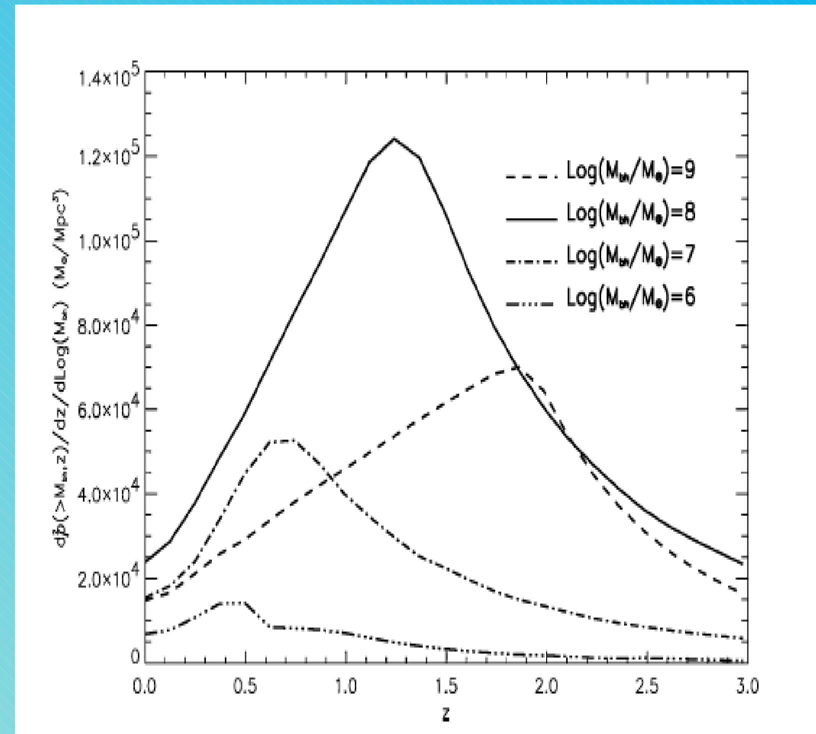




If no subsequent gas infall,
elliptical, otherwise spiral

Anti-hierarchical behavior of luminous matter





Mass increases (funny plot)

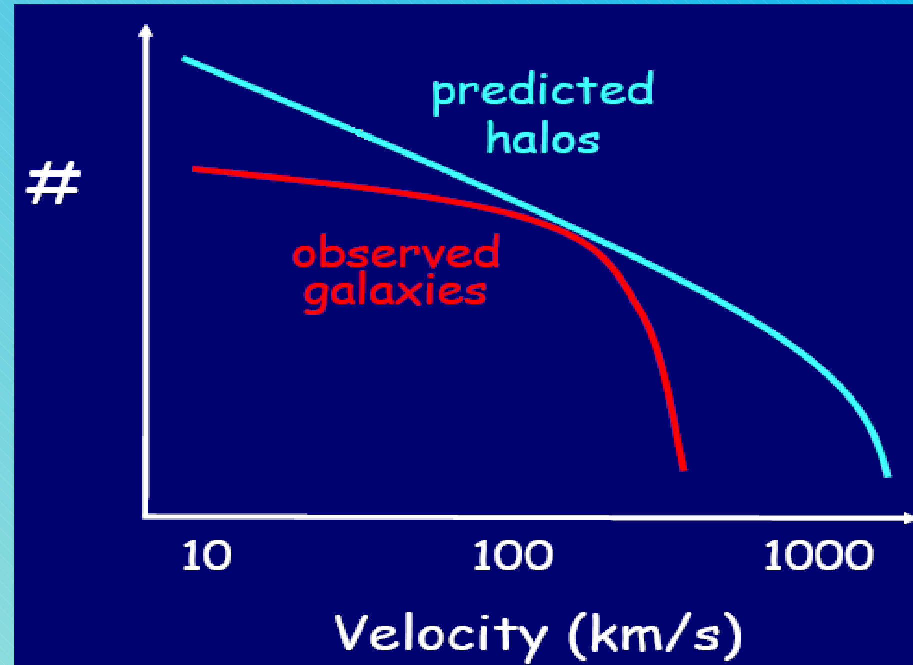
More massive galaxies (SMBH) on place (accrete) earlier and formed stars over shorter periods.

Massive galaxies showing huge activity at high z.

The “trendy” word is downsizing (in time), which is puzzling in hierarchical paradigm.

The luminosity function

The shape of mass function of dark matter halos differs substantially from the LF of galaxies, the most basic statistics of galaxy populations.

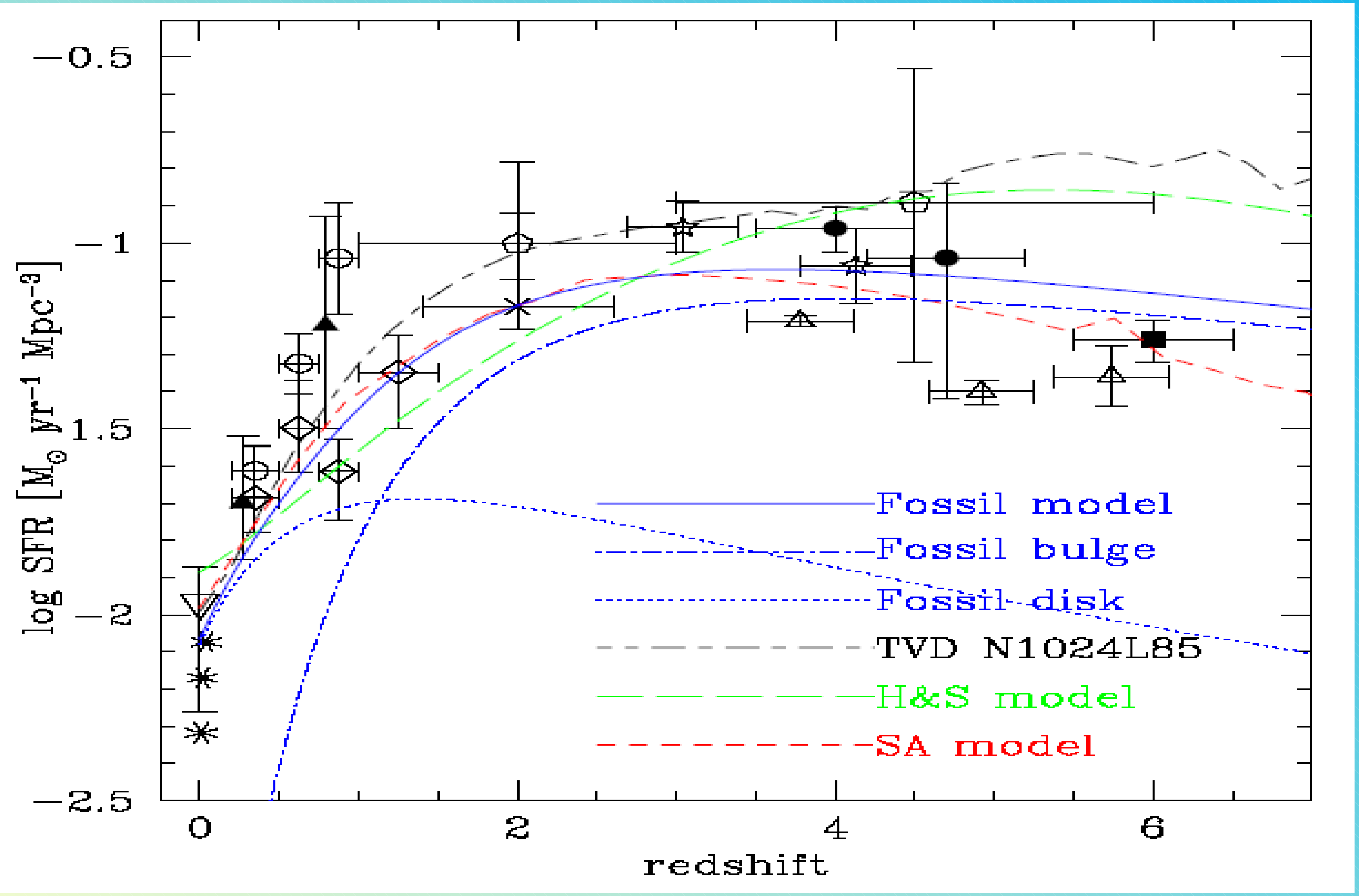


Mass or Luminosity

Questions:

1. Why galaxy formation so inefficient: only » 10% of baryons are in galaxies
2. Why the efficiency peaks at M_{halo} » 10^{12} M–.The sharp drop above is particularly puzzling

What drives the general behavior of the observed cosmic SF(z) ?



Connection between cosmic SF(z) and accretion(z)

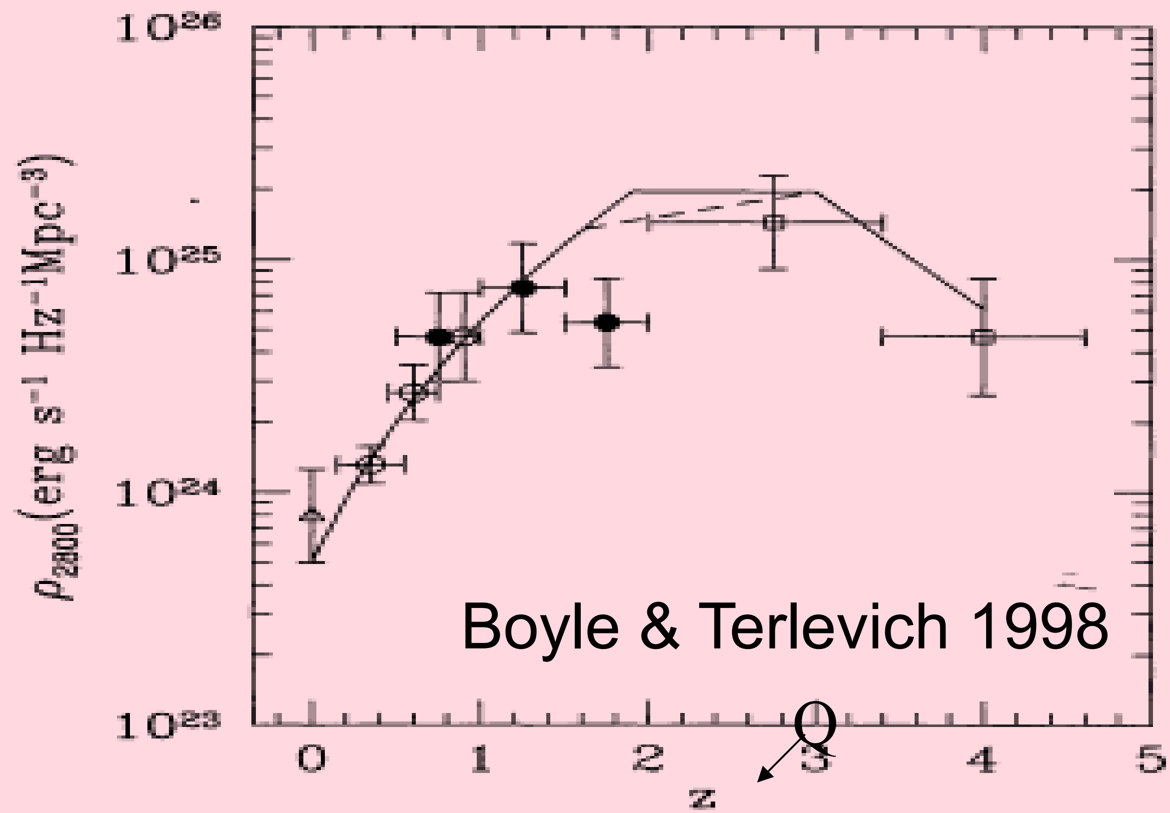
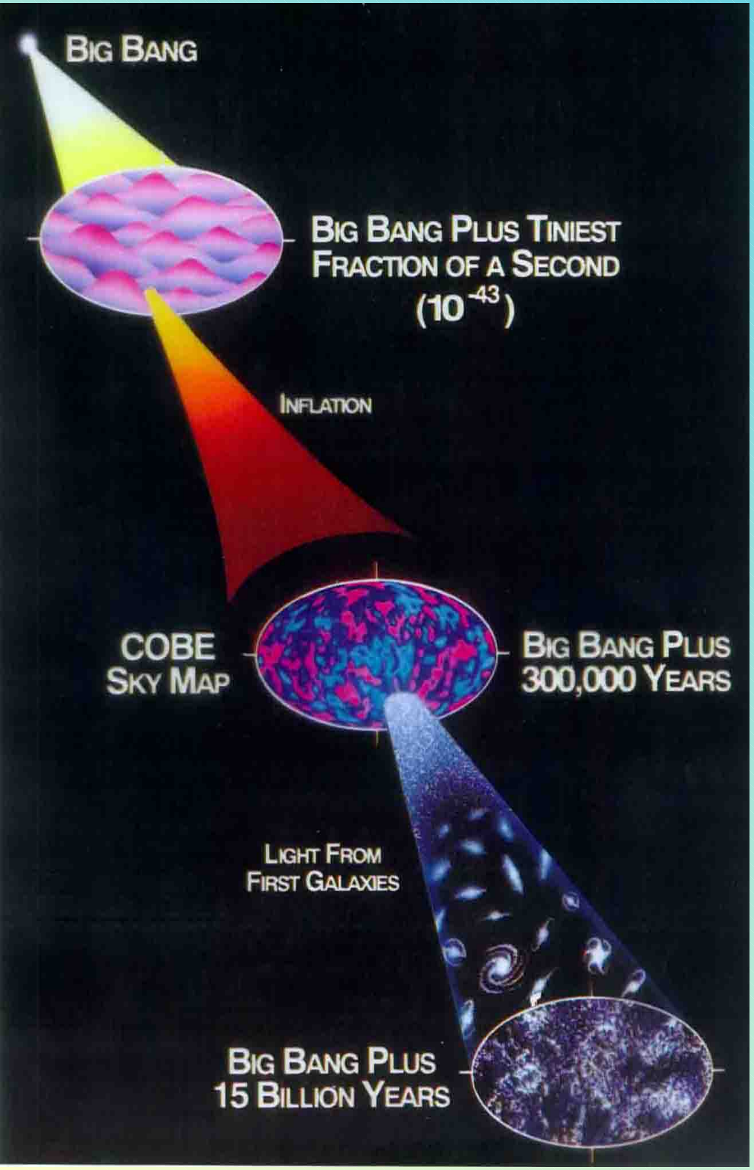


Figure 1. Luminosity density – redshift relation for WSOs (solid line) based on the evolutionary models of Boyle (1993) for $z < 3$ and Schmidt et al. (1995) for $z \geqslant 3$. The dashed line indicates the alternative evolution model of Hewett et al. (1993) for the redshift range $1.6 < z < 3$. The compilation of galaxy luminosity density scaled by 0.025) from Connolly et al. (1997) is also shown. Open circles: Lilly et al. (1996); open squares: Madau et al (1996), corrected for dust extinction by Pettini et al. (1997); open triangle: Guzman et al. (1997); filled circles: Connelly et al. (1997).

The hierarchical paradigm (of dark matter)



- We have a broad outline of the formation of cosmic structures (galaxies and clusters).
- They initially results from gravitational instability of small, primordial density fluctuations (likely quantum ripples boosted to macroscopic scales by inflation)
- Computation of their evolution requires knowledge of initial spectrum, cosmological parameters, and physical processes at work. As for ordinary matter (baryons) the difficulty is quickly formidable.

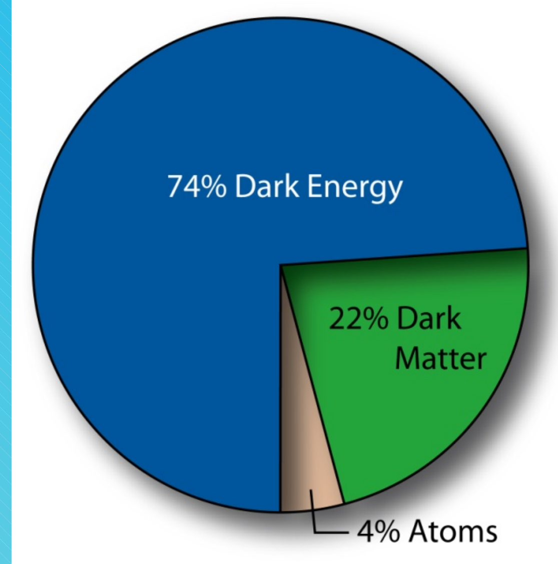
Background cosmology

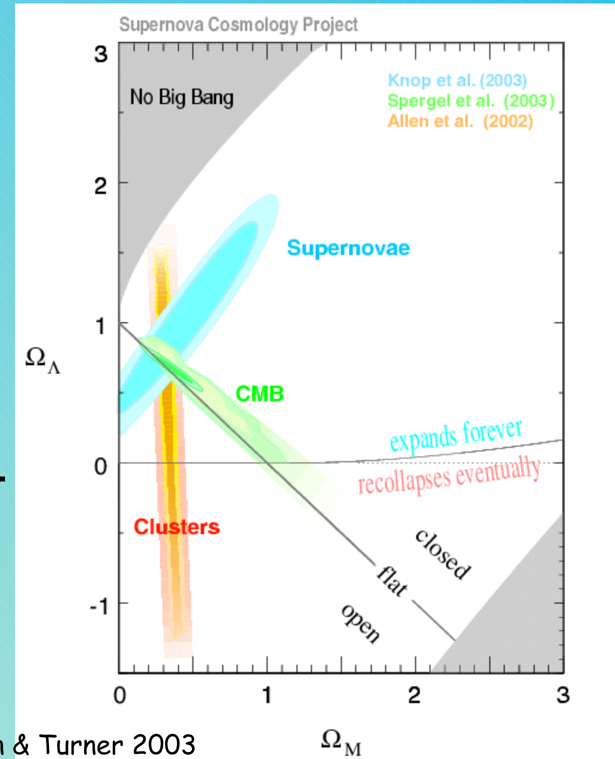
- "Precision era cosmology": cosmological parameters constrained typically at 1% (though mysterious in nature). Concordance ΛCDM model.

∀ $\Omega = \rho / \rho_{crit} = \Omega_{CDM} + \Omega_{bar} + \Omega_{\Lambda} \approx 1$. Universe is spat. flat.

∀ $\Omega_{bar} \approx 0.05$, $\Omega_{CDM} \approx 0.22$, $\Omega_{\Lambda} \approx 073$ (CMBR, large scale clustering , SN m(z) relation, BB nucleo-synthesis).

- And h ≈ 0.71 (HST key project + WMAP)
- Other fundamental quantities are the normalization $\sigma_8 \approx 0.8$ and pl index of the spectrum n ≈ 0.96 of (gaussian) fluctuations (CMBR and clusters).
- As for DM particles, it is sufficient to know if they are cold or hot, i.e. non relativistic or relativistic at matter-radiation equivalence and how they interact
- Several observations (CMB, galaxy clustering, weak lensing, Lyα forest) support cold DM and interaction only with gravity.





From Freedman & Turner 2003

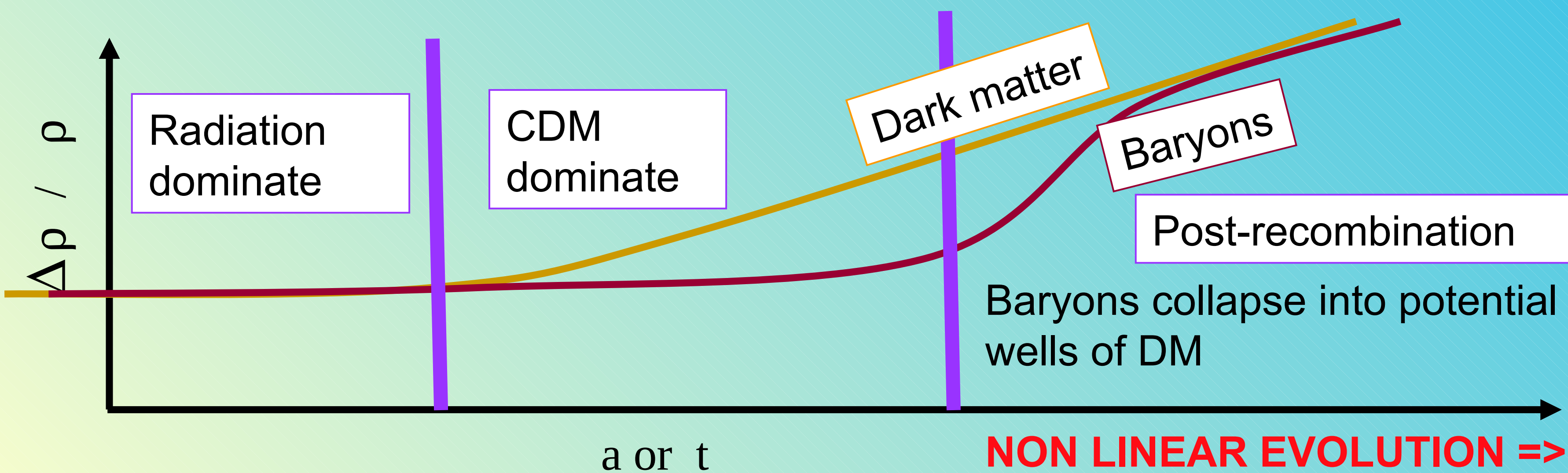
Consequences:

- Now we can concentrate on astrophysics of galaxy formation.
- Baryonic (aka “luminous” or “ordinary”) matter is dynamically minor. The problem can be split to some extent.
    1. First understand how (dark) matter concentrates under the only effect of gravity,
    2. then understand how ordinary matter evolves in DM structures to form galaxies

2 is a much more demanding task

Evolution in the linear regime (δρ /ρ) < < 1

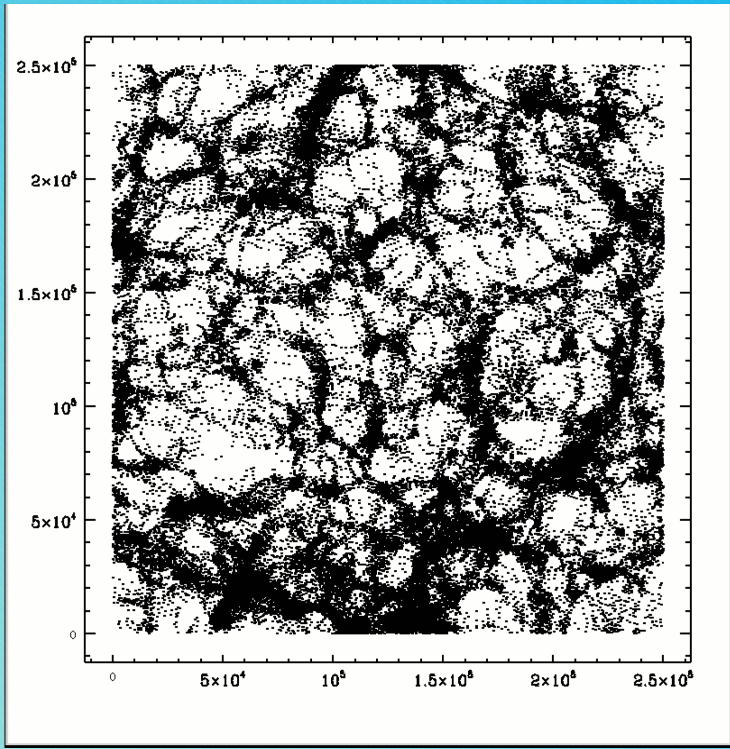
Can be computed exactly and analytically:
Radiation dominated era: perturbations grow in all components for λ>c/H (horizon), while for λ<c/H stall because of too rapid expansion.
Matter dominated era before recombination: DM perturbations grow (because cold), photons and baryon perturbations do not. Those on baryons can't grow due to the strong coupling with the high pressure photon fluid.
Matter dominated era after recombination: the pressure on baryons drops by orders of magnitude, baryons fall into potential well of DM perturbation (grown during the previous phase) and their perturbation soon equalize with those of dark matter. Then perturbations on both component grow proportionally to D(t)



NON LINEAR EVOLUTION => NUMERICAL SIMULATIONS

Sequence of non-linear structure formation

- The evolution beyond the linear regime needs other techniques.
- Approximate analysis suggests (spherical top-hat collapse, Zel'deovich approx..), and N-body gravity only simulations confirm, that significantly over-dense regions collapse to form gravitationally bound objects, Dark Matter Halos (DMH), which cluster and flow along a "web" of walls and filaments.
- DMH are the final state of DM overdensity evolution, near equilibrium and supported against further collapse by random motion of their particles
- They are characterized by large overdensities wrt background at collapse » 200



Sequence of non-linear structure formation

- Non linear collapse on mass scale M occurs when

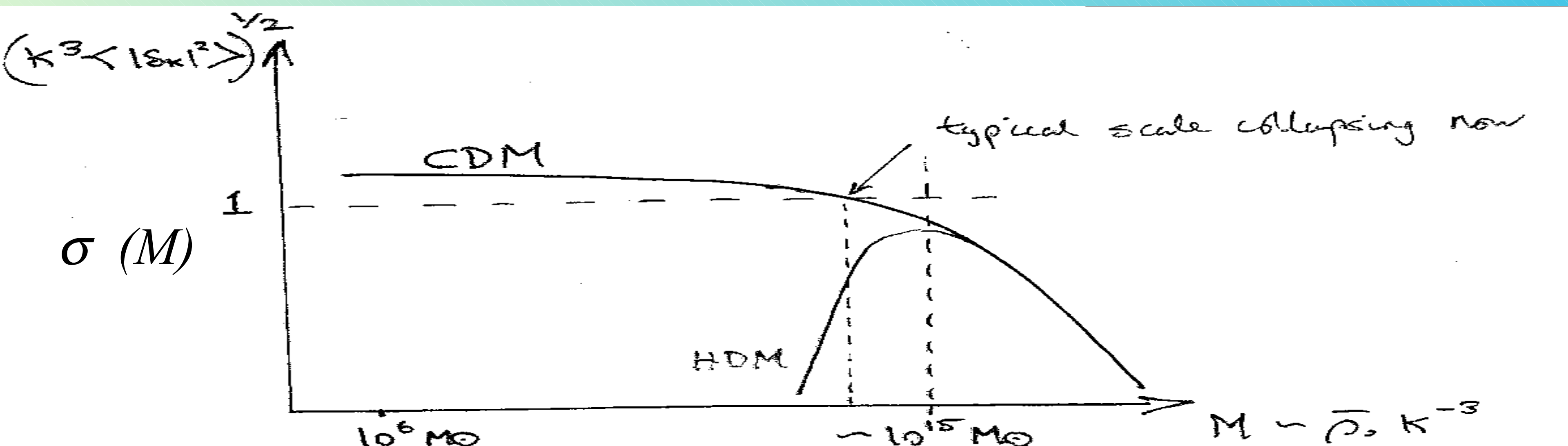
$$\sigma(M) \quad \left\langle \quad \frac{\delta\rho}{\rho} \quad ^{2} \right\rangle^{1/2} \quad 1$$

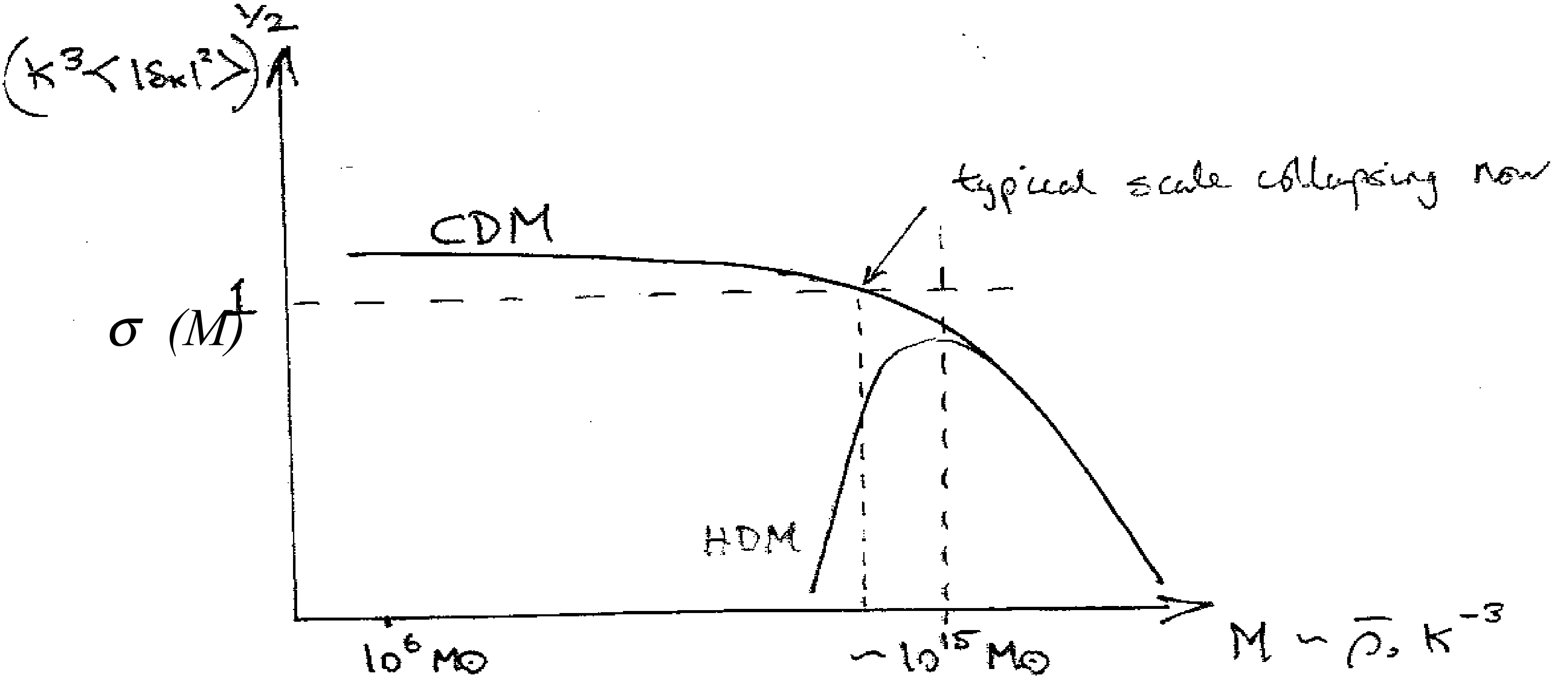
- In CDM σ is a monotonic decreasing function of *M:* small DM objects collapse earlier, forming larger ones also (still unclear how much, eg Genel et al 2010, Wang et al 2010) by clustering and merging

$$\Downarrow$$

Hierarchical clustering

Not clear how to reconcile with downsizing of baryons





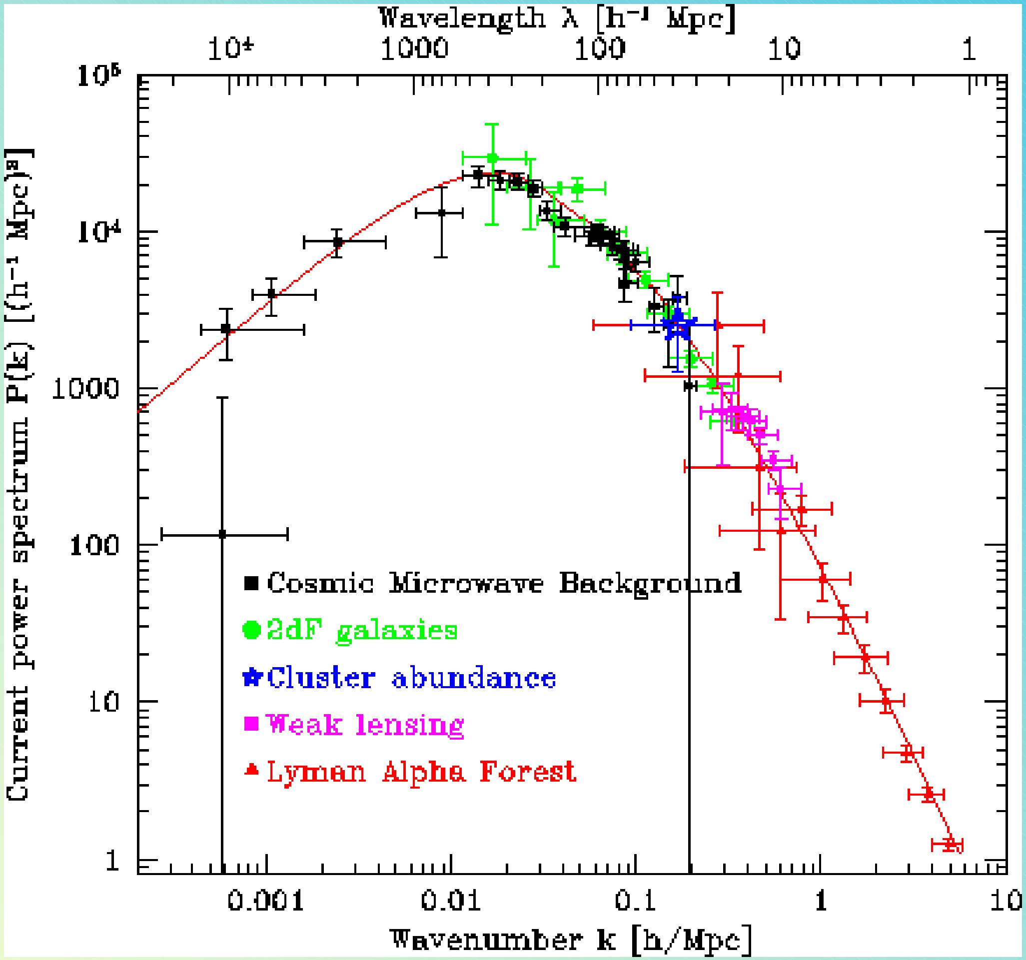
CDM:

forms halos with $10^6 < M/M_{\odot} < 10^{15}$ by present day

HDM:

- first objects to form are DM halos with $M/M_{\odot} \sim 10^{15}$ (galaxy cluster)
- Do not form galaxy size halos ($M/M_{sun} \sim 10^{12}$) at all

Observed power spectrum of linear density fluctuations

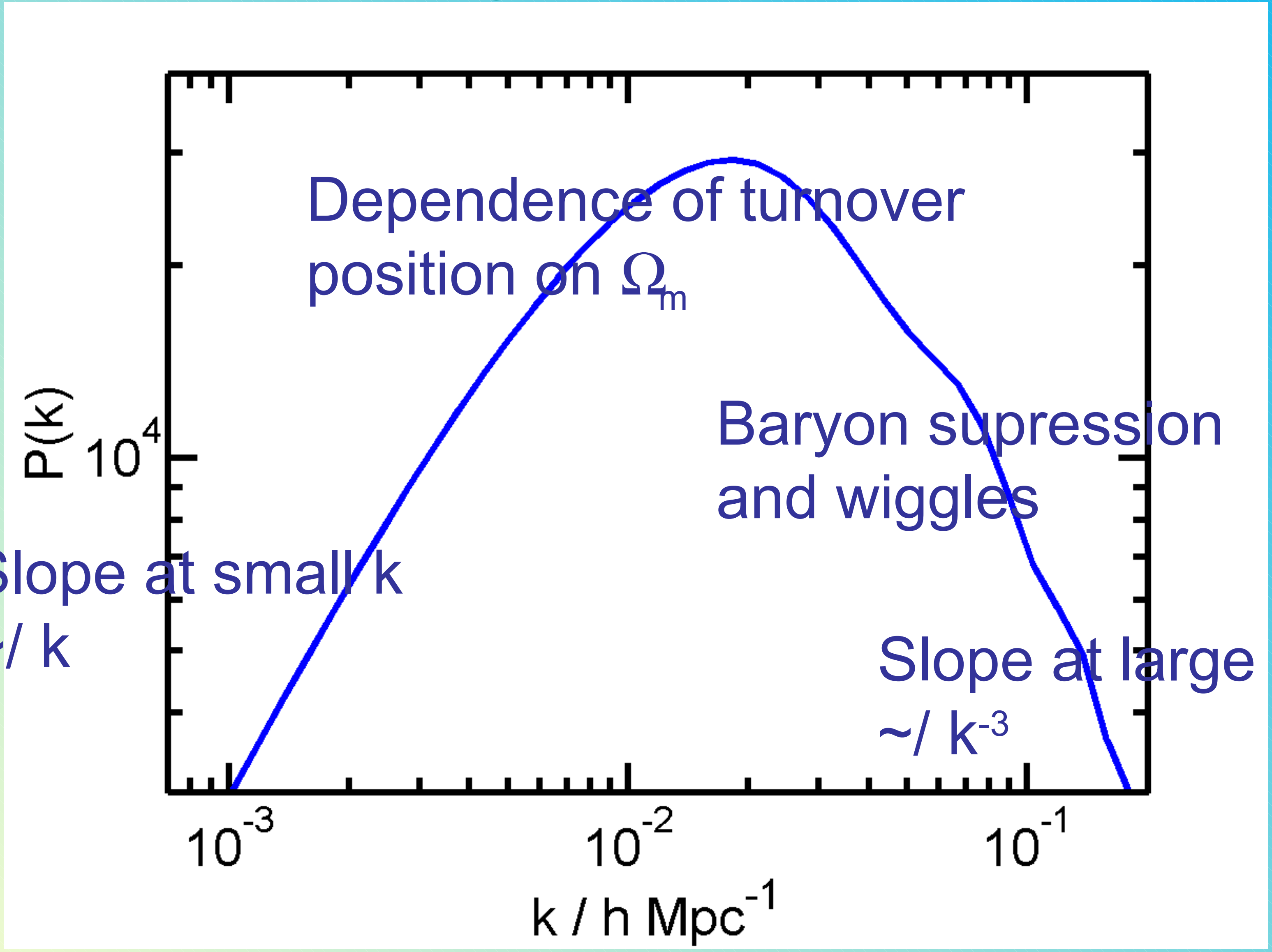


Obs measures of P(k) from:

- CMB
- galaxy clustering
- weak lensing
- Lyα forest

- confirm ΛCDM

Key features



Dependence of turnover position on Ω_m

Baryon supression and wiggles

Slope at small k ~/ k

Slope at large k ~/ k^{-3}

But galaxy formation is much more than gravity

Galaxy formation occur in DM halos and involves a complex web of processes, in which enters all (astro)physics

1. formation & merging of dark matter halos, starting from primordial density fluctuations
2. shock-heating & radiative cooling of gas in DM halos
3. collapse of cold gas & star formation from cold gas
4. energy input into gas from stellar winds and SN explosions ("classical" feedback)
5. chemical enrichment of gas & stars (chem. feedback)
6. galaxy mergers and tidal interactions
7. Formation of SMBH-AGN and its feedback to ISM
8. luminosity evolution of stellar populations
9. absorption of starlight by dust & re-emission in IR+sub-mm

Baryonic processes

Numerical simulations are the main tool for the non-linear regime.

Useful to distinguish 4 subgroups:

1. N-body gravity only simulations; suffice (almost) for DM, but provide only a biased, very rough idea of galaxy formation. However give a "backbone" and "analytic prescriptions" for the physics of galaxy formation.

2. Hydro-dynamical simulations without sub-grid physics; treat also visible matter. Yet not enough resolution (by many orders of magnitude) to understand galaxy formation, good for less collapsed structures. Gives also suggestions for deeper modelling, e.g. on the processes of heating and cooling of gas.

But to compute where galaxies form in DM halos and their properties would require resolution << 1pc (to calculate directly SF and SN feedback) in a box >>10 Mpc to have a representative volume of the universe;

This is by far unfeasible, thus even in the best simulations of galaxy formation in cosmological context, most relevant processes are treated with crude analytical sub-grid prescriptions….

…Two complementary routes are employed:

3.Hydro-dynamical simulations including phenomenological sub-grid physics (aka “gastrophysics”; e.g. star formation); In principle, the best we can do, but results dominated by sub-grid recipes or prescriptions.

4.almost nothing further to lose using this approach for every process involving baryons: semi-analytical models (SAM). Possible advantages over 3 include fast exploration of effects of different assumptions

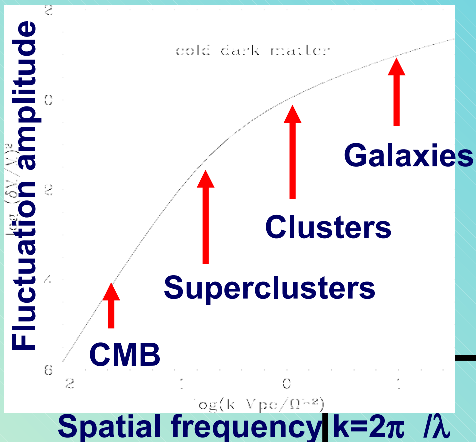
4 is by far the most used technique for extensive comparisons with data

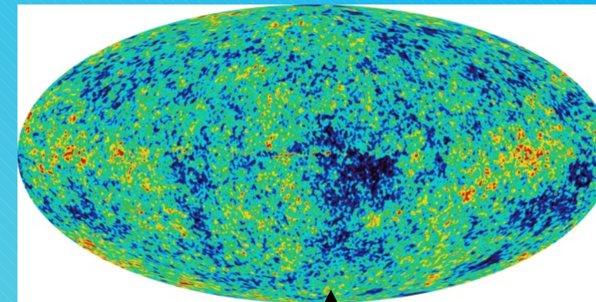
The border between 3 and 4 is becoming more and more fuzzy

In both cases, the predicted outcomes are heavily affected if not dominated by these sub-grid processes

Main message of these lectures is that galaxy formation models are not “first principles models” but rather “toy models”.

Modelling galaxy formation





Cosmological model

(Ω, Λ, h) dark matter

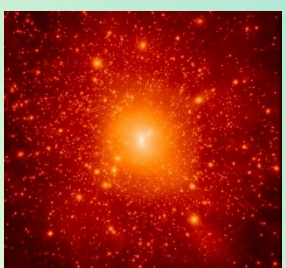
Primordial fluctuations

δρ /ρ (M, t)

Evolution of dark matter halos

Dynamics of cooling gas

N-body simulations



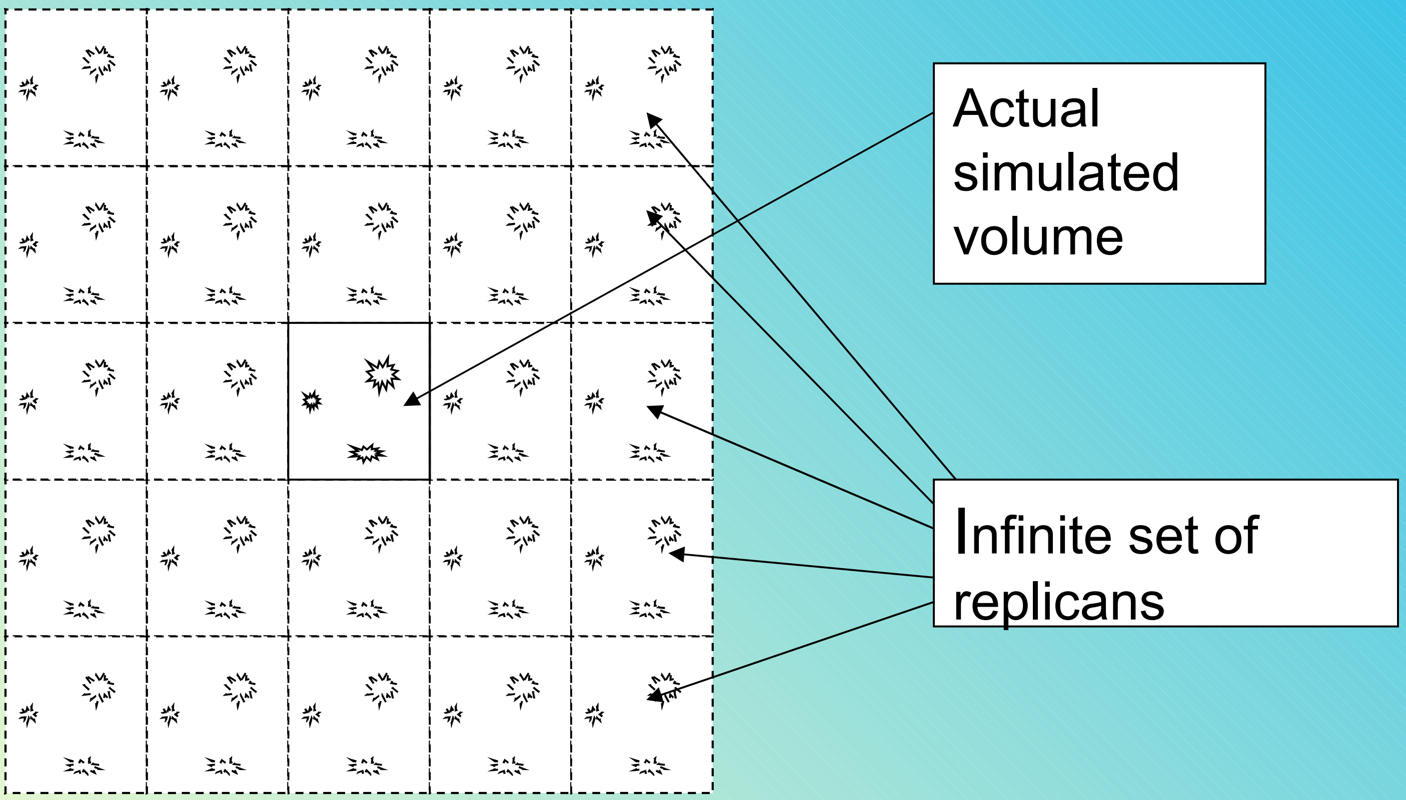
- Gasdynamic simulations
- Semi-analytic modelling

Star formation, feedback, evolution of stellar pops

Galaxy mergers

Formation and evolution of galaxies

The universe is homogeneous on large scales, thus one wants to calculate the evolution of a representative volume, using periodic boundary conditions



•avoids particles feeling edge of the simulation volume
•easily implemented if potential computed using Fourier Transform

Numerical simulations are started at a redshift z_i (typically from a few 10 to »100) before analytical techniques break down, i.e. from initial conditions given by linear perturbation theory

Procedure to compute initial conditions:

1. calculate random realization of primordial density fluctuations in the periodic box (easy for gaussian, cumbersome otherwise)
2. evolve its power spectrum using linear theory (transfer functions; public codes to compute them as CMBFAST)
3. set up uniform distribution of particles in the box on a lattice to represent unperturbed state of uniform density
4. give these particles displacement and velocities using Zel'dovich approximation (which is exact for $\delta<<1$)

4 is required to translate from linear theory (Eulerian) to Lagrangian (particle) numerical simulations

Another, less common, possibility is to perturbate the masses.

1. N-body gravity only simulations

System represented by a set of particles *sampling* the phase space distribution, interacting only by gravity.

Algorithm to evolve system in its essence is trivial:

1. Set up initial conditions
2. Calculate gravitational force on all particles
3. Update position and velocities for small time-step using derivatives
4. Repeat from 2 for many time-steps

Many complications to improve efficiency (e.g. individual Δt, leap frog,….).

Almost all CPU time goes in 2, due to long range nature) a lot of efforts to improve over "brute force" use of Newton's law

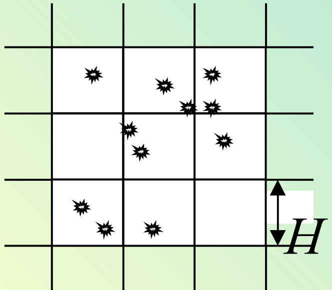
One slide review of N-body methods

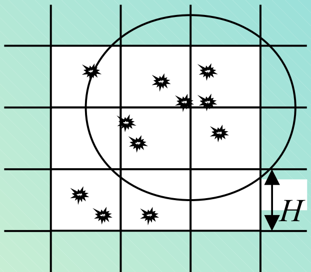
PP (Particle-Particle) or direct summation brute force use of Newton's law for gravitational force. Impractical for the whole system ($T(CPU) \propto N^{1.6 \div 2}$) but used by smarter methods in portions of the system.

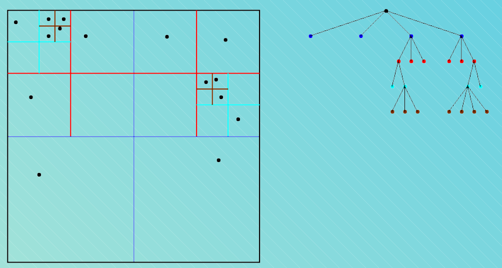
PM (Particle-Mesh or grid) calculates forces on particles from potential obtained solving Poisson equation on a mesh in the Fourier space (where it is algebraic). Not good for CDM.

P^3M and adaptive P^3M use PM for forces due to distant particles, and PP for nearby particles. Transition occurs at (2¥3)H. Speed and high resolution (set by softening) are achieved, provided not to much clustering. Otherwise ! PP.

Tree (approximate!) groups distant particles in pseudo-particles to estimate the force they exert using PP. Space is divided iteratively into cubic cells until a cell contains at most 1 particle.







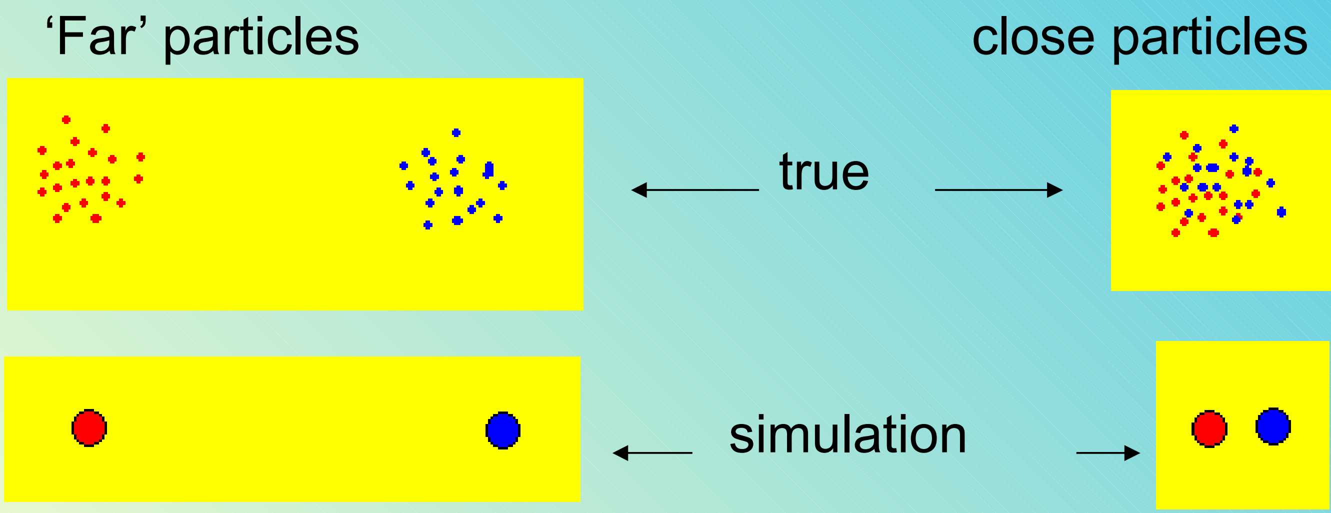
1. N-body gravity only simulations

The system is represented by a set of particles *sampling* the phase space distribution, interacting only by gravity.

The mass of the simulation particle (say not less than 10^6 $M_\odot$ even in single halo sim.) is always >>>>>>>>>> the mass of real DM particles (likely in the range of masses of elementary particles).

The real fluid is much smoother than computer fluid. Thus

- mass resolution
- artificial two body relaxation



The algorithm to evolve system in its essence is:

1. Set up initial conditions
2. Calculate gravitational force on all particles
3. Update position and velocities for small time-step using derivatives
4. Repeat from 2 for many time-steps

Many complications to improve efficiency (e.g. individual Δt, leap frog,….).

Almost all CPU time goes in 2, due to long range nature of gravity.

PP (Particle-Particle) or direct summation method is brute force use of Newton's law for gravitational force

$$\vec{f}_i = \sum_{i \neq j} \frac{G m_j}{\left|(\vec{r}_i - \vec{r}_j)^2 + \varepsilon^2\right|^{3/2}} \left(\vec{r}_j - \vec{r}_i\right)$$

Softening parameter
Avoids artificial two body relaxation
Sets spatial resolution

Impractical for the whole system ($T(CPU) \propto N^{1.6 \div 2}$) but used by smarter methods in portions of the system

Also hard to implement periodic boundary conditions

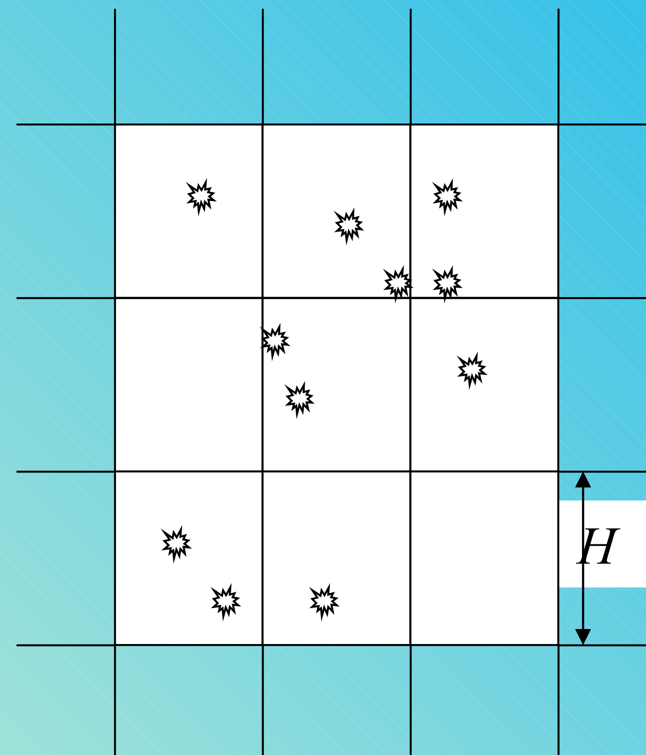
PM (Particle-Mesh or grid) method calculates forces on particles from potential obtained solving Poisson equation on a mesh in the Fourier space (where it is algebraic):

$$\nabla_x^2 \phi = 4\pi G a^2 \delta\rho(\vec{x}, t)$$

$$\phi = \sum \phi_k \exp(i\vec{k}\bullet\vec{x}) \qquad \delta\rho = \sum \delta\rho_k \exp(i\vec{k}\bullet\vec{x}) \quad \Rightarrow \phi_k = -\frac{4\pi G a^2 \delta\rho_k}{k^2}$$

1. particles smoothed to give density field on a grid (NGP, CIC, TSC).
2. potential computed on a grid solving the Poisson equation in the Fourier domain. FFT is used back and forth.
3. Force on particles estimated differentiating (finite difference) and interpolating from grid points

$$\vec{f} = -\nabla\phi$$



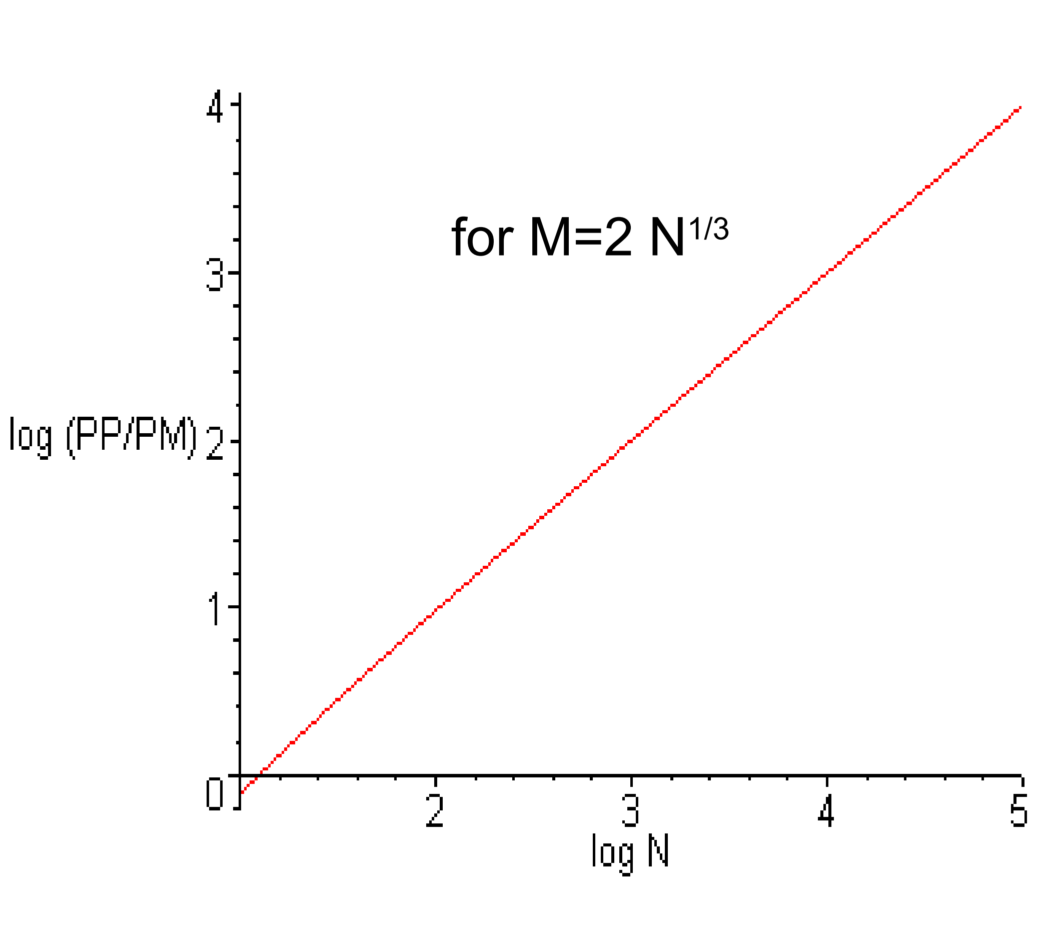
PM can be extremely faster than PP:

T(CPU, PP) $\propto$ $N^{1.6 \div 2}$

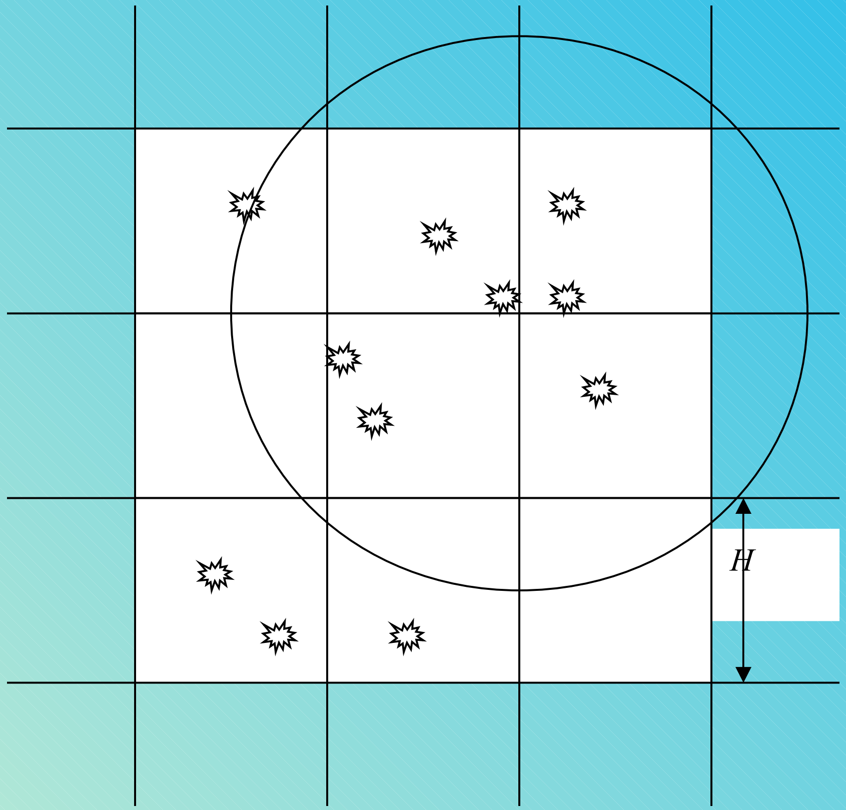
T(CPU,PM) $\propto$ $2N+M^3\log_2 M^3$ where M number of grid points in each direction, usually $M/N^{1/3}$

Also, periodic boundary conditions are automatic (FT)

But forces are softened on a scale ' grid spacing. Good if the potential varies at scales ' a few mesh lengths. OK in HDM (no power on small scales), but not in CDM) Now is out of fashion



Combined methods: P^{3}M and adaptive P^{3}M use PM for forces due to distant particles, and PP for nearby particles. Transition occurs at (2¥3)H. Speed and high resolution (set by ε) are achieved, provided not to much clustering. Otherwise ! PP.



Adaptive P^3M

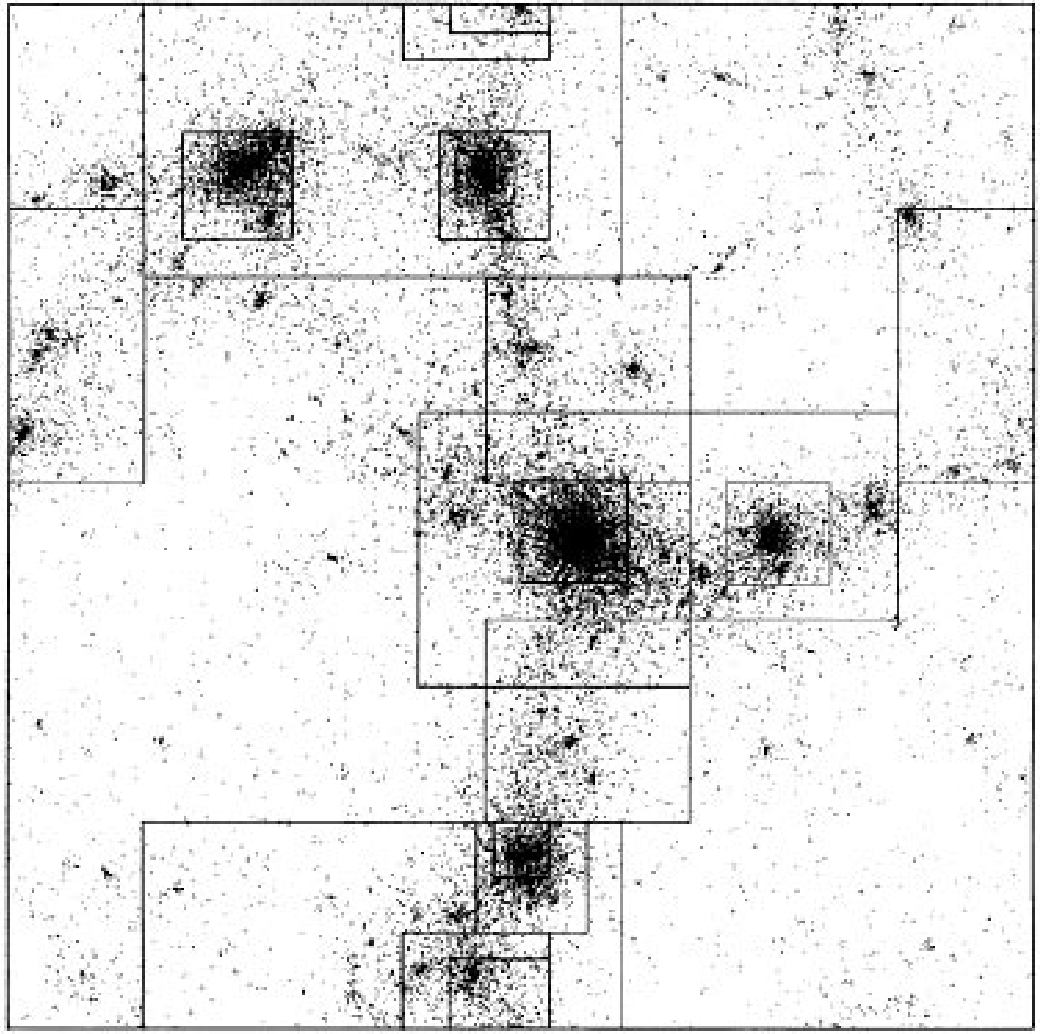


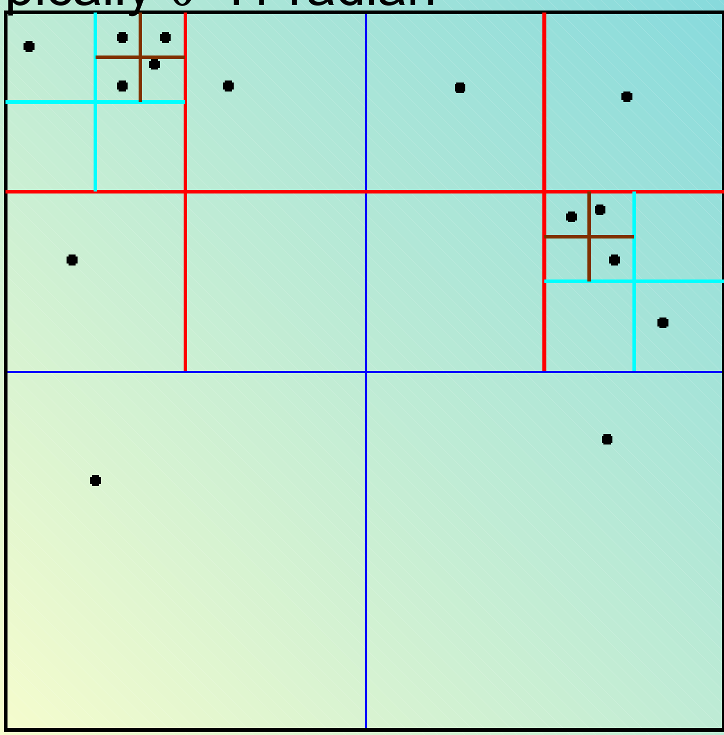
Figure 1 Distribution of grid refinements placed by an adaptive particle-particle/particle-mesh–smooth-particle hydrodynamics (P^3M-SPH) code for the final timestep of a clusters simulation. Gas particles are shown. From Couchman et al (1995).

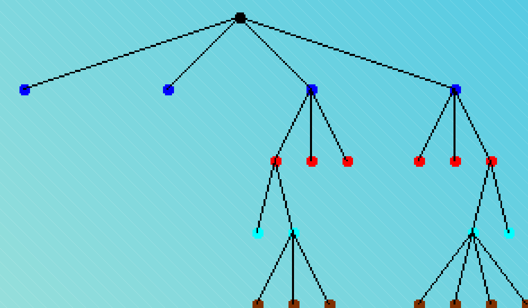
Tree method (an approximate method!) groups distant particles in pseudo-particles to estimate the force they exert using PP. Space is divided iteratively into cubic cells until a cell contains at most 1 particle.

•If distance>size/θ   a cell is "far" and the force due to the particles in it is approximated by their sum located in the centre of mass
•If distance<size/θ   a cell is "near", sub-cells are treated singularly, recursively

$$O(N \log N)$$

Typically θ .1 radian





1) Build the Quadtree,
2) For each subsquare in the quadtree, compute the center of mass and total mass for all the particles it contains.
3) For each particle, traverse the tree to compute the force on it.

Advantages:

- high speed (but slower than P^3M): T(CPU, tree) / N log N
- relatively easy to implement and publicly available algorithm;
- spatial adaptive: resolution automatically refined if needed;

Problems:

- large amount of memory are needed
- does not provide directly periodic boundary conditions
- It is approximated

Slower than P^3M, but better spatial resolution, so often P^3M used simulate large boxes, then individual object re-simulated with a tree code.

Refinement technique:

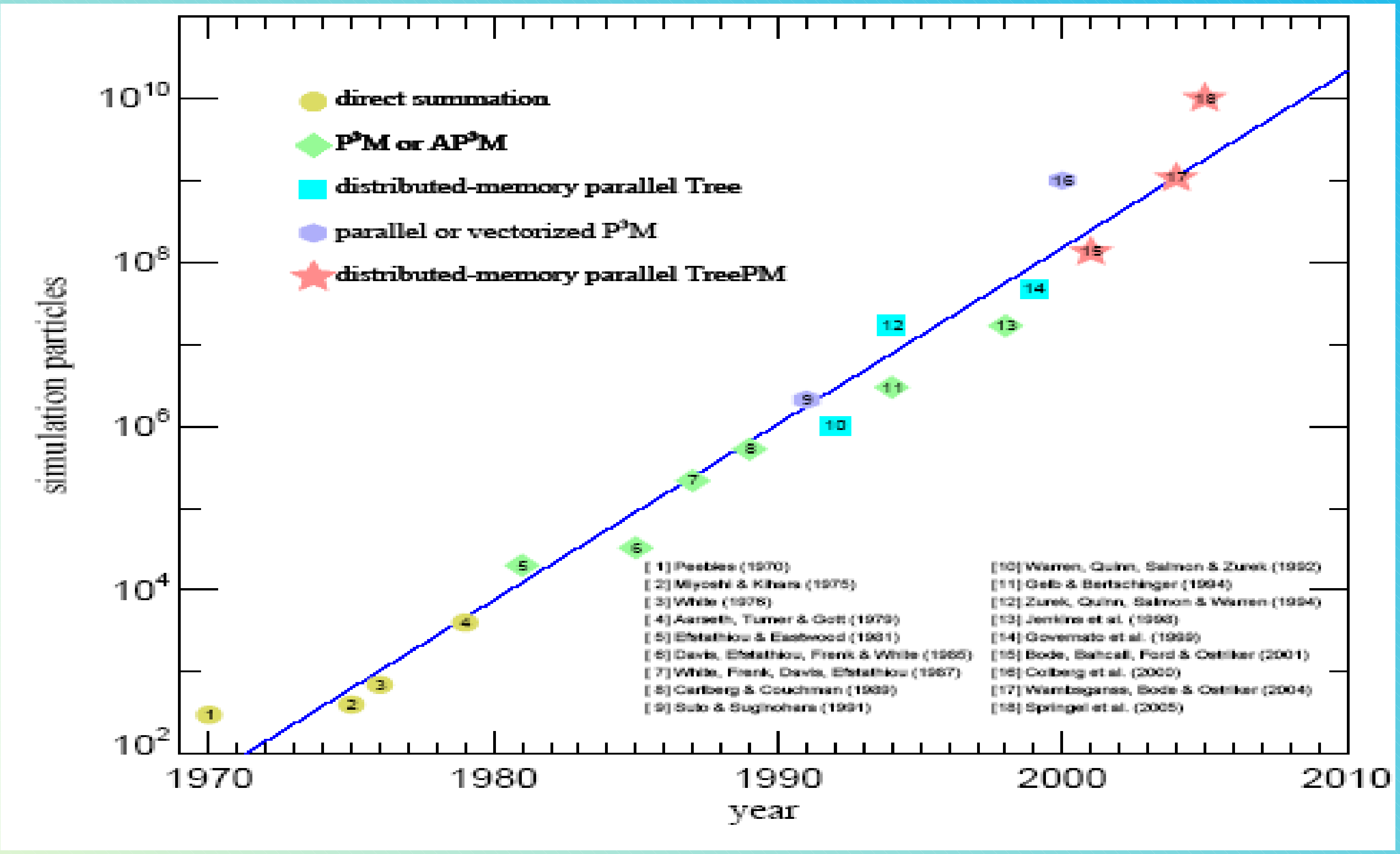
Objects are identified

Their particles tracked back

The initial volume they occupy re-simulated at higher res, adding small scale fluctuations

GADGET is a tree code that can combine directly with P^3M

Huge progresses in past 40 years
faster CPUs + better algorithms



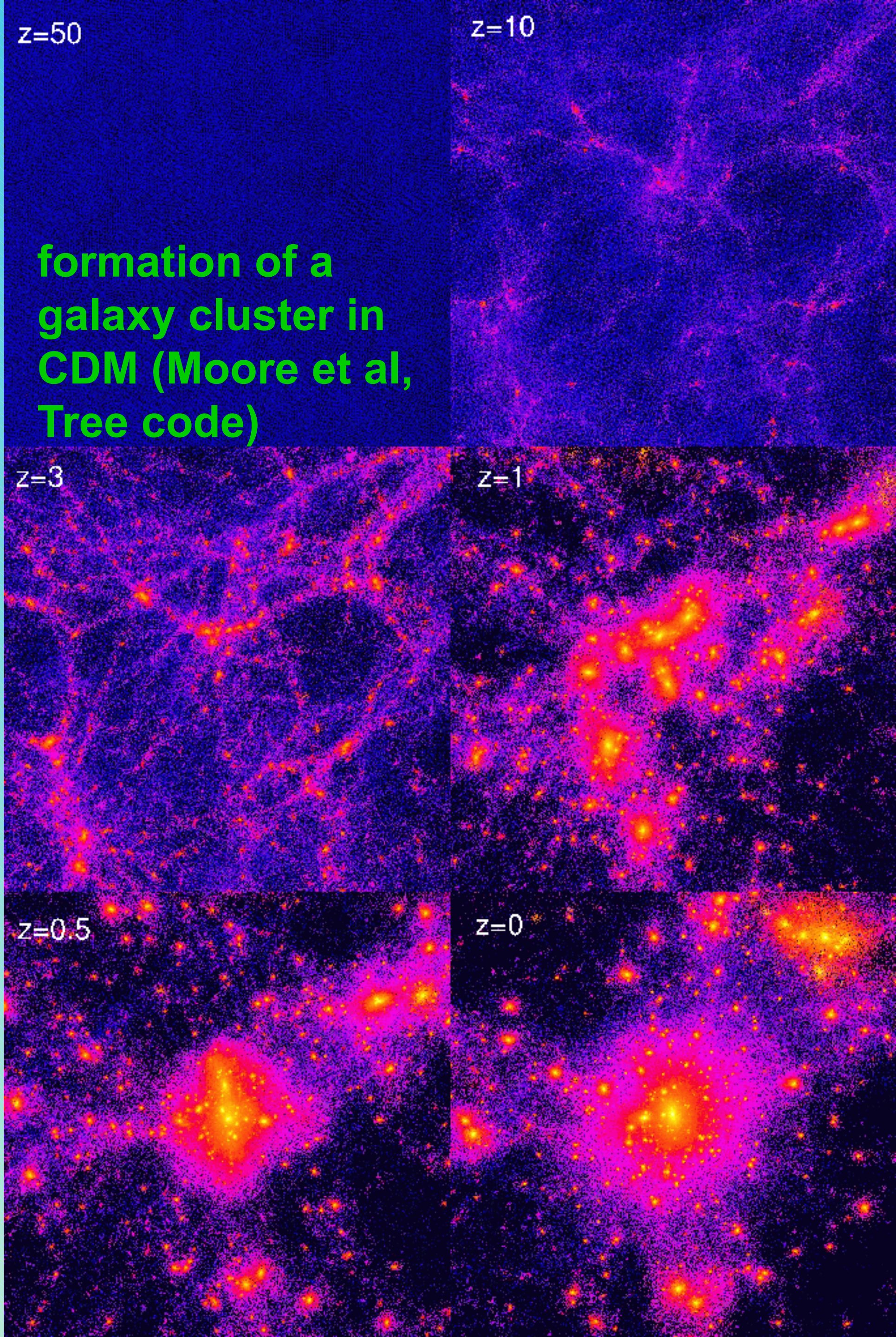
Particle number in highest resolution N-body simulations of cosmic structure formation as a function of publication date.

A small selection of results from N-body simulations

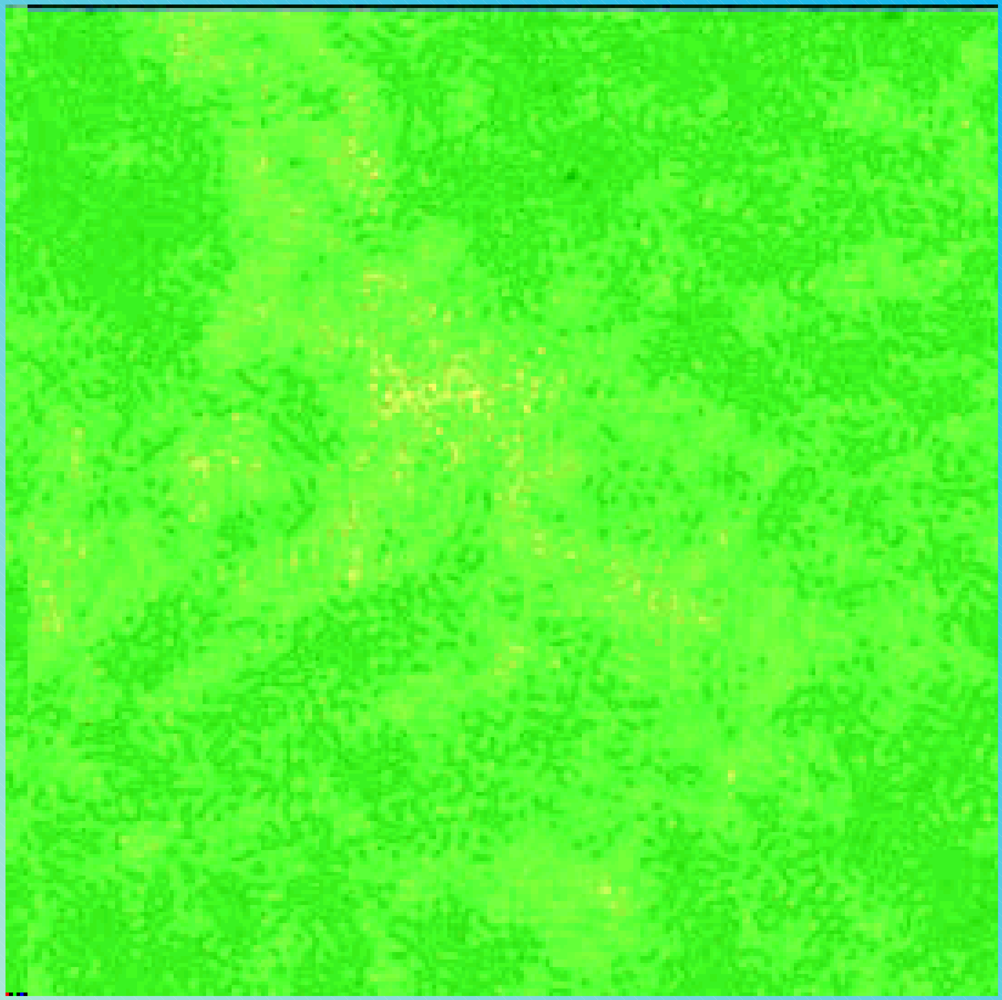
General DM evolution from N-body simulations.

In CDM

1. over-densities collapse into **sheets** ($\delta\rho/\rho \sim 1$) and **filaments** ($\delta\rho/\rho \sim 10$)
2. matter flows along filaments into dark matter **halos** ($\delta\rho/\rho > 100$) (roughly spherical objects in approximate virial equilibrium)
3. halos merge hierarchically to form bigger and bigger halos (this statement is under revision as of 2010..)



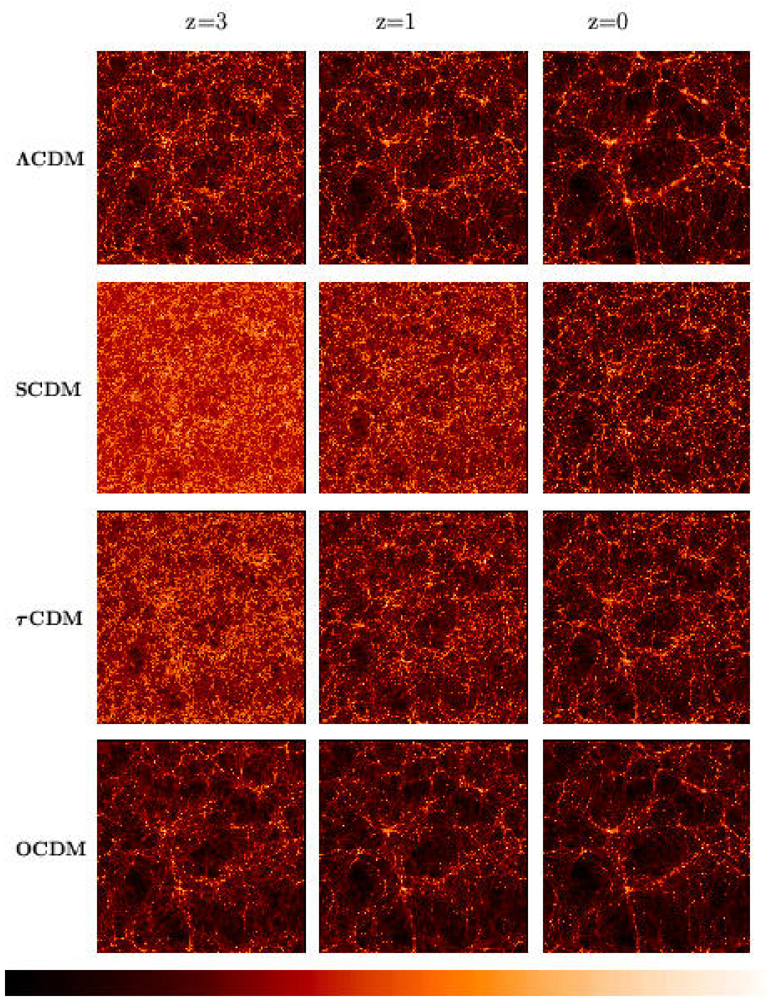
Formation of a galaxy cluster in a CDM simulation. (Starting redshift 20.0)



The Virgo Consortium

By construction the developments of structures is very similar at z=0, but differs significantly at higher redshift. Boxes are 240 Mpc/h

(HYDRA: adaptive P^3M)

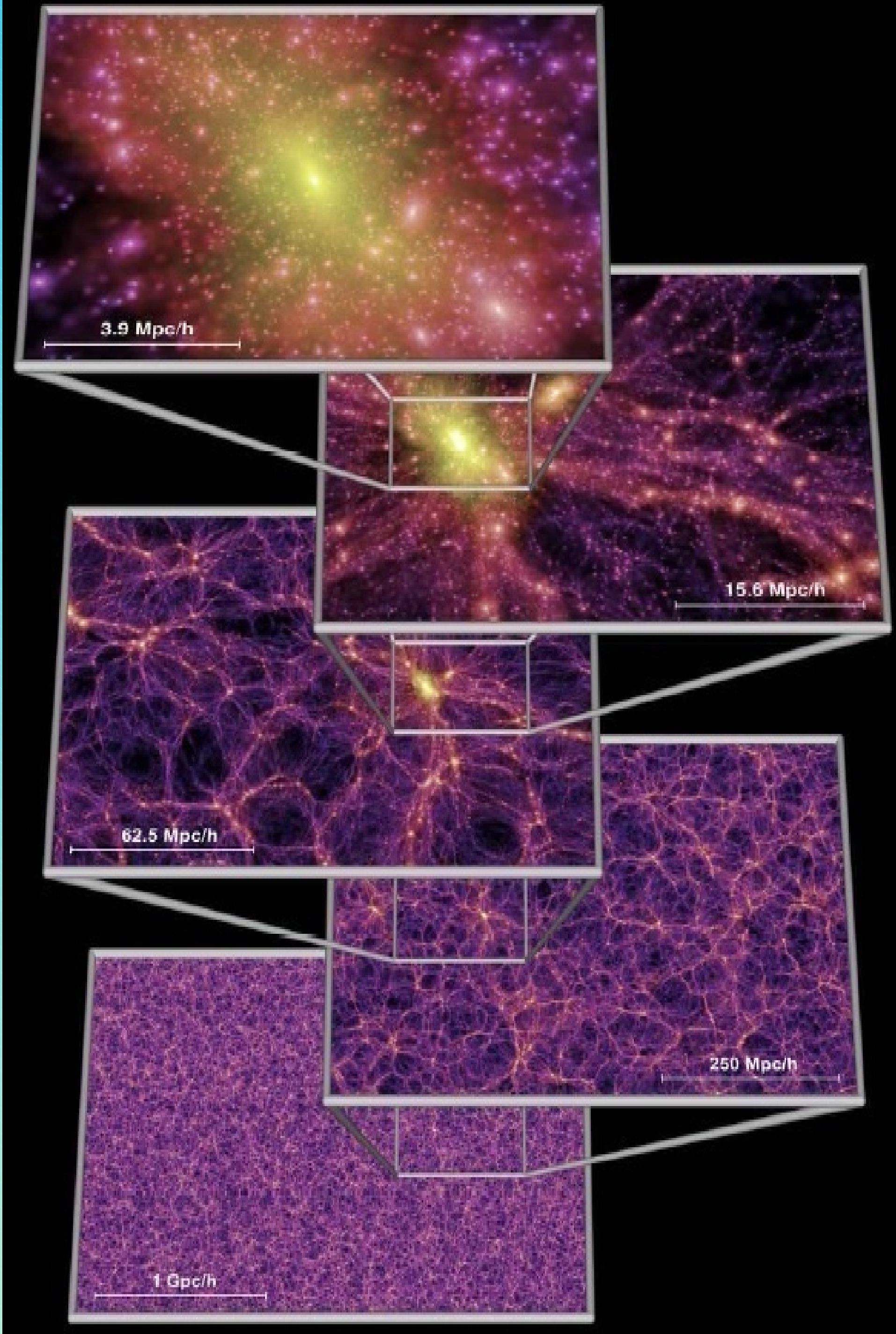


The VIRGO Collaboration 1996

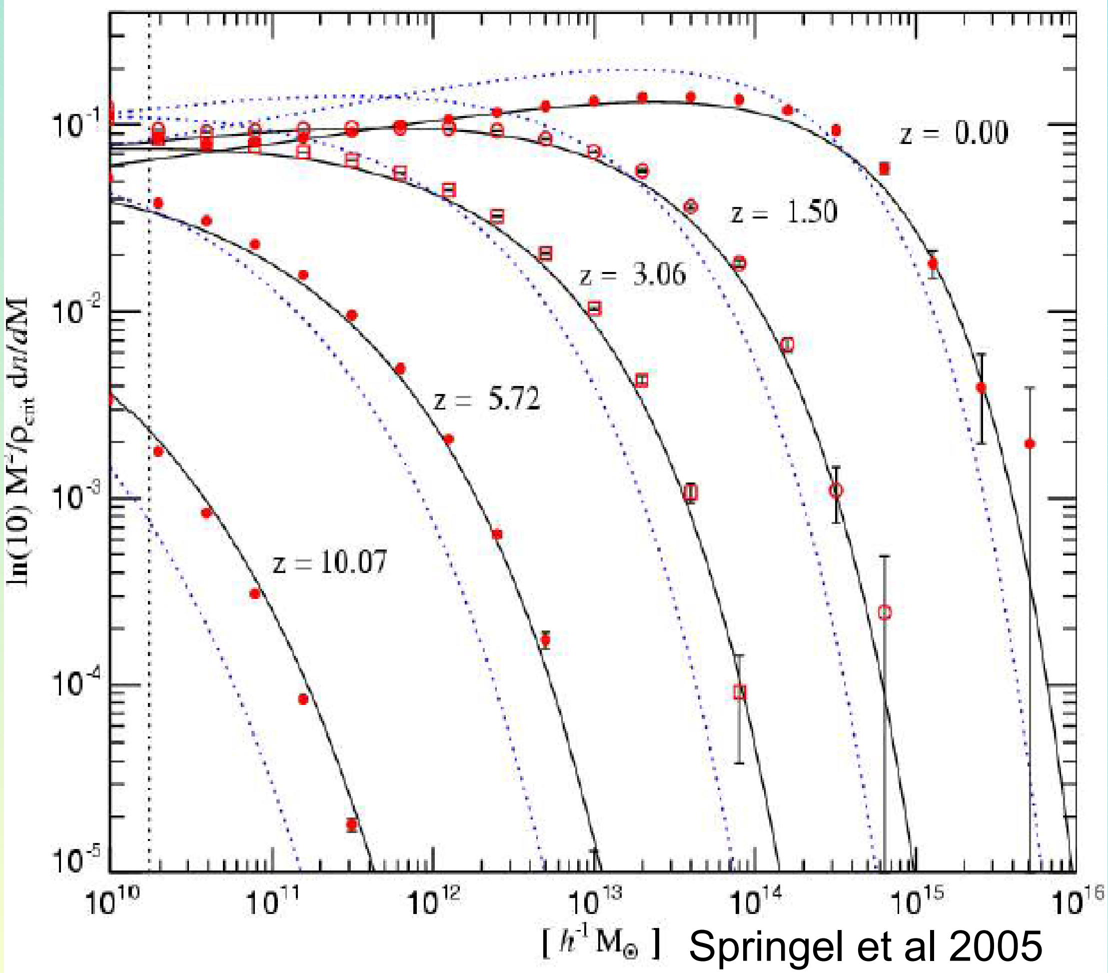
DM distribution in the universe at the z=0, from the *Millennium Simulation* (Virgo consortium), the largest N-body simulation carried out so far (» 10^{10} particles).

The dynamic range of the simulation is 10^5 per dimension.

Problem: it has been run with σ_8=0.9, which now seems » 10% too high!



•The mass function of halos (HMF) is well described by simple formulae that may be derived from approximated analytical treatments, and used in most SAM of galaxy formation
•The DM halos mass function is peaked to larger and larger masses going to lower z



Springel et al 2005

Points: Millenium Run
Solid: Sheth & Tormen
Dotted: Press & Schechter

Radial profiles: density profiles of simulated halos well fit by a "*universal profile*", not dependent on the particular cosmology. Popular fitting formulas:

$$\rho(r) = \frac{\rho_o}{\left(\frac{r}{r_s}\right)^{2-\alpha}\left(1+\frac{r}{r_s}\right)^{1+\alpha}}$$

With α=1 (Navarro, Frenk and White, 1997 NFW) or α=0.5 (Moore et al 1999; Ghigna et al 2000)

With α=1 (Navarro, Frenk and White, 1997 NFW) or α=0.5 (Moore et al 1999; Ghigna et al 2000)

Change of power law slope from -3 when $r>>r_s$ to 2- α when $r<<r_s$ could be related to two phase formation

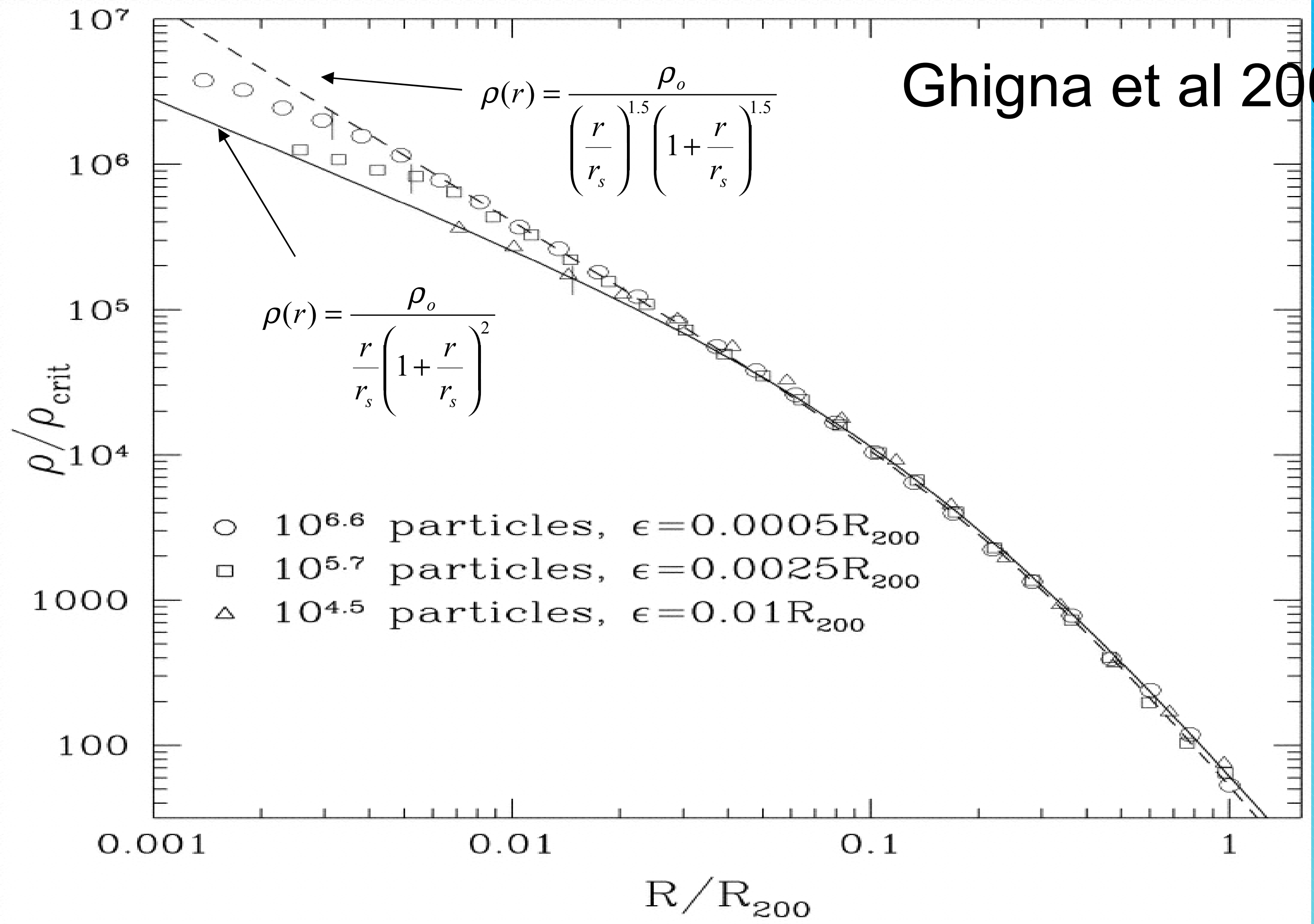
Problem: NFW too much centrally concentrated wrt observed dynamics of spiral and dwarf galaxies, showing cored profiles

This (or similar) result is used in many SAM of galaxy formation, affecting for instance the cooling-collapse recipes, sizes of galaxies etc.

$$\rho(r) = \rho_{-2} \exp\left(\frac{-2}{\beta}\left[\left(\frac{r}{r_{-2}}\right)^{\beta} - 1\right]\right)$$

Other possible choice: Einasto profile

Ghigna et al 2000



With increasing resolution, the cluster's profile continues to approach the curve of Moore et al 1999.

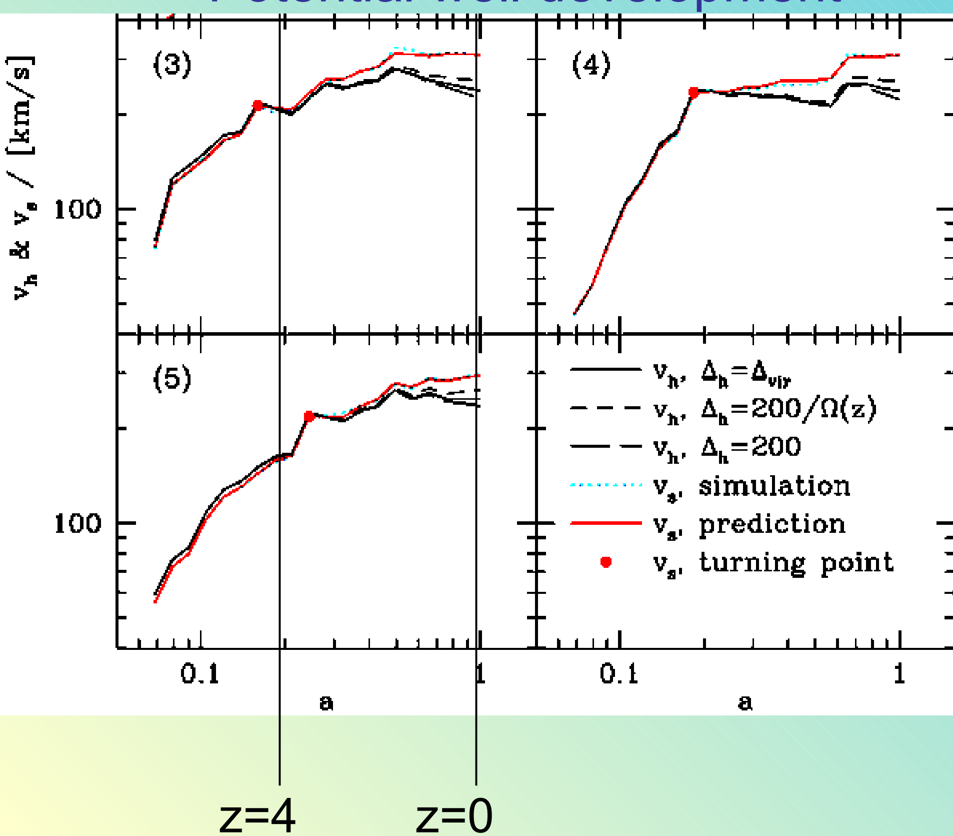
N-body results *(Wechsler et al 2003; Zhao et al. 2003)* indicate a two phase build up of large DMH, which may have consequences for galaxy formation:

1. Fast accretion (» 1 Gyr) by sudden merge of many similar clumps, during which the final potential well is set;

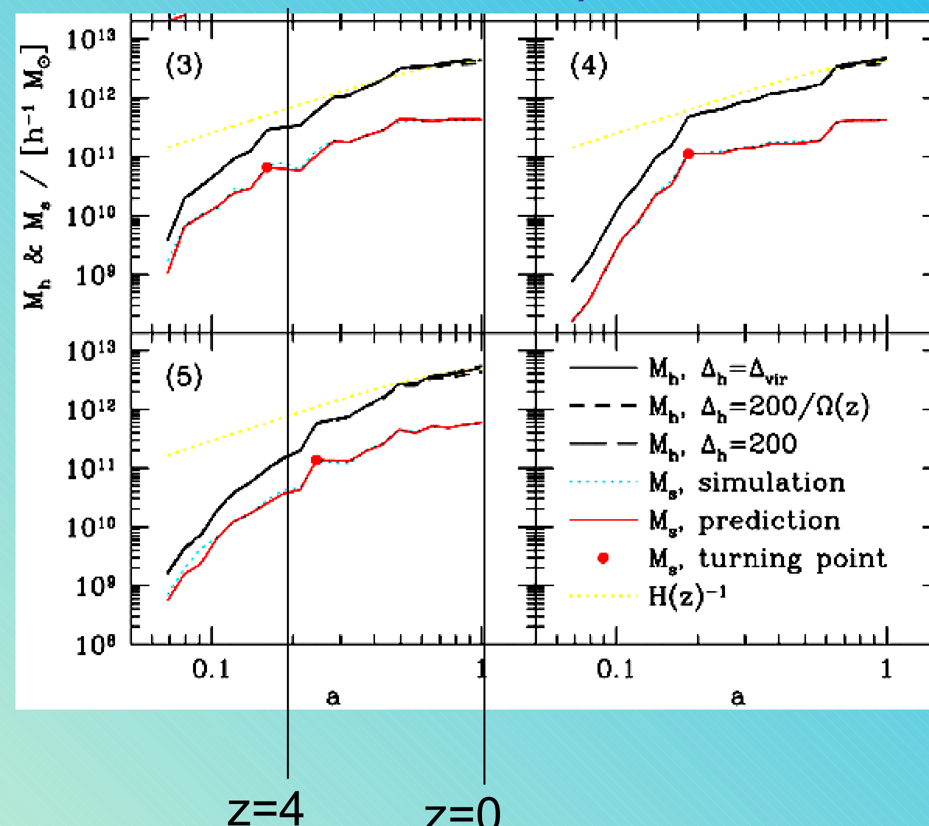
2. Slow accretion of matter in the outskirt (» 10 Gyr), hardly affecting the central region (galaxy?)

Possibly related to the double power law NFW density profile (Li et al 2006).

Potential well development



Mass development



Moore et al 99

5e14 M- ; 2000 kpc (cluster)

2e12 M- ; 39 kpc (galactic)

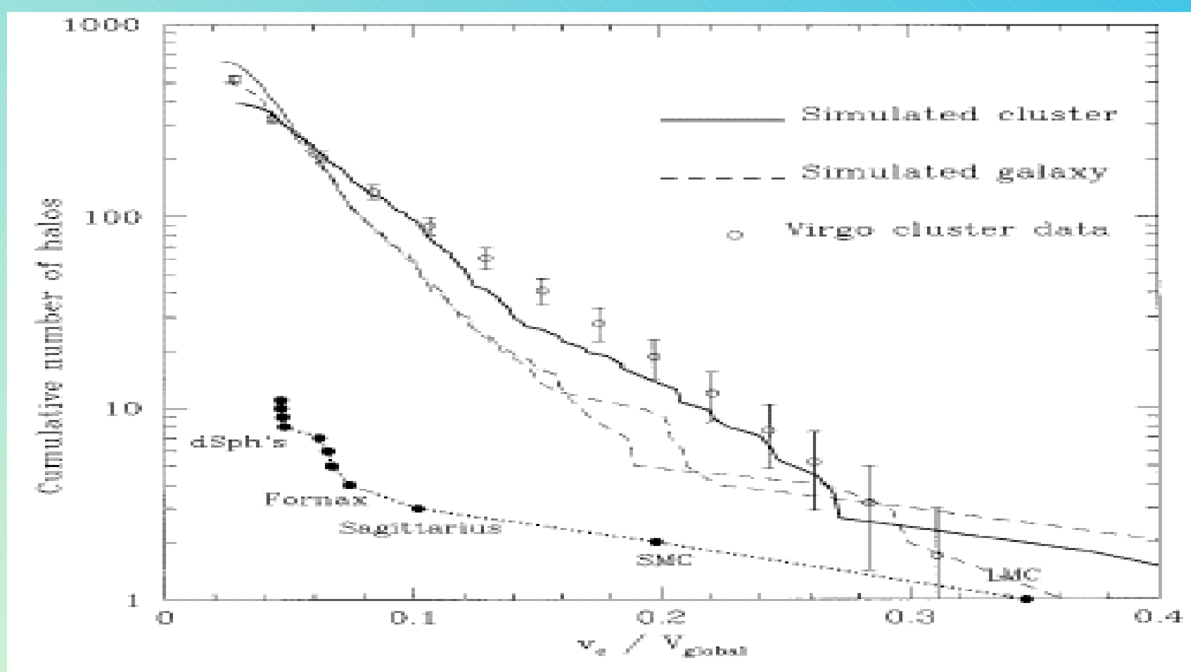
Substructure problem

In DM simulation galaxy halos appear as scaled versions of galaxy clusters, at odd with MW satellites observations.

•Data incompleteness? Indeed SDSS data alleviated problem

•A fundamental problem (eg WDM, less power at small scales)?

•Gas physics and feedback hide 95% of the Milky Way's satellites?



2. Hydrodynamical simulations (without gastrophysics)

DM only simulation shows where the bulk of matter is but not where and how galaxies shine.

To do this, a lot of complex physics should be added, but there are severe practical limitations:

•dissipative collapse of baryon produce very small and dense small clumps) prohibitive dynamical range;
•we lack a theory of star-formation, put into simulations using very simpleminded recipes (and in any case below resolution);
•the energy input from stars (stellar feedback) is essential but again can be introduced only with sub-grid recipes. Not to speak about AGNs

On the positive hand hydrodynamic forces are local

Gas dynamic equations written in two ways, yields to two classes of numerical methods

	Eulerian coord. (Mesh methods):	Lagrangian coord. (SPH methods):
mass:	$\frac{\partial\rho}{dt}+\vec{\nabla}\cdot(\rho\vec{v})=0$	$\frac{d\rho}{dt}+\rho\vec{\nabla}\bullet\vec{v}=0$
momentum:	$\frac{\partial\vec{v}}{dt}+(\vec{v}\cdot\nabla)\vec{v}=-\vec{\nabla}\phi-\frac{\vec{\nabla}P}{\rho}$	$\frac{d\vec{v}}{dt}=-\vec{\nabla}\phi-\frac{\vec{\nabla}P}{\rho}$
energy:	$\frac{\partial u}{dt}+\vec{v}\cdot\vec{\nabla}u=-\frac{P\vec{\nabla}\cdot\vec{v}}{\rho}+\frac{\Gamma-\Lambda}{\rho}$	$\frac{du}{dt}=-\frac{P\vec{\nabla}\cdot\vec{v}}{\rho}+\frac{\Gamma-\Lambda}{\rho}$

Cooling rate

System closed by Poisson equation and state equation (e.g.)

$$\nabla_x^2\phi=4\pi Ga^2\delta\rho(\vec{x},t)$$

$$P=(\gamma-1)\rho u$$

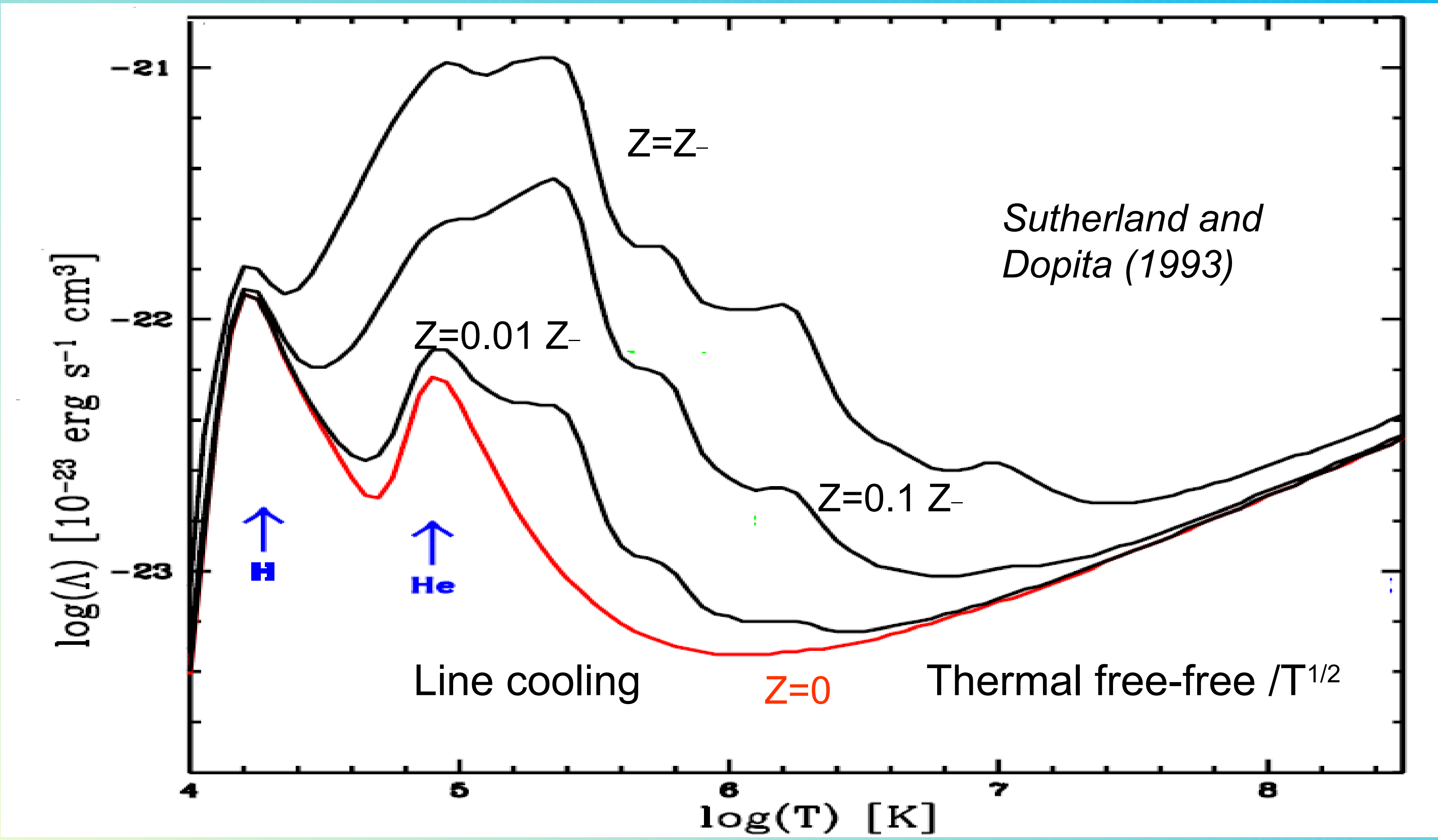
Pre-galactic evolution of baryons

1. Dragged by the gravitational potential of DM to accrete in DMH
2. Kinetic energy thermalized by shocks (heating)
3. Thermal energy radiate away (radiative cooling):
   - Atomic cooling (line and free-free, two body / n^2)
   - Compton cooling on CMB (one body/ n; z & 6)
   - Molecular line cooling (important in very small=cold halos)

Cooling is included in hydro-dynamic equation of energy by means of pre-computed (e.g. with CLOUDY) cooling-function Λ:

$$\frac{du}{dt} = -\frac{P \nabla \cdot \vec{v}}{\rho} + \frac{\Gamma - \Lambda}{\rho}$$

cooling functions

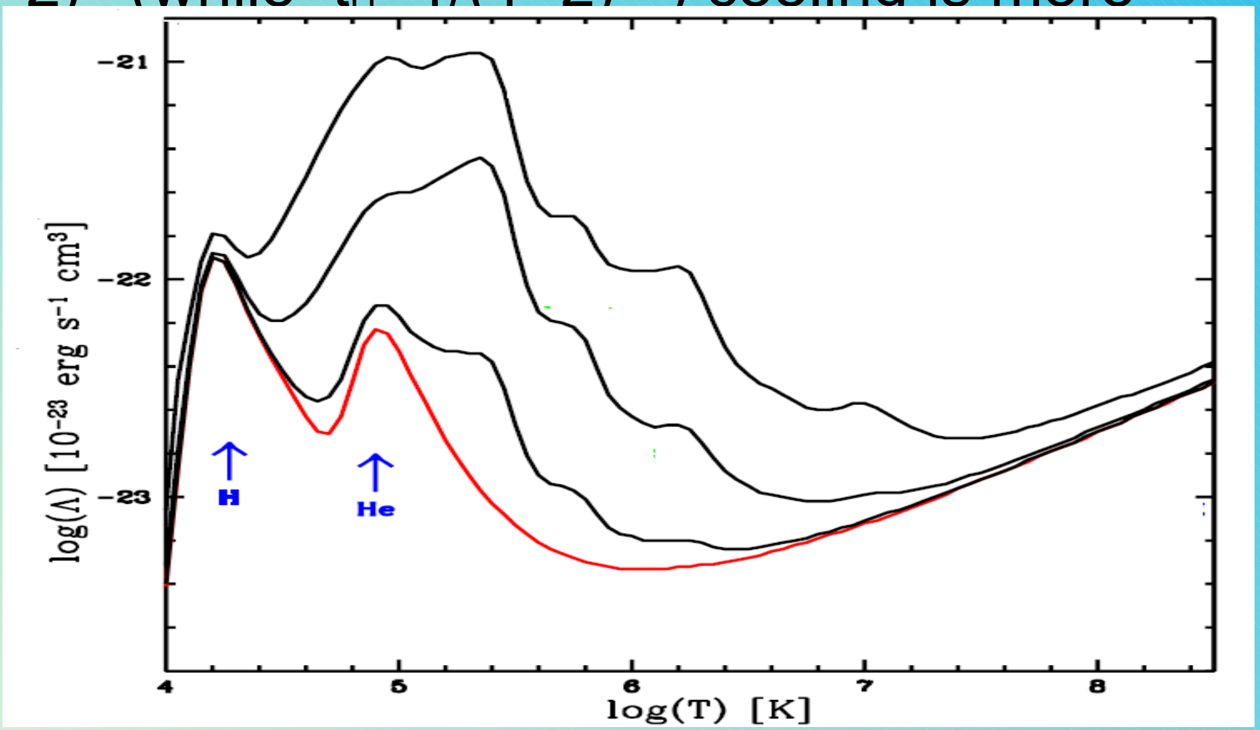


A decent treatment of chemical feed-back from stars would be crucial

General features of cooling

- Atomic cooling cuts off for $T< 10^4$ K) t_{cool} becomes very long
- For $T>10^6$-10^7 K, cooling is dominated by bremsstrahlung,

  $t_{cool} \sim nT/(n^2 T^{1/2}) \sim T^{1/2}/n$ - increases with T (or Mass)
- Then cooling is most rapid for intermediate $T\sim10^4$-10^6 K (for fixed n; say M_{vir} »1e9 -a few 1e11)
- Since $t_{cool}\sim1/n\sim1/(1+z)^3$ (while $t_H\sim1/(1+z)^{3/2}$) cooling is more effective at high-z



Mesh methods (classical) discretize the PDEs on a mesh and solve the corresponding finite difference equations.

In most popular implementations (TVD and PPM), the eulerian fluid equations are written in form of conservation laws for the various quantities:

densities → fluxes

$$\frac{u^i}{t} = -\frac{F^{ij}(u)}{x^j}$$

Then fluxes are computed across cell boundaries to update the cell averaged density over a timestep.
Modern codes use Adaptive Mesh Refinement.

Advantages (wrt SPH):

- Faster
- periodic boundary conditions automatically implemented.
- superior with shocks and turbulence

SPH (Smooth Particle Hydrodynamics) is a Lagrangian (particle tracking) method: the fluid elements are represented by fixed mass particles, characterized by the fluid variables (baryon density, velocity, temperature, etc). The hydrodynamics equations rewritten in terms of forces acting on these particles.

A generic fluid variable *f* is evaluated for any particle as a smoothed estimate, i.e. a smoothed sum over nearby particles, using an interpolating function or kernel *W*.

$$< f_i > = \sum_j \frac{m_j}{\rho_j} f_j W(|\vec{x}_i - \vec{x}_j|; h)$$

The kernel is a function of particle distance and depends on a parameter, the smoothing length *h*. It must satisfy:

$$\int_{all\ space} W(r;h)\, d^3x = 1 \qquad \lim_{h \to 0} W(|\vec{x}_i - \vec{x}_j|; h) = \delta(|\vec{x}_i - \vec{x}_j|)$$

The mathematically ideal kernel would be a gaussian, but in practice is better to use an algebraic approximation with compact support ($\xi = r/h$)

$$W(r,h) = \frac{1}{h\sqrt{\pi}} \exp(-r^2/h)$$

$$W(r,h) = \begin{cases} 1 - \frac{3}{2}\xi^2 + \frac{3}{4}\xi^3 & 0 \quad \xi \quad 1 \\ \frac{1}{4}(2-\xi)^3 & 1 < \xi \quad 1 \\ 0 & 2 < \xi \end{cases}$$

Best to have $h/\rho^{-1/3}$, adjusted to keep »10 particles in the sum.

There are well studied recipes to translate the physical equations into SPH formalism. Gradients of physical quantities are replaced by gradients of the smoothing kernel, known and analytical:

$$\nabla f_i = \sum_j \frac{m_j}{\rho_j} f_j \nabla_i W(|\vec{x}_i - \vec{x}_j|; h)$$

Thus for instance:

$$\frac{d\vec{v}_i}{dt} = - \sum_j m_j \left(\frac{p_j}{\rho_j^2} + \frac{p_i}{\rho_i^2} \right) \nabla_i W(|\vec{x}_i - \vec{x}_j|; h)$$

Momentum eqz.

Advantages

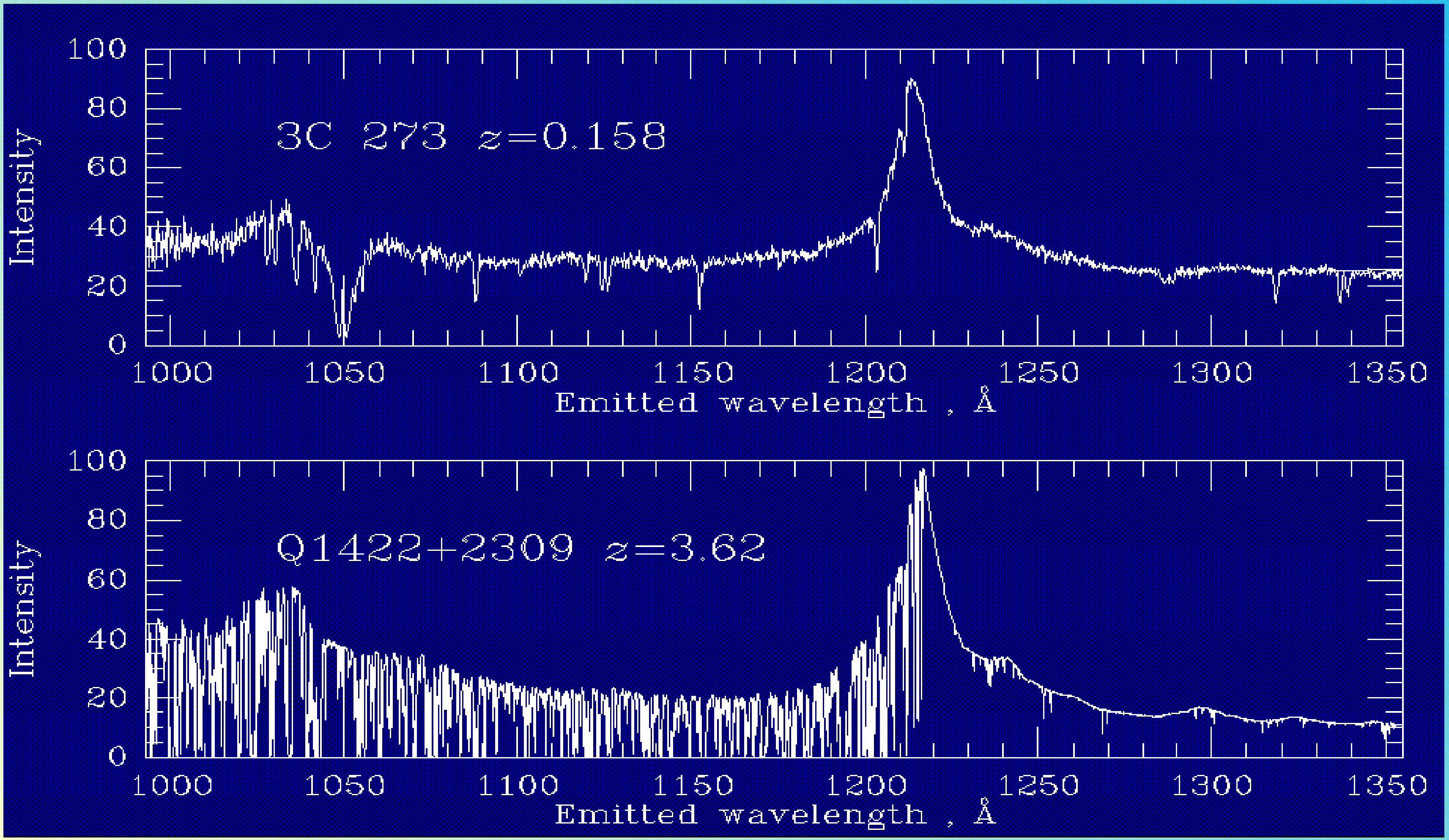
•Easy to combine with N-body gravity methods (it's a sort of extension);

•Spatial resolution automatically increased in denser regions;

SPH is the most used in AP, but the optimal choice is problem dependent.

Resolution not sufficient to study galaxy formation.

Interesting results in the study of properties of gas which has not yet collapsed to form galaxies, a less non-linear problem. Neutral H in IGM produces absorption features in QSO spectra, the Lyα forest



An empty selection of results from hydro-simulations without sub-grid physics the distribution of column densities is quite well reproduced

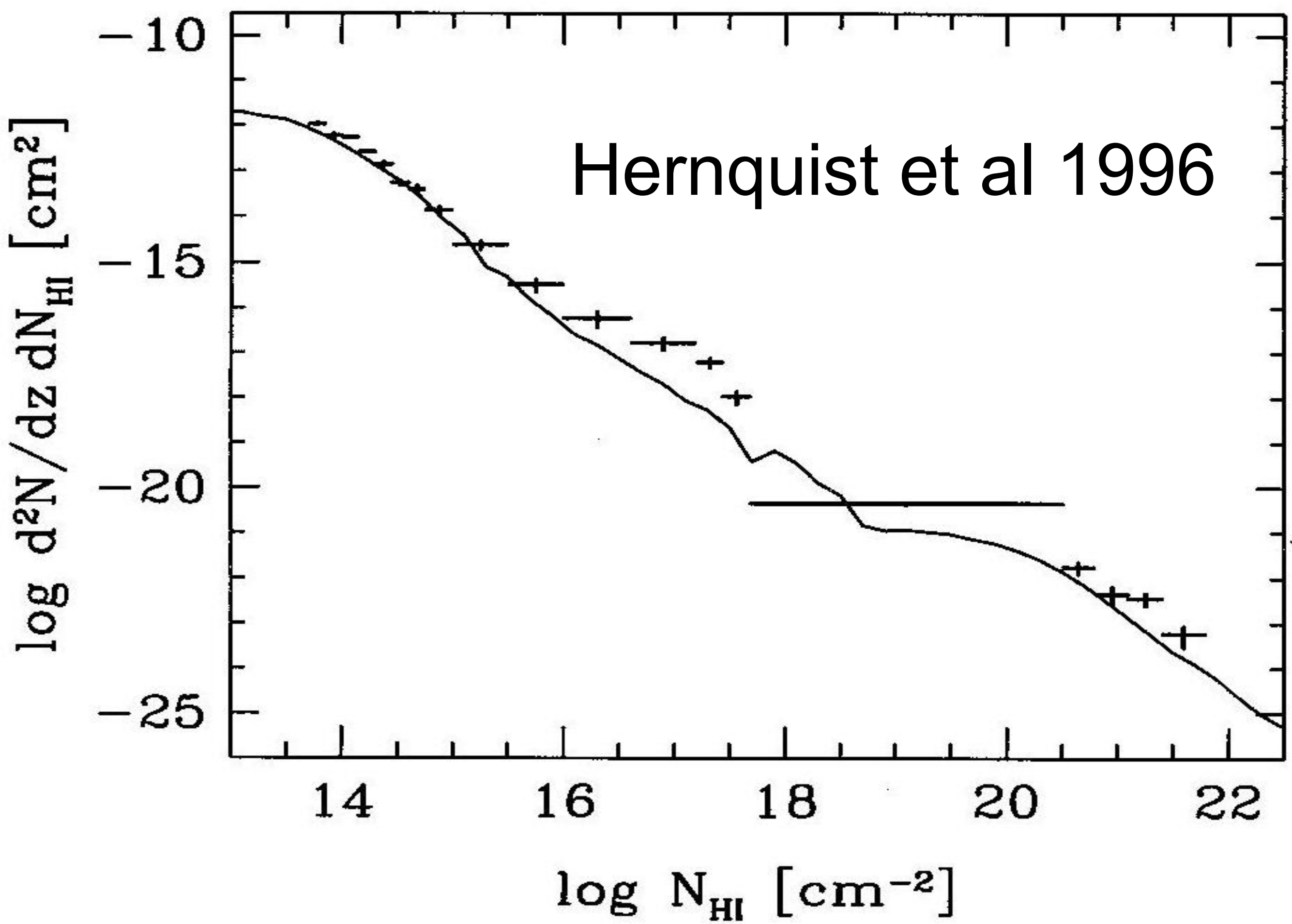


FIG. 2.—Distribution of neutral hydrogen column densities. The solid line shows the simulation results at $z = 2$. Points with error bars are taken from Petitjean et al. (1993); we multiply their values by $1 + z = 3$ to convert from number of lines per "absorption distance" interval ΔX to number of lines per redshift interval Δz.

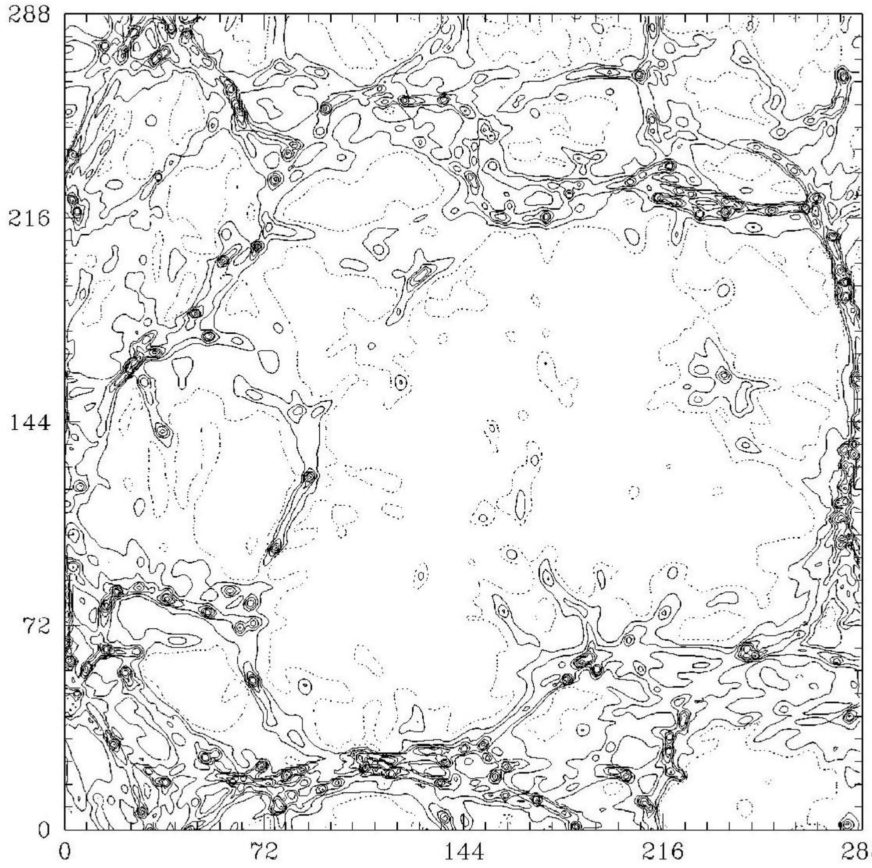
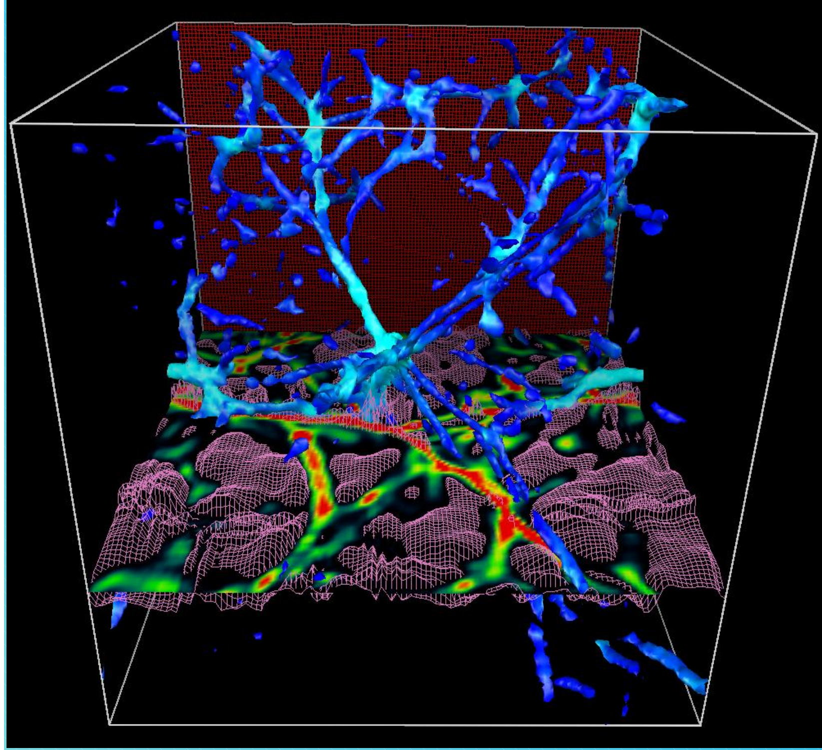


Figure 2 H I column density contours for a slice of the 10 h^{-1} Mpc (comoving) box from a cold dark matter model with a nonzero cosmological constant Λ (ΛCDM) by Miralda-Escudé et al (1996).



Gas density at z=3 from a hydrodynamics simulation of the Lya forest. The box is 2.4 Mpc. The surfaces represent baryons at 10 the mean ***ρ*** (typical filamentary structures) and are color coded to the gas T (dark blue = 3 10^4 K, light blue = 3 10^5 K). (Zhang et al.)

log *N*(H I)<14 sheet-like
14<log *N*(H I)<16 in filaments
log *N*(H I)>16 are spherical

3. Hydrodynamical simulations with phenomelogical treatment of sub-grid physics

To simulate formation of galaxies substantial sub-grid physic is required.

Most notably star formation and its feedbacks

More recently recognized, also SMBH growth and AGN feedback

Star formation

A crucial ingredient in any galaxy formation models. The problem is not only that the relevant scale is well sub-grid, but also that it is poorly understood. Typical prescriprion is (Katz 92):

$$SFR = c_* \frac{\rho_{gas}}{\tau}$$

Efficiency ~ parameter adjusted typically to a few %

Local dynamical timescale / ρ $^{-0.5}$

Provided that some conditions are satisfied (Cen & Ostriker 92)

$$\rho_{gas} > \rho_{threshold} \quad \text{(dense enough)}$$

$$\mathrm{T_{gas}} < \mathrm{T_{threshold}} \quad \text{(cold enough)}$$

$$\mathrm{v_{gas}} < 0 \quad \text{(contracting)}$$

The masses of collisional (gas) and collisionless (stars) simulation “particles” (a few 10^6 stars) are evolved accordingly

$$m_* = m_{gas}\left(1 - \exp\left(-\frac{c_* \Delta t}{\tau}\right)\right)$$

Stellar feedback

Energetic and chemical feed-back from stars are also included with naïve approximations (when included), e.g. over a timestep:

$$R\,\Delta m_* \text{ returned to gas}$$

$$y\,\Delta m_* \text{ returned to gas as metals}$$

$$\varepsilon\,\Delta m_* c^2 \text{ injected as thermal and/or kinetic energy}$$

Wherein stellar lifetimes are neglected (Instantaneous Recycling Approx.), greatly reducing computational demand.

R, y and ε depends on the adopted IMF.

Leaving IRA, they depend also on age of stellar populations and thus a convolution integral over past star formation history is required. This complicates computations, attempted only in a few recent papers (Kobayashi et al. 2007; Tornatore et al et al. 2007).

SMBH growth and its feedback

Introduced in some hydro sim since » 2005 to compare also with AGN-SMBH populations and to account their feedback on galaxy formation.

SMBH treated as a sink particle, accreting mass at minimum between Bondi accretion rate (rough prescription to ensure fuel availability) and Eddington accretion rate (rough prescription to ensure gravity overcame rad pressure)

$$\dot{M}_{BH} = \min(\dot{M}_{bondi}, \dot{M}_{Edd});$$

$$\dot{M}_{bondi} = \alpha \frac{4\pi (G M_{BH})^2 \rho_{gas}}{(c_s^2 + v_{BH-gas}^2)^{3/2}}; \quad \dot{M}_{Edd} = \frac{4\pi G M_{BH} m_p}{\eta_{rad} \sigma_T c}$$

SMBH growth and its feedback

a fraction ε ' 5% of the radiative energy produced

$$L = \eta \dot{M}_{BH} c^2$$

is assumed to be thermally deposited in the surrounding gas.

ε controls the normalization of M_{BH}-σ or M_{BH}-M_{sph} relations and is adjusted to fix observations (also α).

Kinetic feedbacks (jets, BAL) are not explicitly included

SMBH growth and feedback in hydro-simulations I

- Since 2005 this process began to be “gastro-physic included” in HS, in particular the wide-spread public code GADGET2 by Springel et al.
- SMBH modeled as collision-less sink particle that can accrete gas from their surroundings.
- First technical problem: the SMBH can “materialize” suddenly only when $M_{SMBH} >> M_{gas\ particle}$, otherwise the kick due to momentum absorption could even (artificially) eject the SMBH from the galaxy.
- Typically a “seed” SMBH is placed at the center of a galaxy when it surpass some mass threshold depending on resolution. Even in the best cases $M_{seed} > 10^5$-$10^7\ M_{\odot}$ i.e. a large IMBH!
- Then the SMBH begins to growth by Eddington-limited Bondi accretion:

$$\dot{M}_{BH} = \min(\dot{M}_{bondi}, \dot{M}_{Edd}); \quad \dot{M}_{bondi} = \alpha \frac{4\pi (G M_{BH})^2 \rho_{gas}}{(c_s^2 + v_{BH-gas}^2)^{3/2}}; \quad \dot{M}_{Edd} = \frac{4\pi G M_{BH} m_p}{\eta_{rad} \sigma_T c}$$

- In Bondi treatment (spherical simmetric!) α(dimensionless) depends on gas equation of state and should be of order 1. Instead a number >100 (!!) is used, “justified because it produces a reasonable black hole mass at the end of the simulation”

SMBH growth and feedback in hydro-simulations II

- Two BH particles are assumed to merge when their separation has fallen to the spatial resolution
- A fraction »5% of radiative power $\eta dM/dtc^2$ is coupled (only thermally and isotropically) with the nearby SPH particles. The fraction chosen to match observed the $M_{BH} - \sigma$ correlation.
- These crude prescriptions have been used in tens of papers since Di Matteo et al 2005, to investigate the effects of “quasar mode” AGN feedback from merging of pairs of galaxies to cosmological volumes (e.g. Di Matteo et al 2008).
- More recently (e.g. Sijacki et al 2008) it has been considered also the “radio mode” AGN feedback. When the accretion rate falls below » 1% of the Eddington rate, ALL the accretion energy is assumed to power a radio bubble.

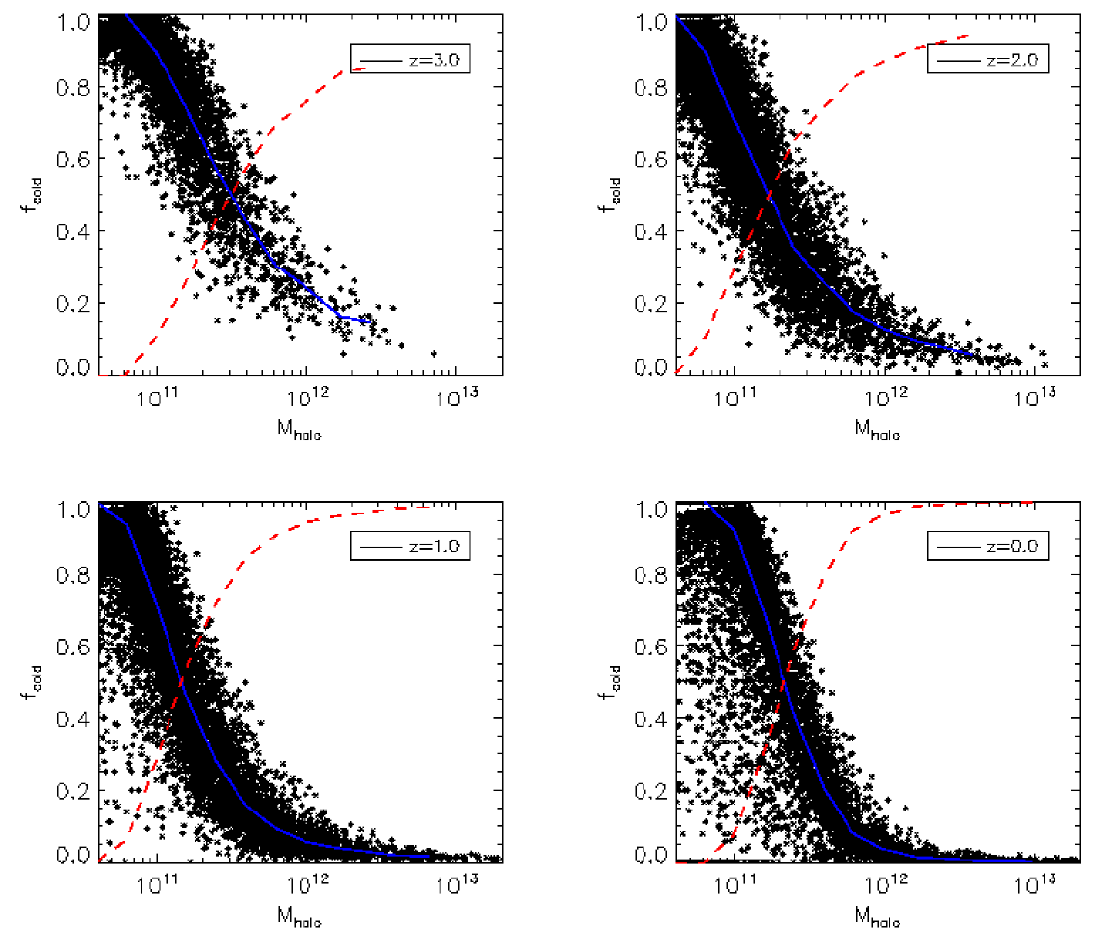
Most popular hydrodynamic codes and physical processes included

Feature	GADGET-3[1]	GASOLINE[2]	HART[3]	ENZO(ZEUS)[4]	FLASH[5]
Gravity	Tree	Tree	Tree	Multi-grid	Multi-grid
Hydrodynamics	SPH[6]	SPH[6]	AMR[7]	AMR[7]	AMR[7]
→ Multiphase subgrid model[8]	✓[9]	×	N/A	N/A	N/A
Radiative Cooling	✓	✓	✓	✓	✓[10]
→ Metal dependent	✓[11]	×	✓[12]	✓[13]	✓[10]
→ Molecular chemistry	✓[14]	×	✓[12,15]	✓[16]	×
Thermal Conduction	✓[17]	×	×	×	✓
Star formation	✓[18]	✓[19]	✓[12]	✓[20]	×
→ SNe feedback	✓[18]	✓[19]	✓[12]	✓[20]	×
→ Chemical enrichment	✓[18]	✓[19]	✓[12]	✓[20]	×
Black hole formation	✓[21]	×	×	×	✓[22]
→ AGN feedback	✓[21]	×	×	×	×
Radiative transfer	OTVET[23,24]	×	OTVET[23]	✓[25]	✓[26]
Magnetic fields	✓[27]	×	×	✓[28]	✓[29]

A selection of results from hydro-simulations with sub-grid physics

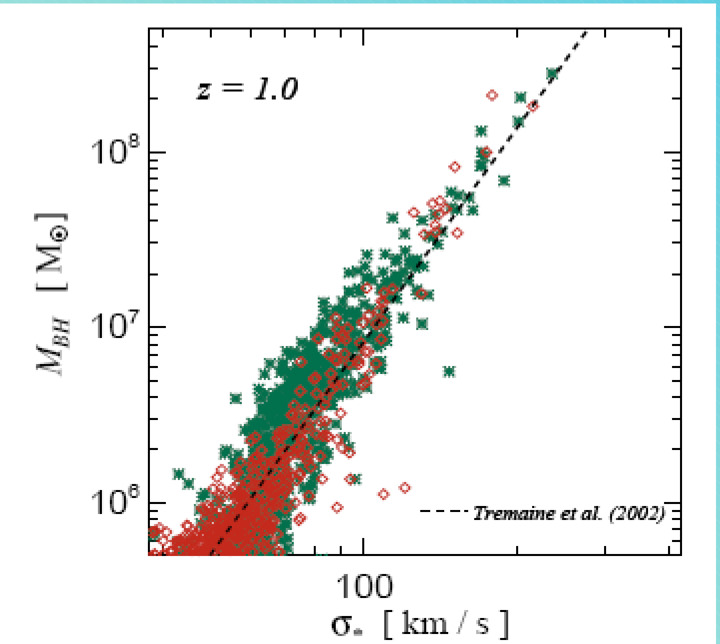
Hydro sim. (with gastrophysics.. Keres et al 2005, 2008) are suggesting revision of recipes of gas accretion used so far in semi analytic models.

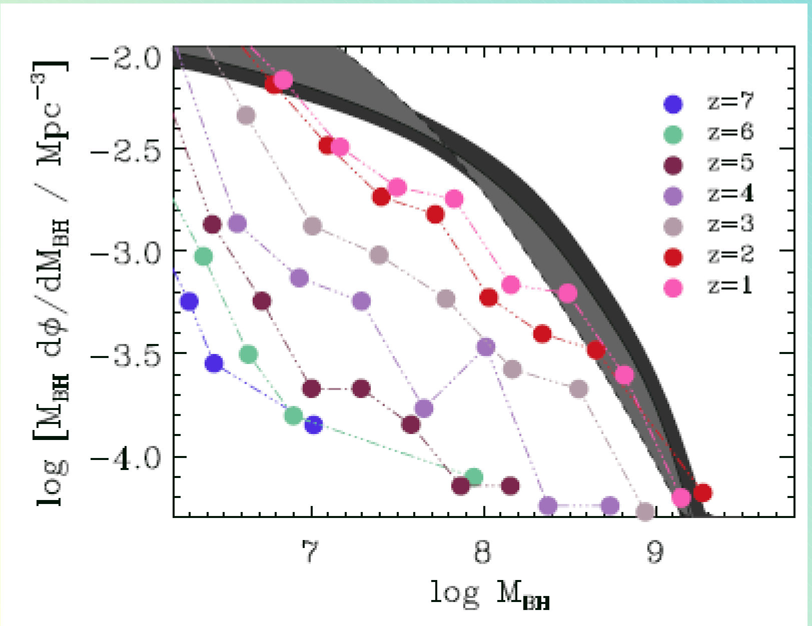
A significant fraction of gas is acquired by galaxies in “cold mode”: flows along filaments avoiding the heating to the virial temperature of the halo.

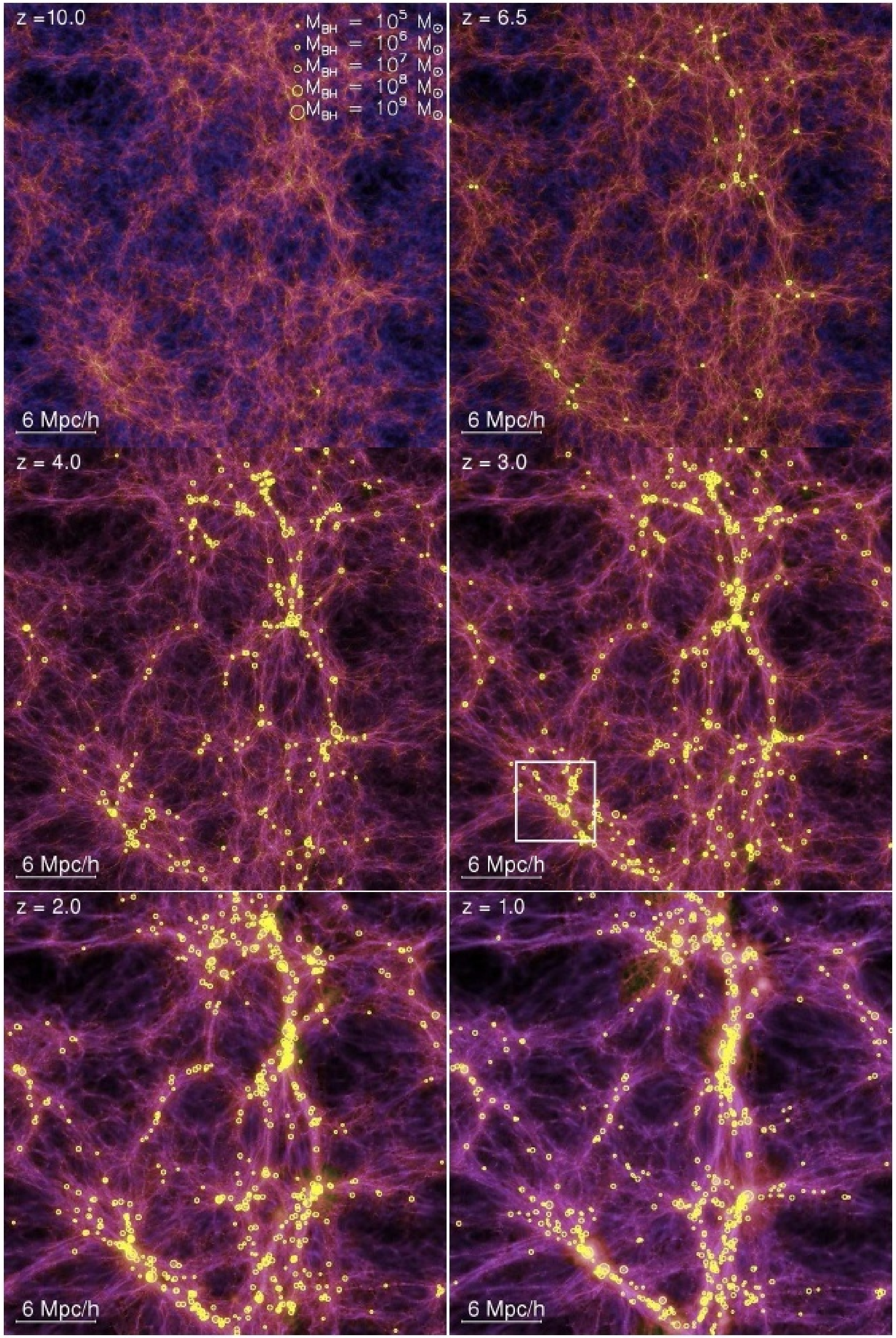


Keres 2008

Di Matteo et al 2008





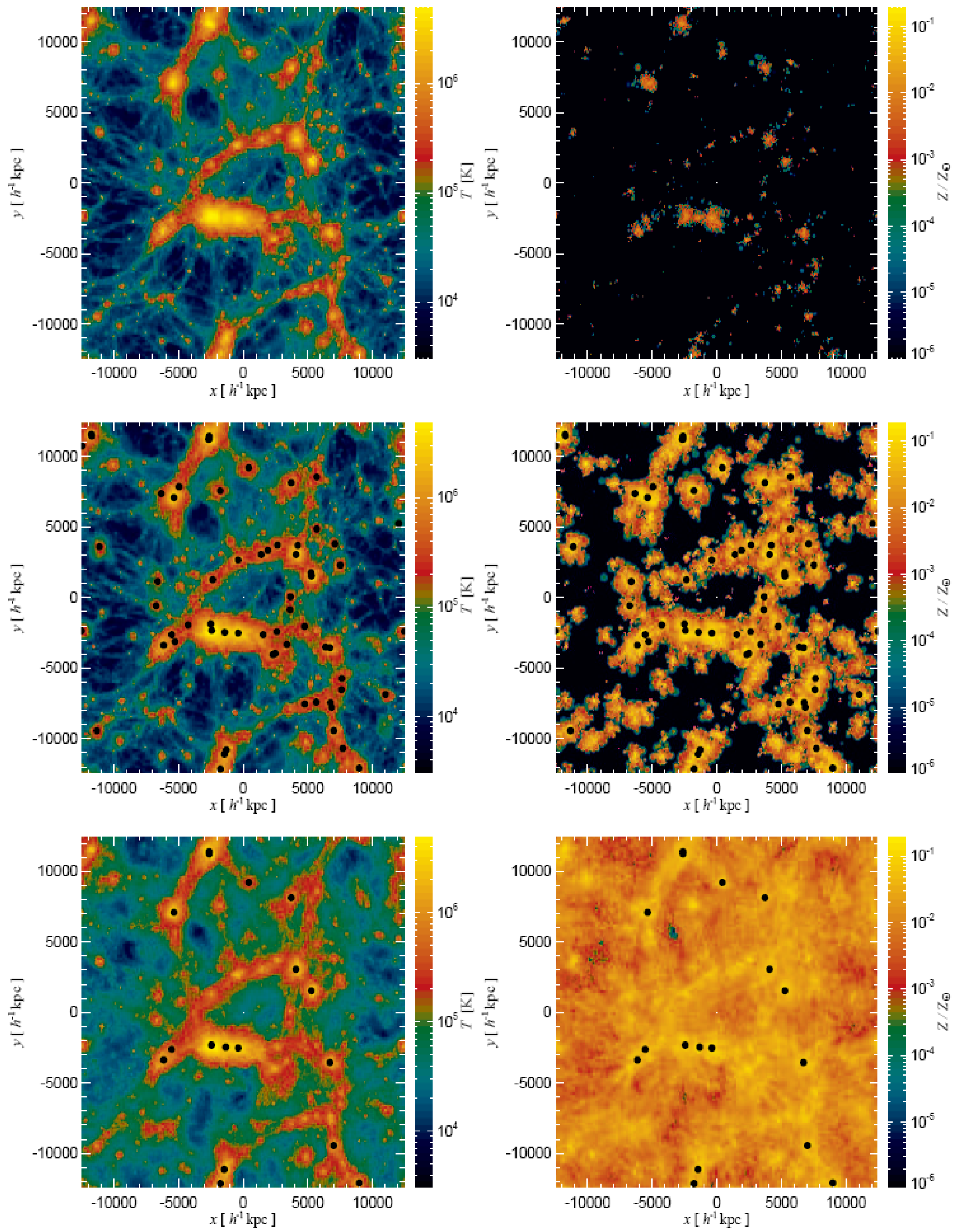


Sijacki et al 2008

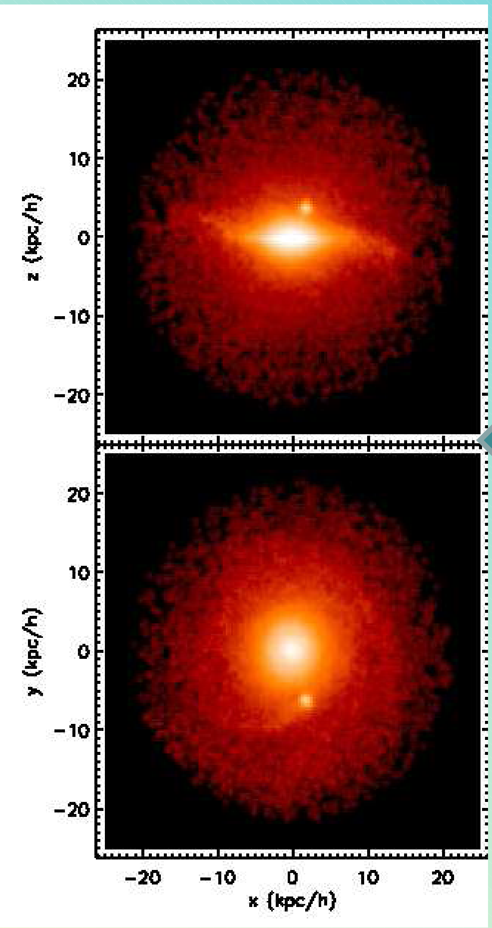
Only cooling and SF

…added BH model

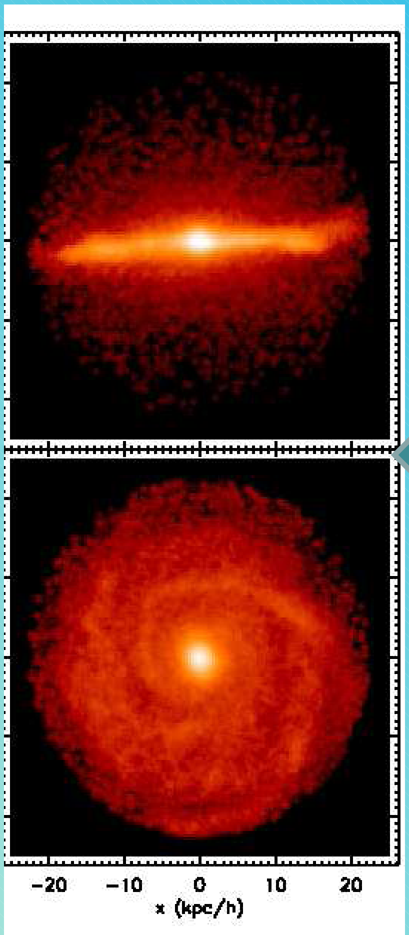
..also galactic wind on



Unsurprisingly, even in simpler non-cosmological simulations, star formation histories and final morphologies are **determined** by the very uncertain sub-grid recipes



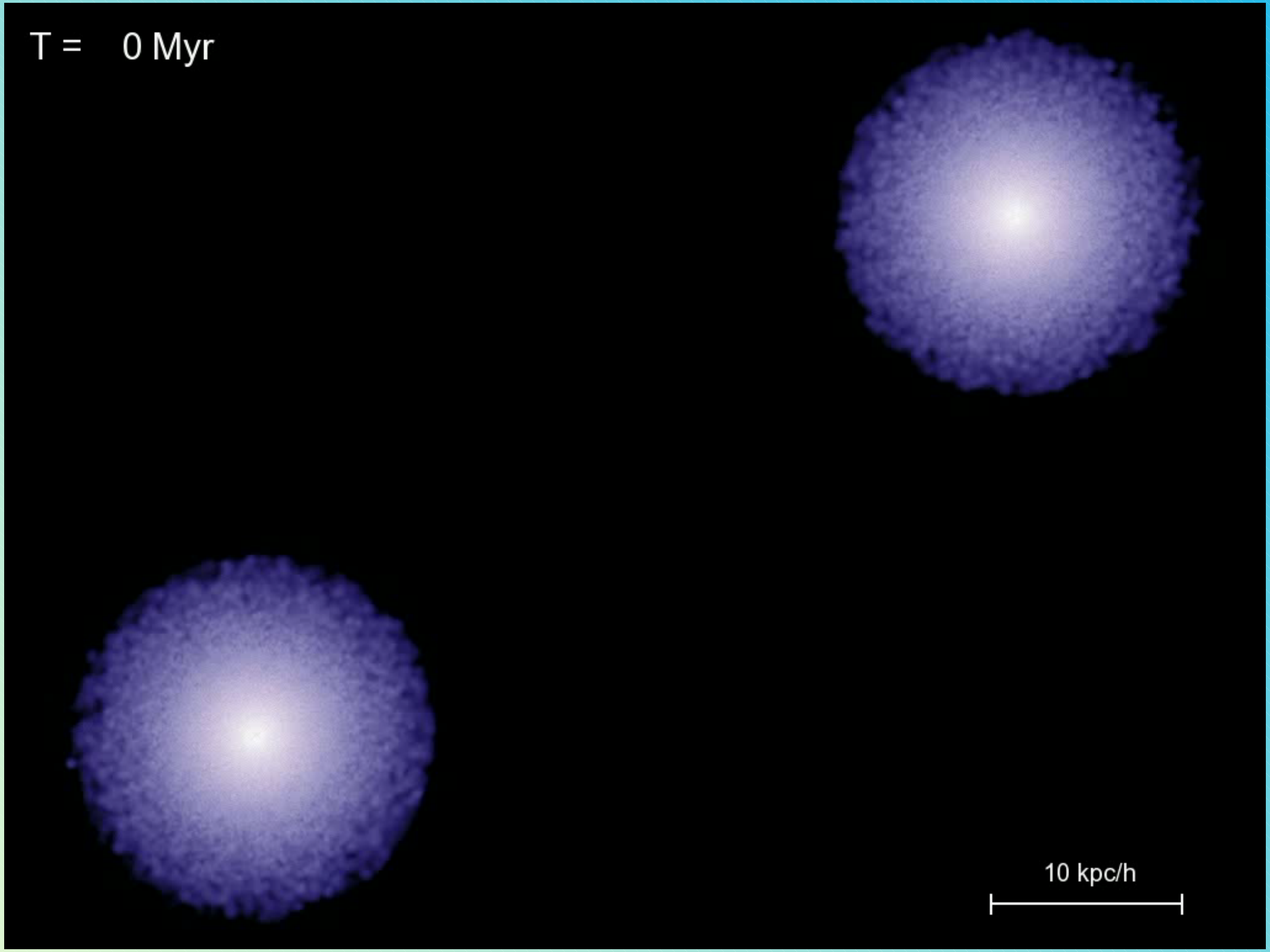
Weaker early feed-back) bulge dominated



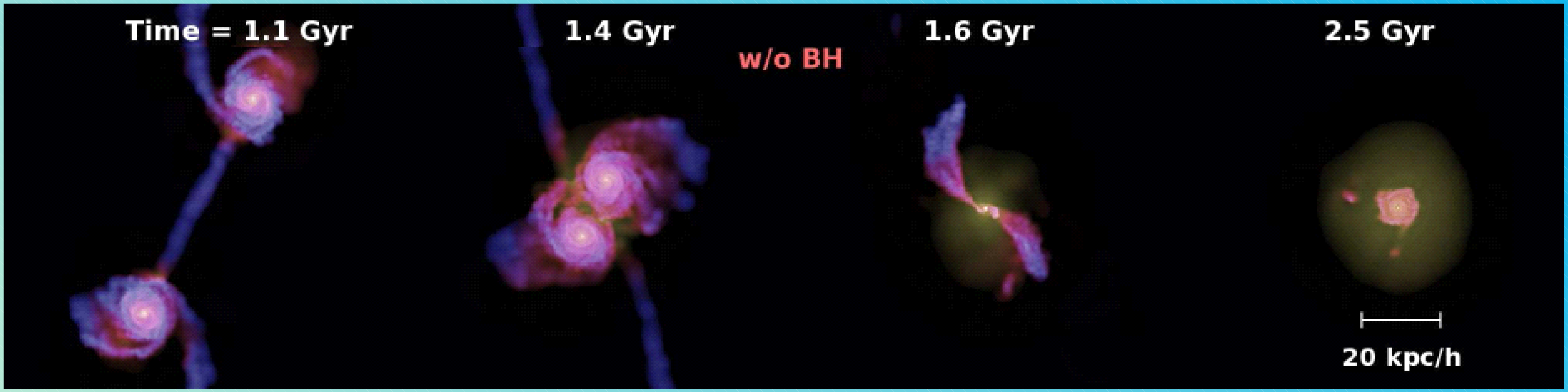
Stronger early feed-back) disk dominated

Example (Zavata et al 07): Simulation of disks formation from identical initial conditions, but different sub-grid physics for feed-back

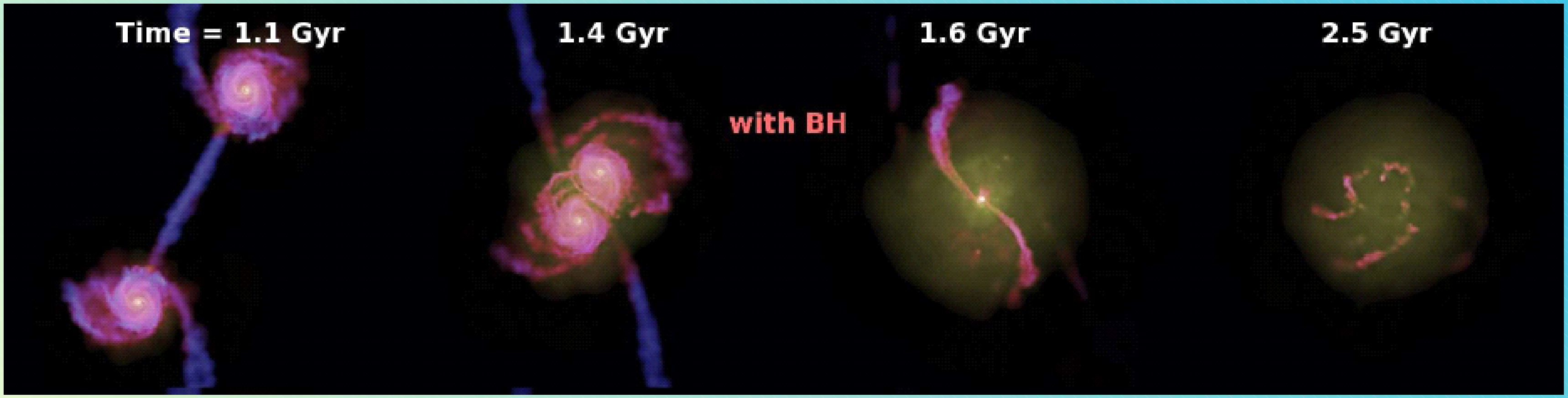
Another example Di Matteo et al 05



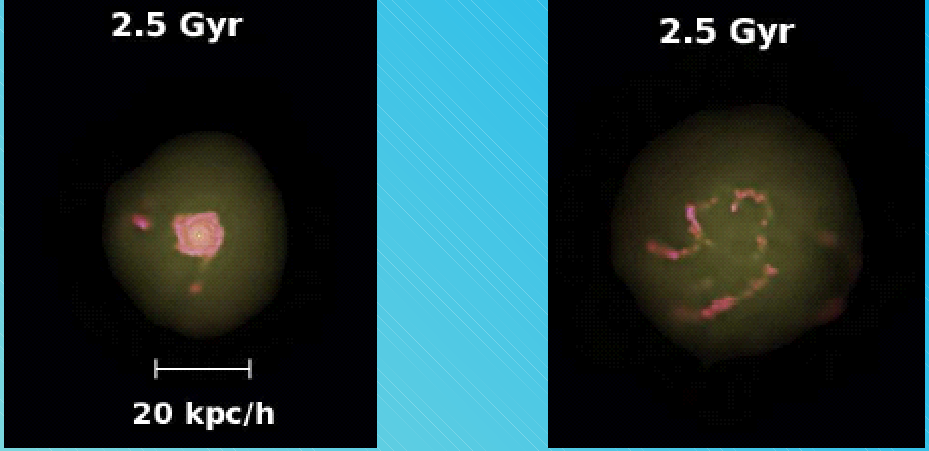
Another example Di Matteo et al 05



Simulation of merging of spirals without treatment of induced QSO activity…



..and WITH (very crude and uncertain sub-grid treatment of) induced QSO activity and ensuing feedback on ISM



Fate of initial gas	Without AGN	With AGN
In stars	89%	52%
Cold SF gas	1.2%	0%
Hot halo gas	9.8%	11.1%
Expelled from halo	0.05%	38%
In SMBH	-	1.6%

The predicted star formation histories and final morphologies are completely different. final morphologies.

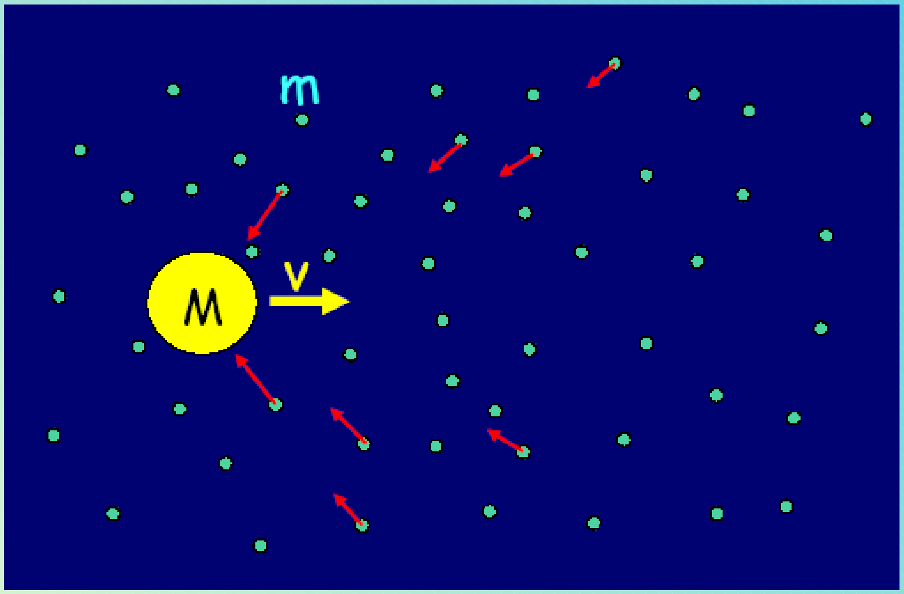
This strong interference of sub-grid prescriptions reinforce rationale for fully sub-grid modelling) Semi Analytic Models

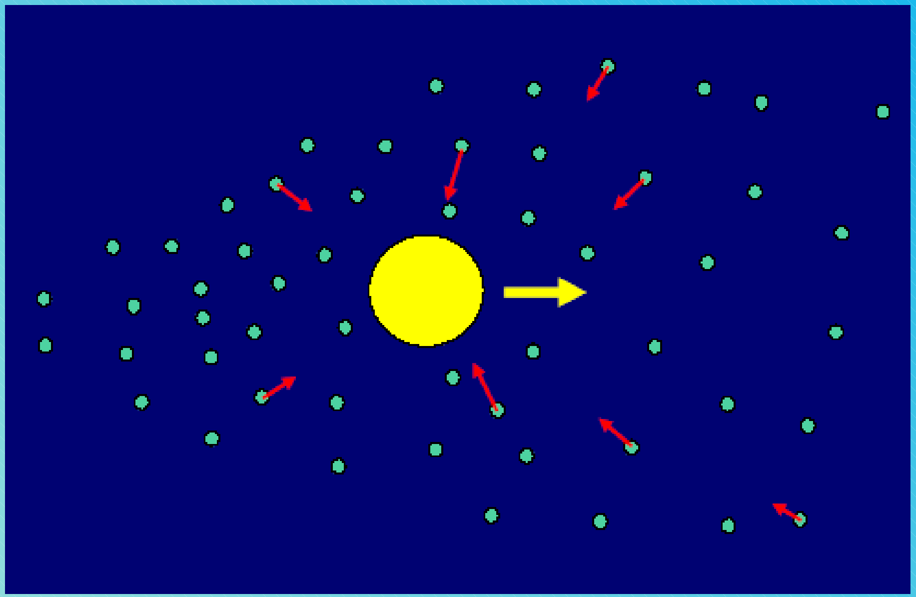
Two main problems affects simulated disk galaxies in CDM scenarios (e.g. Navarro et al 1995, Navarro and Steinmetz 1997

1. Overcooling problem: to much gas cools and forms galaxies: almost all baryons cool and collapse in dense clumps, whilst in the real universe no more than 10% do it.
2. Disk angular momentum problem: simulated disks have angular momentum and corresponding radii about 10% of observed values. It happens because cold gas clumps lose angular momentum to dark matter halos by dynamical friction before merging to form galaxies.

Perhaps problems are connected and can be solved with more realistic treatment of star formation and feedback in simulations (e.g. Maller & Dekel 2002).
2 can also be solved with better mass resolution (Governato et al 2004).
Simulations have also problems in producing old ellipticals

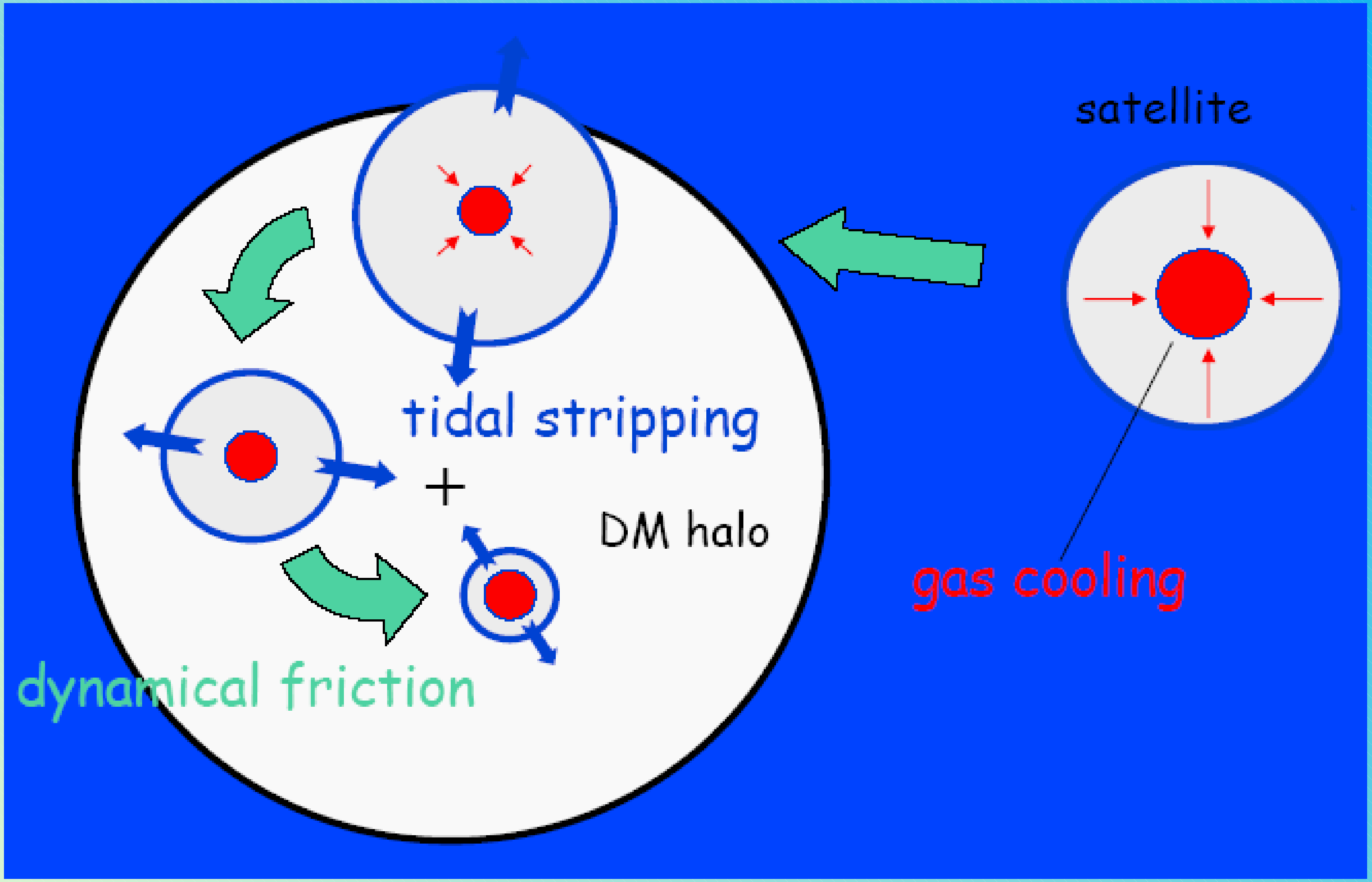
Dynamical friction *m<<M*



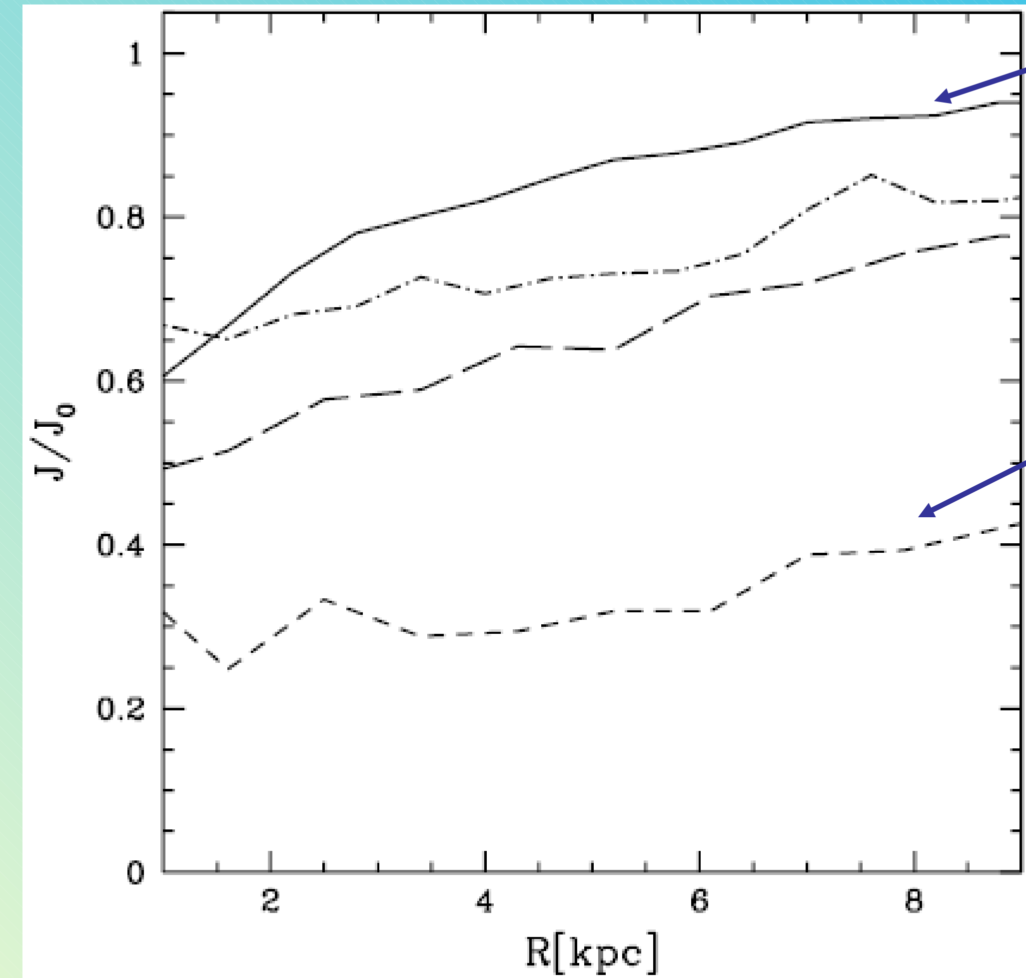


As M moves in a background of much less massive particles, causes a concentration of them past it. Thus it slows down.

If M is orbiting, orbits decay.



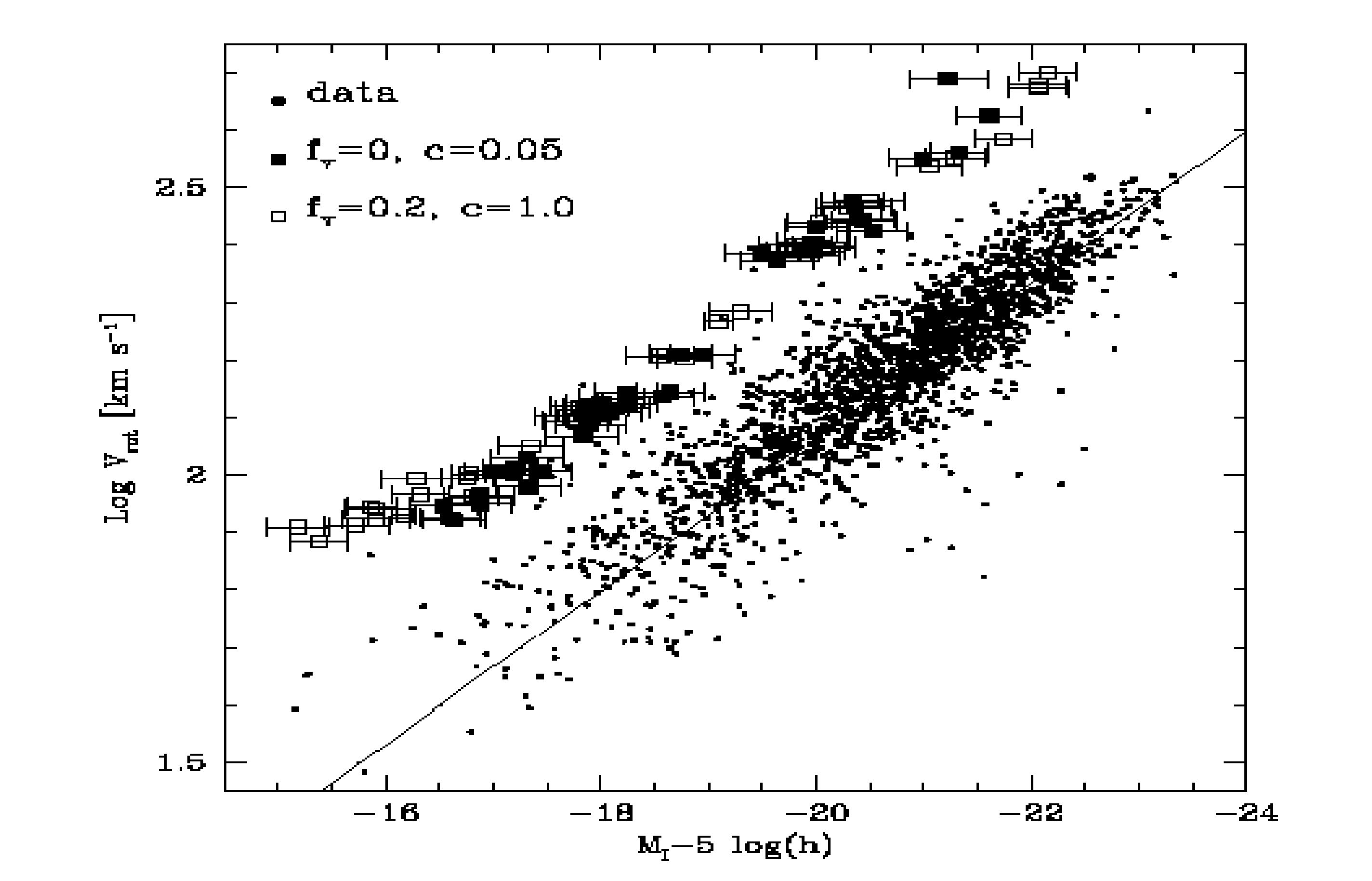
Feedback could save the day, contrasting cooling in small halos which makes gas immune to tidal stripping (Maller & Dekel 2002)



N_{tot} ' 3£ 10^5

N_{tot} a factor 25 less) catastrophic angular momentum loss

Angular momentum loss in a disk galaxy model as a function of radius obtained with progressively decreasing mass resolution, up to a factor 25 (Governato et al 2004)



The slope and scatter of the simulated TF relation are in agreement with the observational data, but the zero-point is in serious disagreement.

But this is connected to the fact that predicted halo profiles (NFW) are too much concentrated

Semi-analytical technique and models (SAMs)

- Extensive comparisons between different scenarios and galaxy data are done by means of fully Semi-Analytic Models (SAM) for baryons.
- The present fashion is post-processing of gravity-only simulations for DM.
- SA technique use simplified analytical descriptions for ALL the gas processes which are thought to be relevant.
- This implies “by definition” an a-priori (more or less physically motivated) choice of processes and analytical forms of relationship between fundamental, typically integrated, quantities of the system (e.g. global masses in certain components, average densities, sizes etc)
- Danger is you get what you put in
- Relationships contain many (tens) fitting free parameter

Semi-analytical technique and models (SAMs)

Most SAMs assume a disk galaxy merger driven sequence of processes leading to present day galaxy populations

1. The outcome of gas cooling in DMH is gaseous rotation supported disk, with mild SF (Rees & Ostriker 1977, Silk 1977, White & Rees 1978….);
2. Disk mergers are the only driver of bursty SF and of the main path for the formation of spheroids (White & Rees 1991, Cole 1991, … omissis…Cole et al 2000).

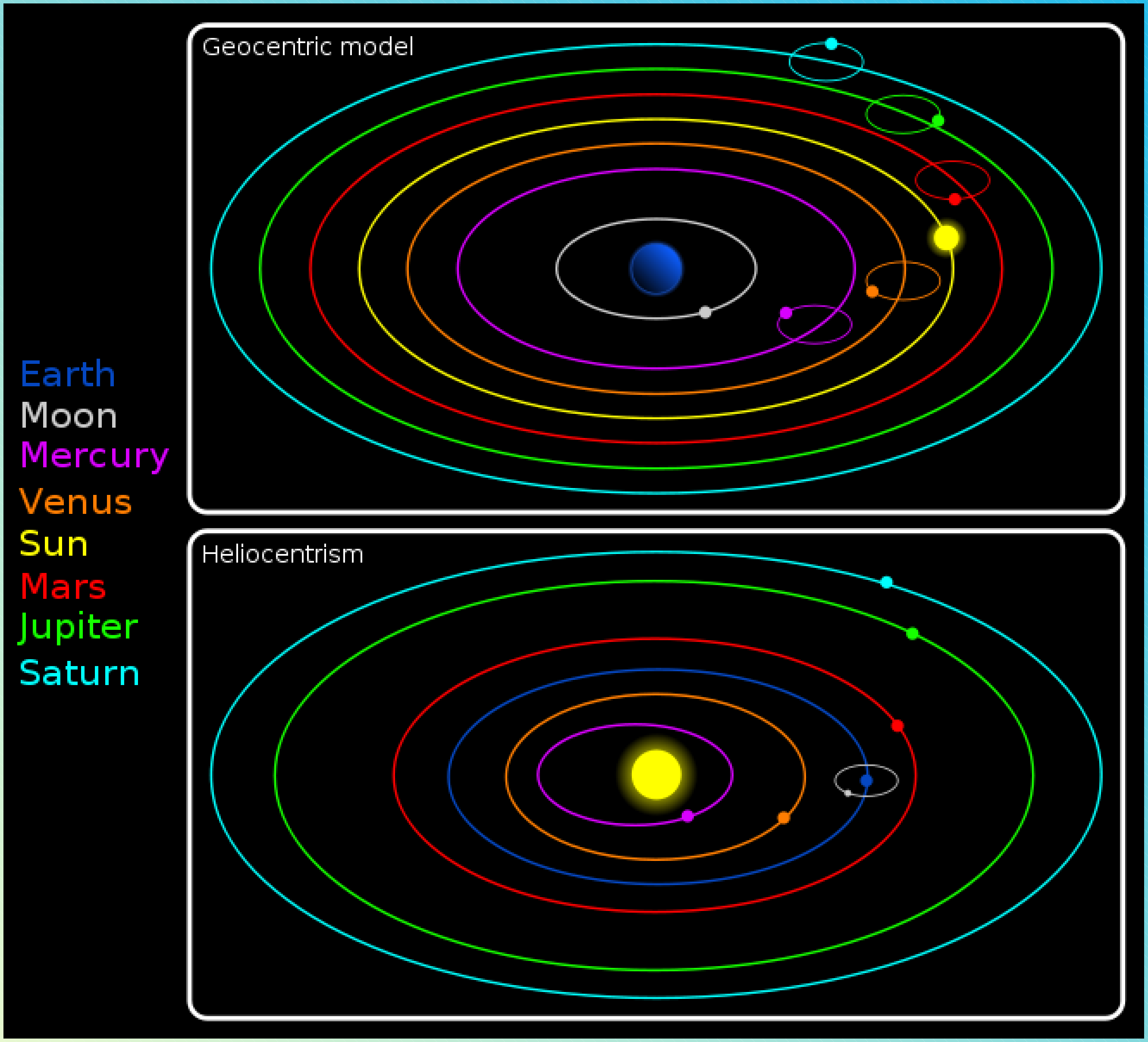
As a result in this scheme

•Baryons tend to follow the hierarchical behaviour of dark matter

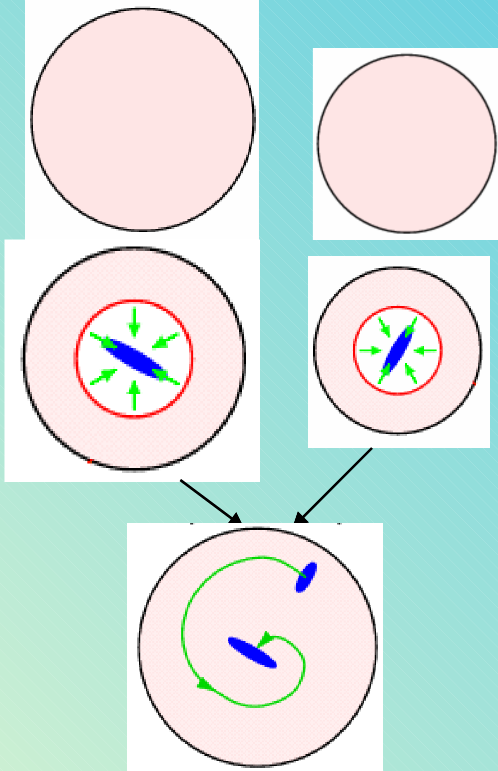
•There is no link between star formation history and morphology

This generates "tension" with several pieces of evidence related to observed downsizing, so far partly cured with a "Ptolemaic approach", i.e. complicating more and more models rather than with a "Copernican revolution". i.e. revisiting the basic assumptions.

Ptolomaeus vs Copernicus



Standard semi analytical models

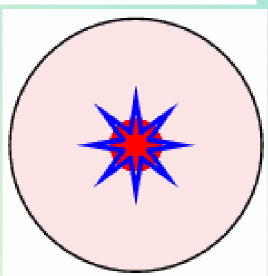


DM Halos form, gas shock heated to virial T»3e6σ_{200} K

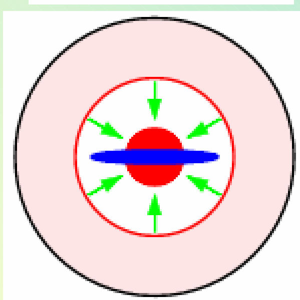
Gas cools and settles into disks with low SF

Halos merge

Galaxy orbit decay by dynamical friction leading to galaxy mergers



If gas rich major merger then starburst and spheroid formation (recently discussed also “dry” mergers)



New disk start to form around spheroid

NB: no build in link between SF history and present day morphology

In practise, a set of differential equations is numerically integrated over time-steps (hence the name SA?) along each DMH merger tree, basically:

$$M_{vir} = M_{DM} + M_{hot} + M_{cold} + M_{stars}$$

$$\dot{M}_{hot} = f_{bar}\,\dot{M}_{vir} - \dot{M}_{cool} + \dot{M}_{reheat}$$

$$\dot{M}_{cold} = \dot{M}_{cool} - \dot{M}_{stars} - \dot{M}_{reheat}$$

where

$\dot{M}_{vir}$ from merger history

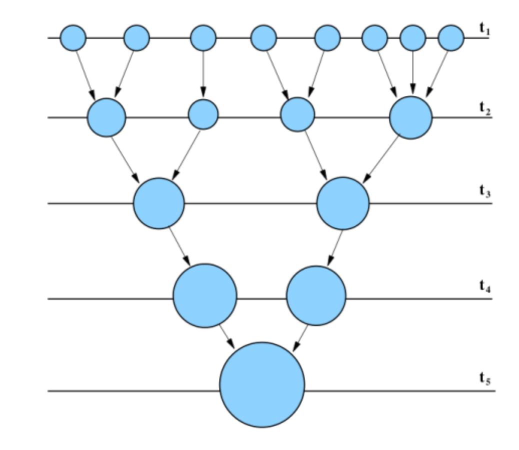
$\dot{M}_{cool}$ from cooling recipe

$\dot{M}_{stars}$ from star formation recipe

$\dot{M}_{reheat}$ from feedback recipe(s)

"Recipe" means simple formula

Halo merger trees: the backbone



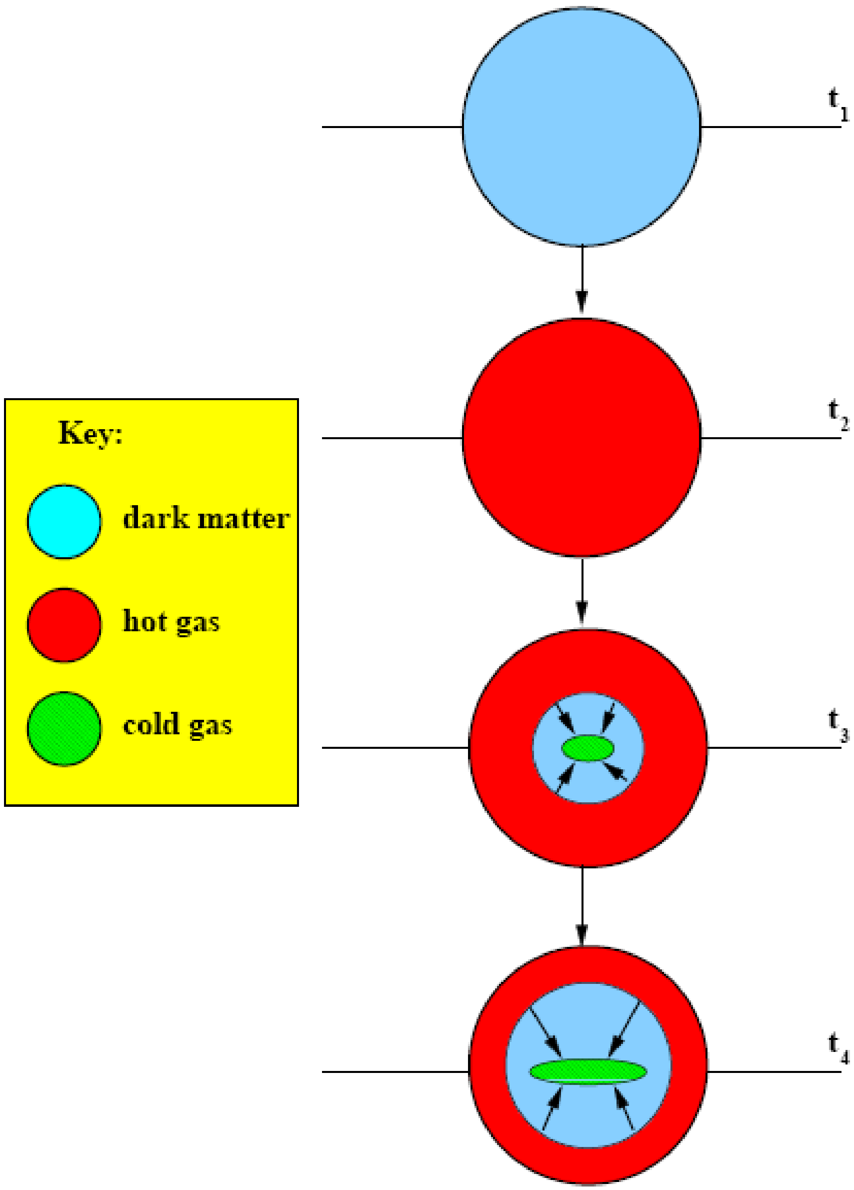
dM_{vir}/dt is obtained computing the hierarchical build-up history of the halo, the halo merger tree. Two alternative approaches to build it:

1. construct halo merger trees using Monte Carlo method based on Press-Schechter approximation and sample halos from PS mass function (e.g. Cole et al 2000). Spherical collapse + assumption of gaussian density field allow an estimate of HMF in reasonable agreement with N-body sim. Faster thus traditionally more employed in SAMs.

2. Extract halo formation histories directly from N-body simulations (aka **post-processing;** e.g. De Lucia et al 2005, Croton et al 2005 using Millenium sim). Non trivial issues are linked to proper halo identification (e.g. sometimes the halo accretion history is not monotonically increasing). Required by studies of environmental dependences.

Halo properties (e.g. distribution of spin, density profiles..) assigned guided by results of N-body simulations, and galaxy formation is followed through each branch of the tree with following prescriptions…

Schematic of gas heating and cooling in standard SAMs



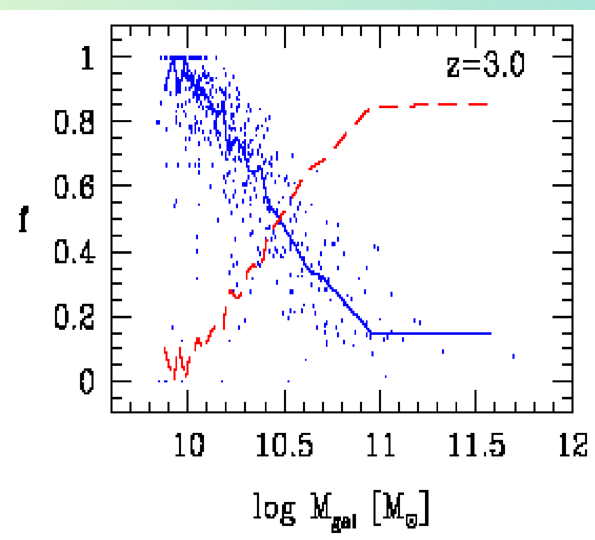
- Baryons fall into the potential well of the dark matter halo.
- gas is assumed heated by shocks to the virial equilibrium temperature
- the inner parts of the hot gas halo cool, forming a rotationally supported disc on the dynamical timescale
- the cooling radius expands and the cold gas disc grows until $r_{cool}=r_{vir}$ or the halo merge with another

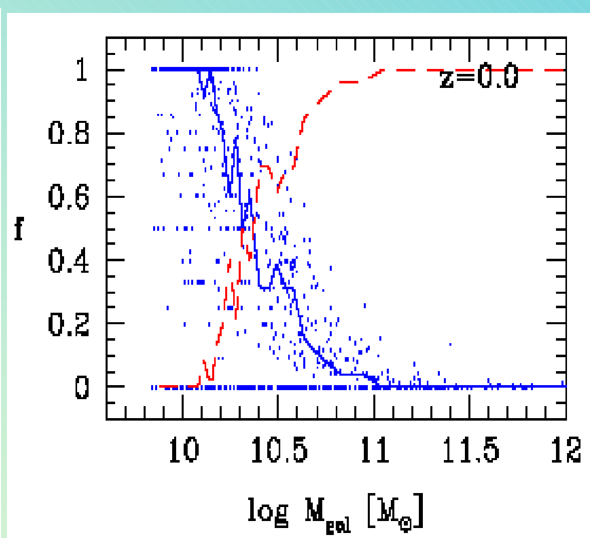
Gas heating

The gas entering the DM halo is assumed to be heated by shocks to the virial temperature of the halo, defined to bring it in virial equilibrium within the halo potential well

$$T_{vir} \quad \frac{1}{2}\frac{\mu m_p}{k}V_{vir}^2 \quad 35 \quad \frac{V_{vir}}{\mathrm{km/s}}^{\;2}$$

This assumption is now questioned by SPH simulations (e.g. Keres et al 2004, 2008, Birnboim & Dekel 2003). A significant fraction of gas accretes to the halo in cold mode. Could be connected to bi-modality in SDSS. Since 2006 SAM begun to explore possible consequences.





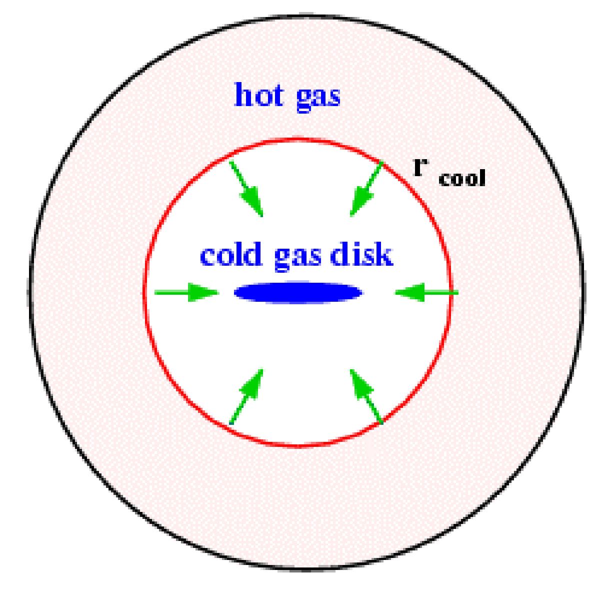
Median hot fraction

Median cold fraction

Gas cooling & disk formation I

After shock heating, the gas starts to radiatively cool, looses P support and assumed to settle on a rotation supported disk.

The hot gas is assumed to have a reasonable radial density profile ρ (r) within the halo, decreasing with r. The cooling rate is higher for denser gas, which therefore cools inside-out:



$$\frac{dM_{cool}}{dt} = 4\pi\, r_{cool}^2\, \rho(r_{cool}) \frac{dr_{cool}}{dt}$$

The cooling radius r_{cool} is given by the condition

$$\tau_{cool}(r_{cool}) = \text{halo age}$$

and the cooling time τ_{cool} is

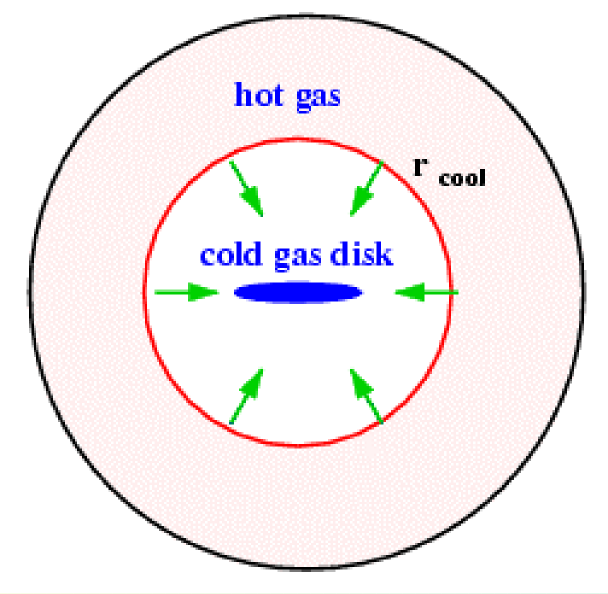
$$\tau_{cool}(r) = \frac{E}{dE/dt} = \frac{3/2\, nkT}{n_e n_i \Lambda(T,Z)}$$

← Cooling function

Gas cooling & disk formation II

Gas without pressure support requires a finite free fall time to reach the center. Thus τ_{cool} is replaced by max(τ_{cool},τ_{ff}) in prev equations.

Cool gas collapses to rotationally supported gas disk, conserving angular momentum (contrary to result of most hydro simulations)

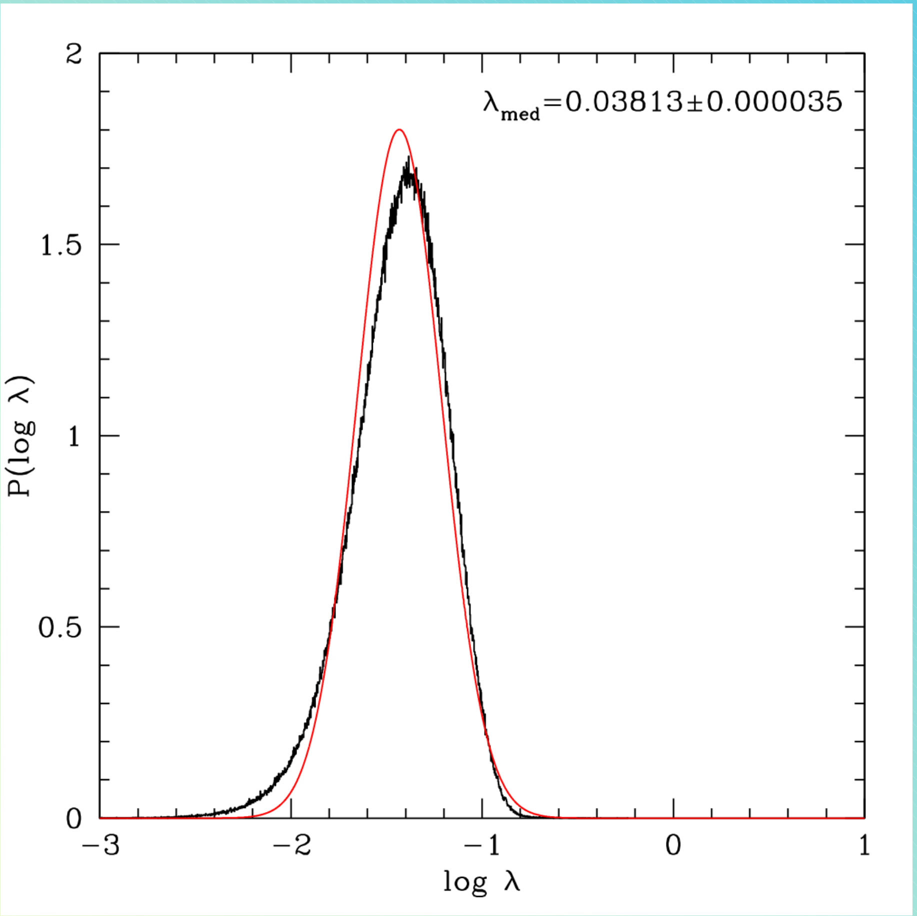


$$\Rightarrow r_{disk} \approx \lambda_H r_{cool}$$

Halo spin parameter Randomly assigned according to lognormal distribution derived from N-body sim

Sizes are important in computing SFR and optical depth, among other things.

Spin parameters of halos in numerical simulations close to lognormal



Bett 2006

Star formation

Fundamental process, poorly understood.

SAMs considers two modes of Star Formation:
quiescent SF from gas disks
bursts quickly consuming all available gas during mergers provided that the mass ratio of the merger is "major", i.e. mass ratio above some threshold (parameter).

Models use prescriptions usually inspired by the observed Schmidt law

SFR for the whole galaxy

$$SFR = \frac{dM_*}{dt} = \frac{M_{cool}(r_{cool})}{\tau_*}$$

SFR timescale τ_* may be a fixed value but usually depends on galaxy and/or burst properties e.g. dynamical time.

Star formation

The dependence introduces free parameters.

A popular choice is $$\tau_* = \varepsilon_*^{-1}\tau_{dyn} = \varepsilon_*^{-1} r_{gal} / V_{gal}$$

Another is $$\tau_* = \varepsilon_*^{-1} \; \frac{V_{vir}}{V_0}^{\alpha_*}$$

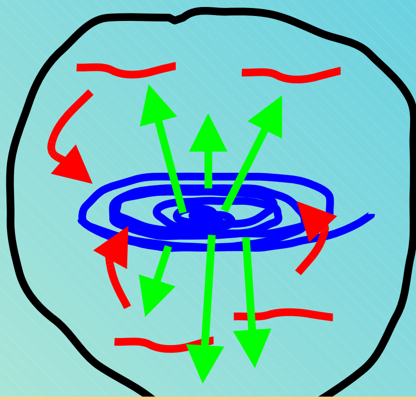
The Durham group GALFORM combines the two

$$\tau_* = \varepsilon_*^{-1}\tau_{dyn} \; \frac{V_{vir}}{V_0}^{\alpha_*} = f\left(V_{vir}, z\right) \text{ since } \tau_{dyn} \quad \frac{V_{vir}}{r_{vir}(z)}$$

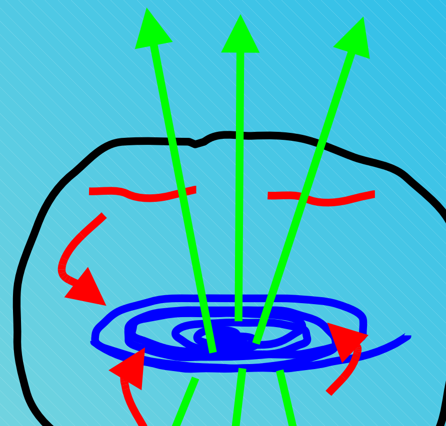
Parameters are calibrated to match the local gas fraction.
(The additional dependence on the velocity is important in GALFORM to reproduce the observed gas fractions in spirals as a function of luminosity.)

But different laws yield very different predictions for gas masses and SFRs at high-z

Supernova feedback



Retention: energy reheats disk material to halo temperature



Ejection: energy expels disk material completely so it is no longer available for cooling.

Energy injection by supernovae & stellar winds, causes cold gas reheated to M_{hot} (retention) and/or ejected from the halo (rejection).

The process is physically extremely complex.

Ejection rate expected to depend on

1. SN rate $\propto$ SFR (for type II SN)
2. Efficiency ε_{SN} = frac of SN energy not lost in radiation
3. Escape velocity (stronger mass loss in smaller galaxies)

Supernova feedback

So a typical recipe is

$$\frac{dM_{reheat}}{dt} = \varepsilon_{SN} E_{SN} \eta_{SN} \left(\frac{V_0}{V_{vir}} \right)^{\alpha} \frac{dM_*}{dt}$$

Where
•E_{SN}» 10^{51} erg is the energy produced by each SN event
•η_{SN}» $10^{-2¥3}$ is the number of SNae produced by each mass solar mass of gas converted to stars, which depends on the IMF.

At least 3 adjustable parameters

Energy balance between wind velocity, if assumed proportional to V_{vir} (thus to escape velocity) and SNae energy would yield α=2.

Galaxy mergers

almost all SAMs assume galaxy mergers have a key role in galaxy formation

•Cause transformations of galaxy morphologies

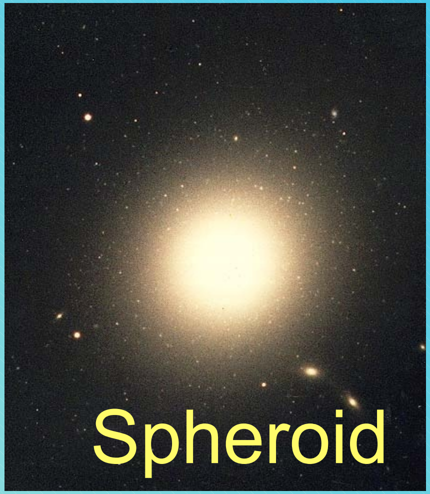
(2 disks ! 1 spheroid). Main channel to form spheroids in most SAM

•Trigger starbursts

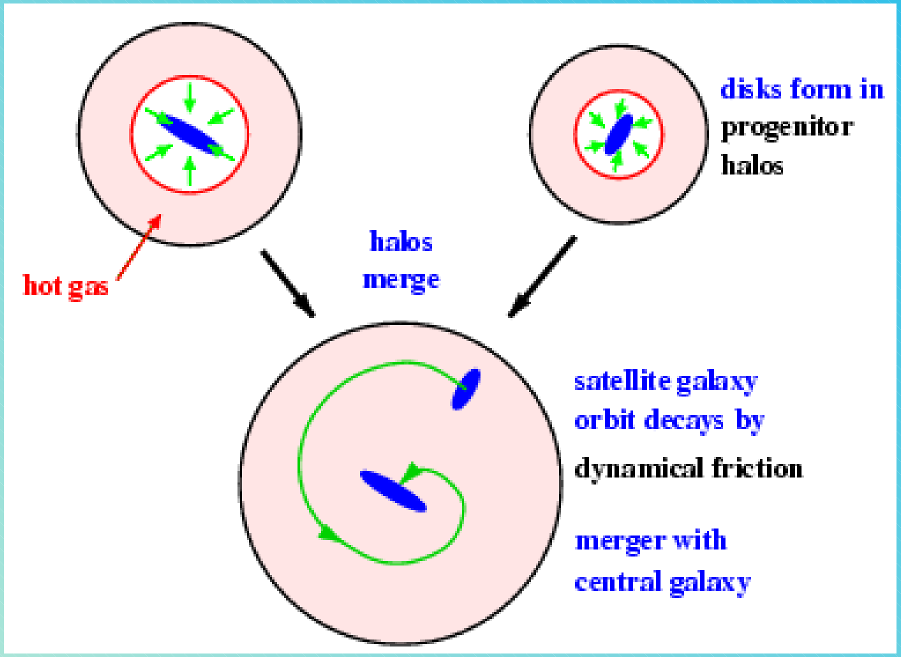
•Modify galaxy mass function

+

=



Galaxy mergers by dynamical friction

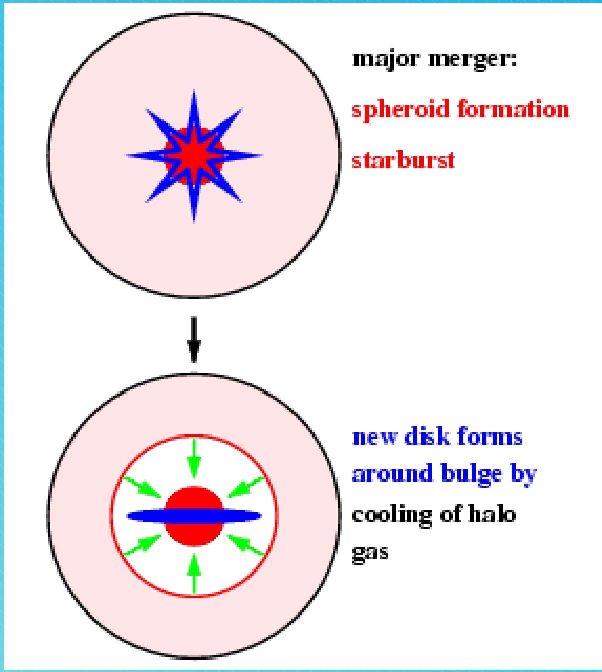


$\tau_{merge}(M_H, M_{sat}, orbit)$ from Chandrasekhar dynamical friction formula (questioned by recent numerical works).

$$\tau_{merge} = \frac{10^{10}}{\ln \Lambda} \quad \frac{r}{60\mathrm{kpc}} {}^{2} \quad \frac{v_c}{220\,\mathrm{km\,s}^{-1}} \quad \frac{2 \quad 10^{10} M_{Sun}}{M}$$

Most models ignore satellite-satellite mergers (encounters too fast)

Morphological transformations in galaxy mergers



Much numerical work on merging of 2 galaxies under different initial condition (gas content, orbital parameter) suggest the a rule of thumb, more or less employed by all SAMs

- Major mergers ($M_1 \sim M_2$):
  - Major mergers of stellar disks completely disrupt disks, producing stellar spheroid
- Minor mergers ($M_2 << M_1$):
  - Small satellite galaxy falling into stellar disk does not disrupt disk, just makes disk thicker
- Dividing line is around $M_2/M_1 \sim 1/3$, treated in SAM as an adjustable parameter
- Size of new spheroid usually \determined by conservation of energy and virial theorem.

Chemical Evolution

A reasonable treatment of the evolution of chemical content of gas in galaxies (and stars formed from that gas) is required to

- Proper estimate of cooling rate, strongly dependent on Z
- Computation of starlight, which is affected by atmospheric abundances
- Comparison with fundamental chemical abundances, which give clues on duration of major star formation phases
- Estimate of dust content and importance of its reprocessing of starlight

Most models uses Instantaneous Recycling Approximation (IRA) which neglects stellar lifetimes: each amount of mass M_* converted to stars immediately returns a fraction $R\, M_*$ to the ISM and produces YM_* of new metals. The remaining $(1-R)M_*$ is assumed to live forever….

The return fraction R ' 0.2¥ 0.5 and the Yield Y ' 0.005 ¥ 0.05 depend on the IMF and are quite uncertain (especially Y).

IRA is un-sufficient for most purposes, however only recently some SAM (e.g. Granato et al 2004, Nagashima 2005, 2008, Arrigoni 2009) began to use full treatments, traditionally employed in so called “monolithic models”.

Population synthesis

If dust reprocessing where negligible (very irrealistic), then the galactic SED at each time would be given by a simple sum over the SED of individual stars alive at that time, which depends on their mass, age and metallicity.

In the simple “monolithic-one-zone” case, the metallicity is an univocal function of galactic age, then

$$F_\lambda(t) = \int_0^t d\tau \int_{m_{\min}}^{m_{\max}} dm\, f_\lambda(m, Z, t-\tau)\, IMF(m)\, SFR(t)$$

In hierarchical models a galaxy is in general the result of merging of sub-units which before merging have had different SF-enrichment histories: stars of a given age have a distribution of metallicities:

$$F_\lambda(t) = \int_0^t d\tau \int_{m_{\min}}^{m_{\max}} dm \int_{Z_{\min}}^{Z_{\max}} dZ\, f_\lambda(m, Z, t-\tau)\, IMF(m)\, \frac{dM_*}{dt\, dZ}(t, Z)$$

Absorption & emission by dust

But the true complexity comes from the fact that it is now clear that dust reprocessing is very important especially at high-z.

Dust modify the SED of galaxies transferring power from the optical-UV, where dust is very effective in absorbing and scattering photons, to the IR, where the absorbed energy is thermally re-emitted

The state of the art in SAMs is to treat the effect with the code GRASIL (Silva et al 1998). The first SAM to do that has been GALFORM by the Durham group.

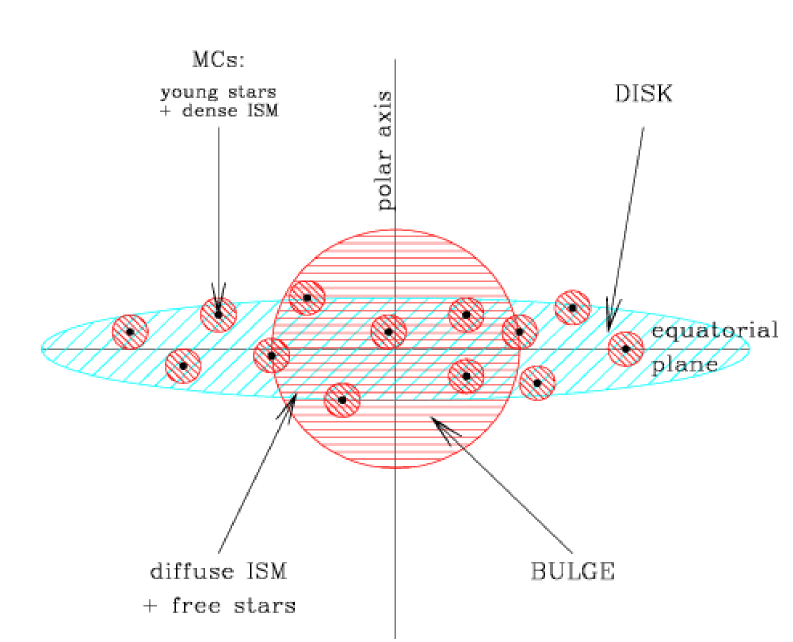
Main GRASIL features

•Radiative transfer when required

•stars in disk + bulge

• dust in molecular clouds + diffuse medium (M_{cl} ~ 10^6 Mo, r_{cl} ~ 20 pc)

• stars form in clouds & leak out (t_{esc}~few Myr)



GRASIL geometry

Less common/classical ingredients

Satellite-Satellite Mergers Satellite galaxies can also merge among themselves, but less effectively. (e.g. Santa Cruz & Galics)

Spheroid Formation from Disk Instabilities: Disks that are too compact are unstable. They form a bar which may later 'dissolve' into a bulge. (e.g. van den Bosch 1998, Cole et al 2000, Galics)

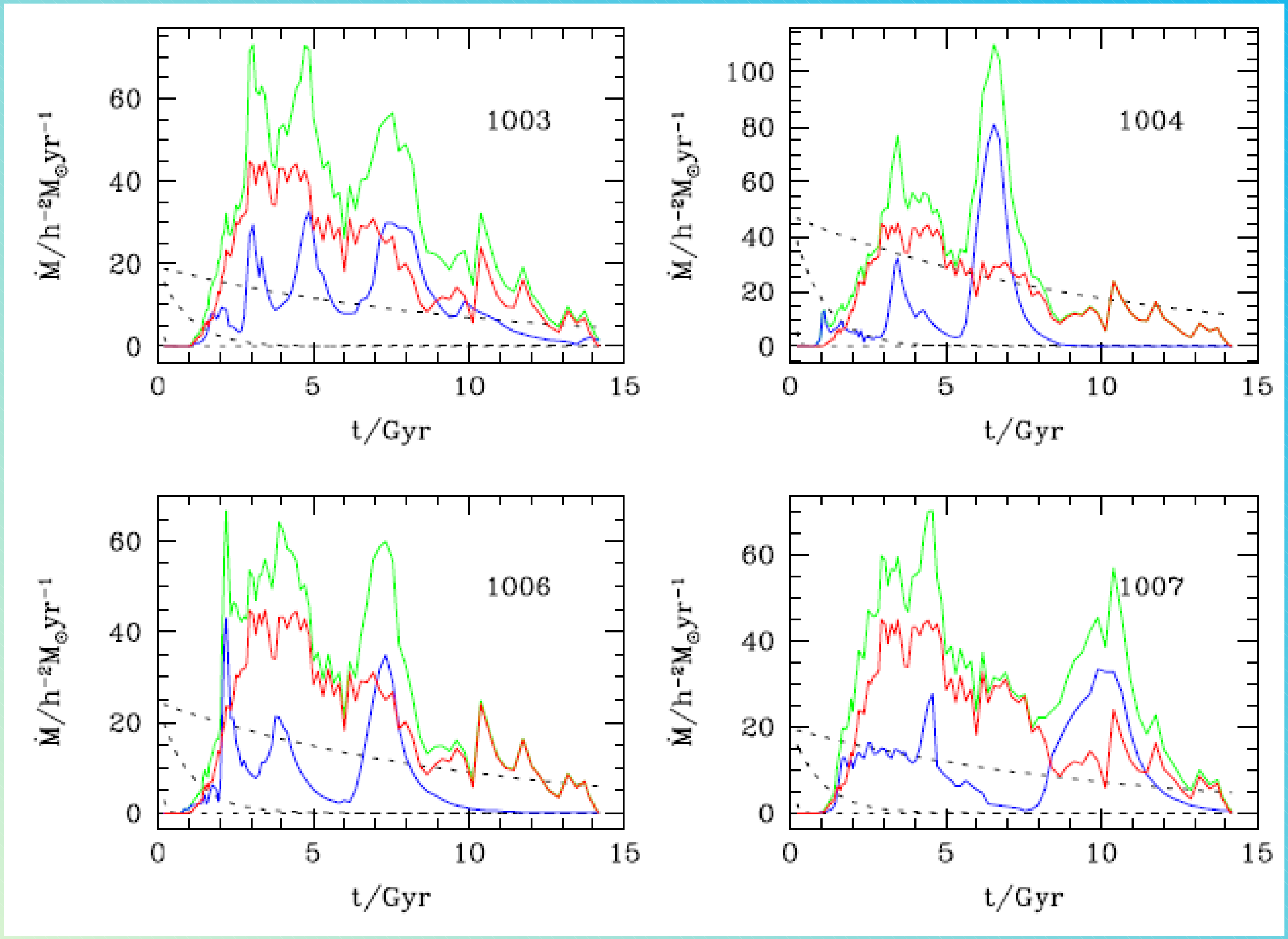
SMBH grow and AGN feedback: the growth of SMBHs is treated considering merging of SMBH and accretion. Observationally, the latter process should dominate. The ensuing feedback on ISM has been recognized only recently as a key mechanism, now all SAMs incorporate it (Granato et al 2004, Monaco & Fontanot 2005, Bower et al 2006, Croton et al 2006, Menci et al 2006, Somerville et al 2008)

Most popular SAMs and physical processes included

(from Benson 2010)

Feature	Model: DURHAM[1]	MUNICH[2]	SANTA-CRUZ[3]	MORGANA[4]	GALICS[5]
Merger Trees					
→ Analytic	Modified ePS[6]	ePS[7]	ePS	PINOCCHIO[8]	×
→ N-body	✓[9]	✓	✓	×	✓
Halo Profiles	Einasto[10]	Isothermal	NFW	NFW	Empirical[11]
Cooling Model					
→ Metal-dependent	✓	✓	✓	✓	✓
Star Formation	✓	✓	✓	✓	✓
Feedbacks					
→ SNe	✓	✓	✓	✓	✓
→ AGN	✓[12]	✓	✓	✓	✓[13]
→ Reionization	✓[10]	×	✓	✓[14]	✓[15]
Merging					
→ Substructure[16]	N-body[17]	N-body[17]	DF[18]	DF[18]	N-body[17]
→ Substructure–Substructure[19]	✓[20,10]	×	✓[21,22]	×	✓[21]
Environments					
→ Ram Pressure Stripping	✓[23]	✓[24]	×	×	✓[25]
→ Tidal Stripping	✓[10]	×	✓	✓	✓
→ Harassment	×	×	×	×	×
Disks					
→ Disk Stability	✓	✓	✓[26]	✓	✓
→ Dynamical Friction[27]	✓[28]	×	×	×	×
→ Thickness	✓[28]	×	×	×	×
Sizes					
→ Adiabatic contraction	✓	×	✓	✓	×
Chemical Enrichment	✓ [delayed[10]]	✓ [instant[29]]	✓ [delayed[30]]	✓ [instant]	✓ [delayed[31]]
Dust	GRASIL[32]	Screen[33]	Slab[34]	GRASIL[32,35]	Slab[34]

A small selection of results from the big industry of SAMs



Examples of predicted SFR in model galaxies in a standard SAM (Durham): red quiescent in disk, blue burst in mergers, green total

Gas fraction in L* spirals

Gas fraction dependence on L

Faint end of optical LF

Solar neighbourhood

Morphological mix at L*

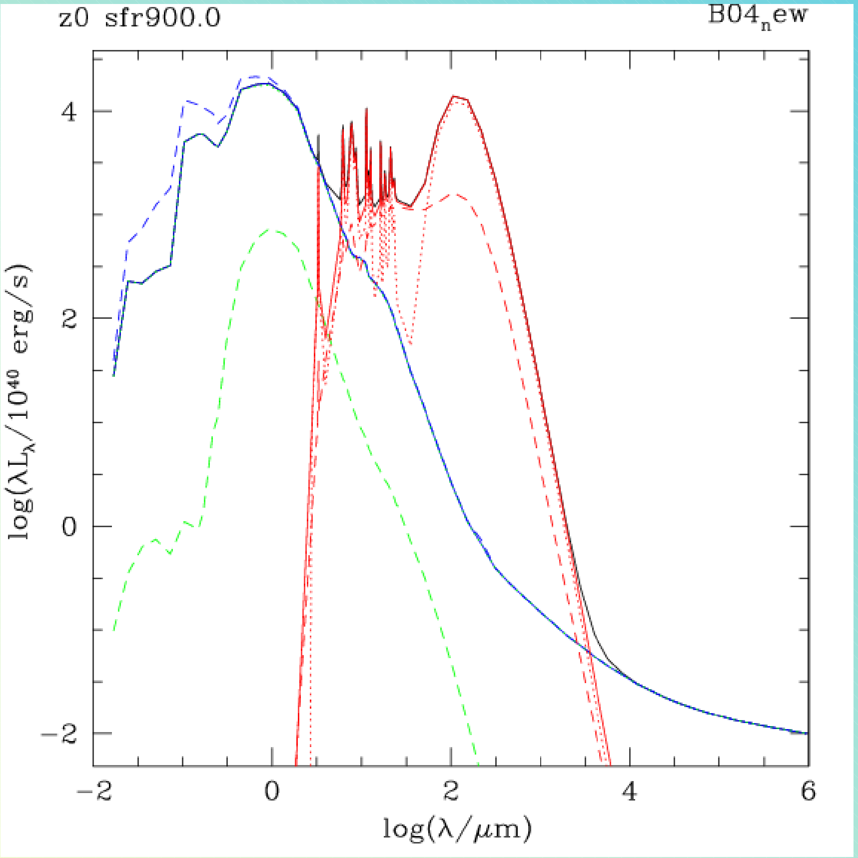
bright end of FIR LF

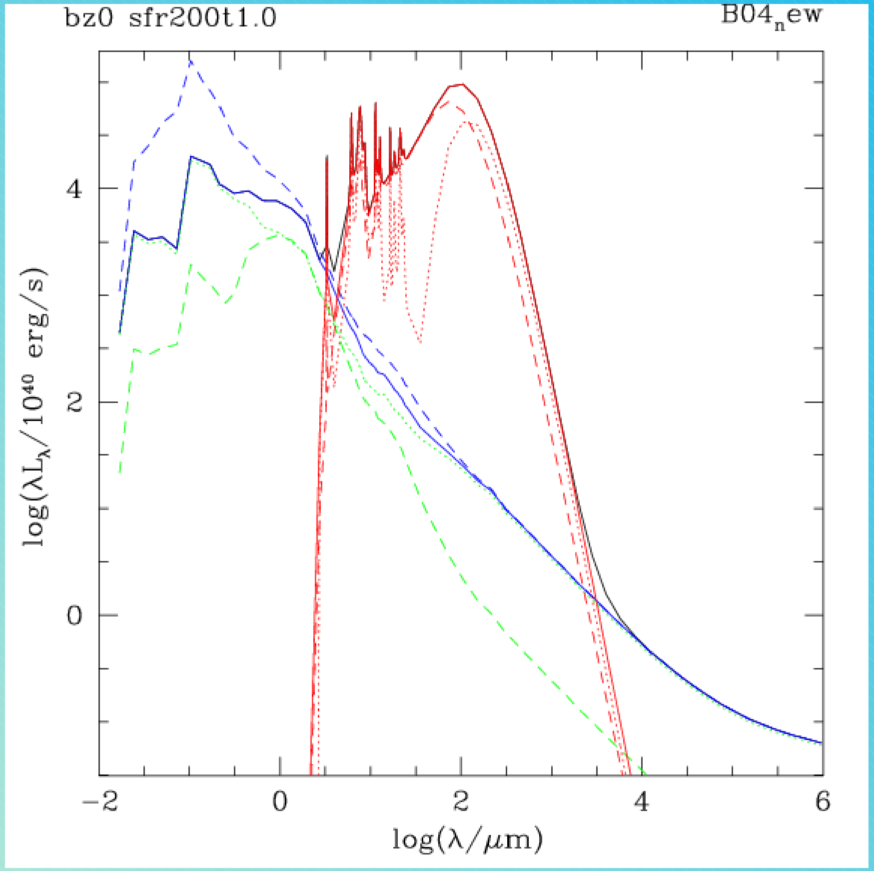
Star Formation and Feedback	
$\epsilon_{*\mathrm{disk}}$	6.7×10^{-3}
α_{*}	-1.5
α_{hot}	2.0
V_{hot} (km s^{-1})......	150.0
Stellar Populations	
IMF	Kennicutt (1983)
Υ..................	1.4
p	0.02
R..................	0.29
Mergers and Bursts	
f_{ellip}	0.3
$\epsilon_{*\mathrm{burst}}$	0.5
η	0.1

Example SEDs of model galaxies

Quiescent spiral

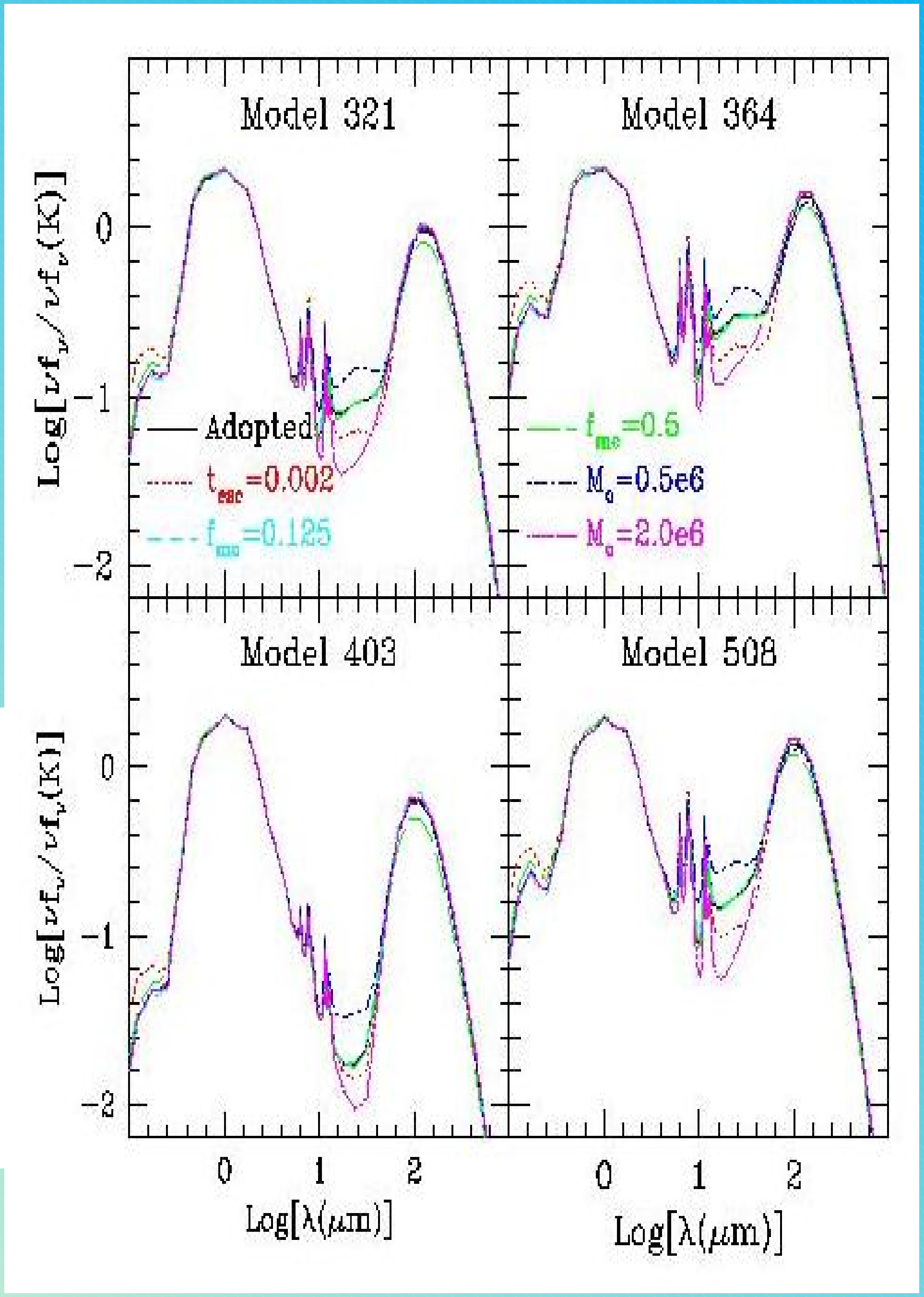
Ongoing burst



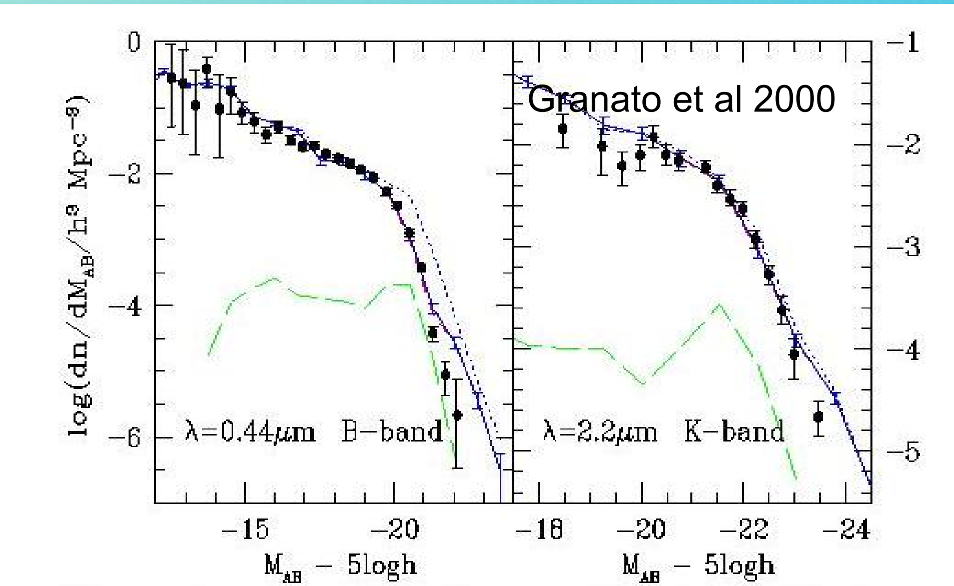


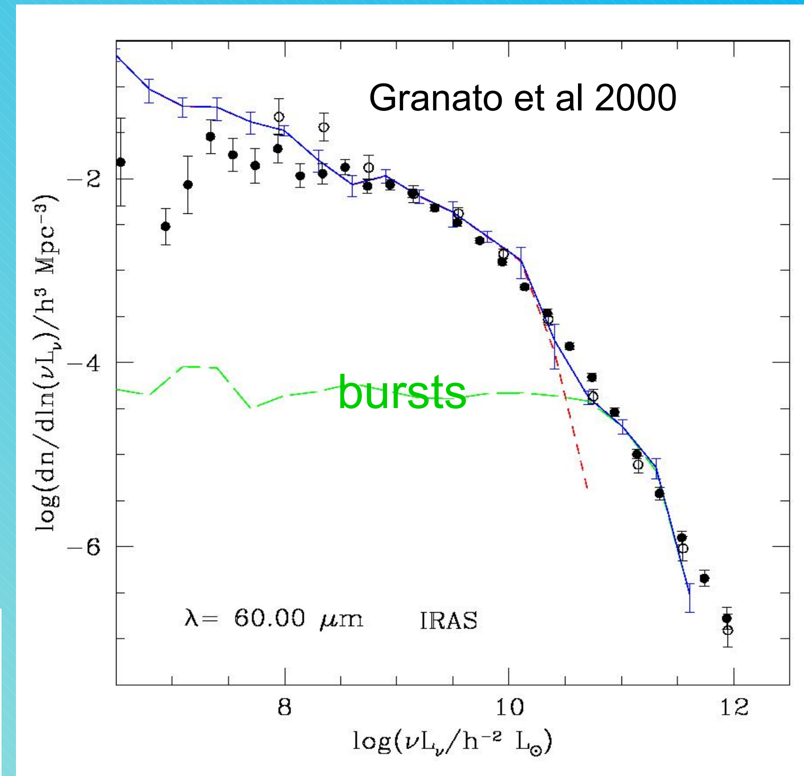
Given a GALFORM model, “only a few” GRASIL parameters remains to be set, affecting mostly MIR and FUV

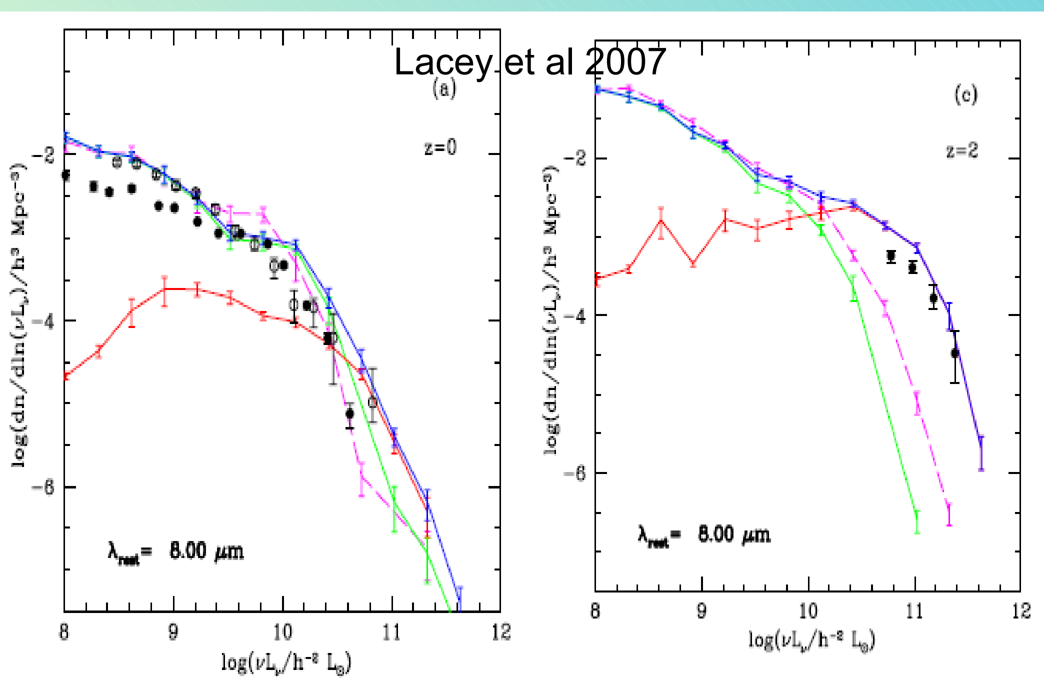
$h_z/h_R(disk)$	0.1
$h_z/h_R(burst)$	0.5
$h_z(dust)/h_z(stars)$	1
f_{mc}	0.25
M_c/r_c^2	$10^6 M_\odot/(16\text{pc})^2$
tesc(*disk*)	2Myr
tesc(*burst*)	10Myr

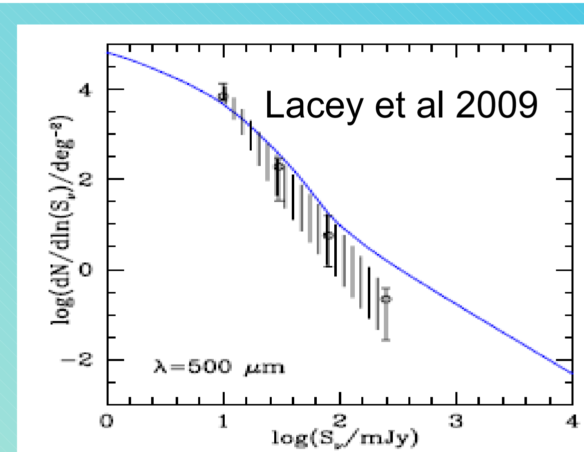


Standard test for SAM: LFs at various z and number counts



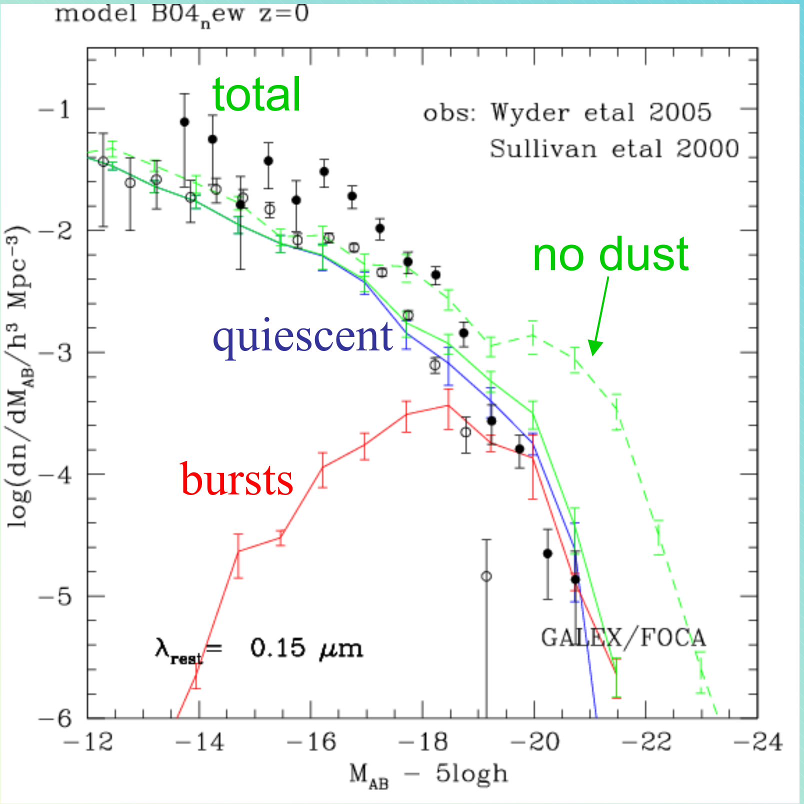




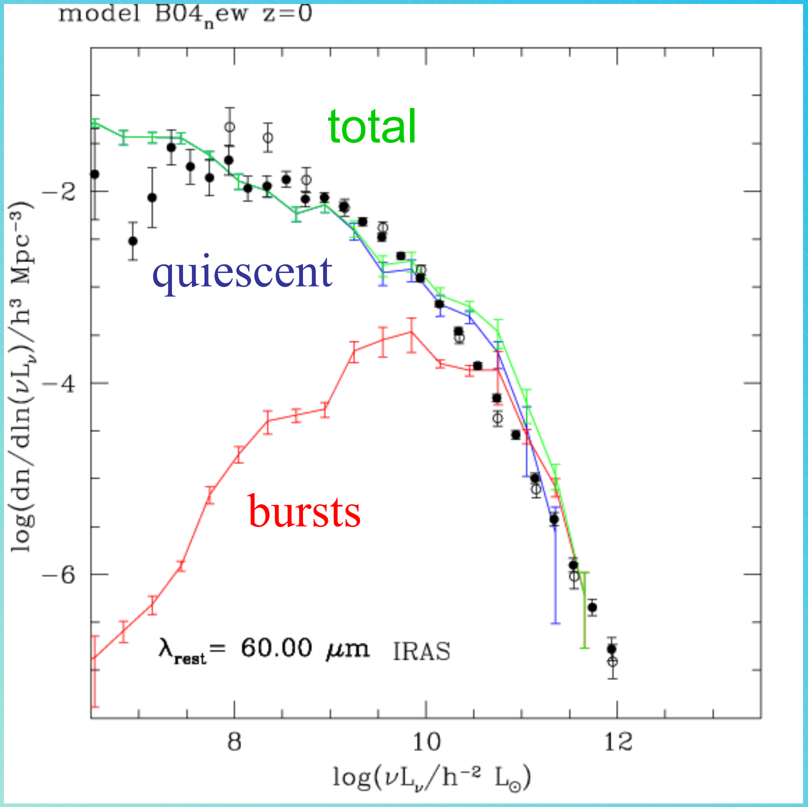


Present day luminosity functions in far-UV & far-IR (z=0)
importance of computing dust reprocessing

Far-UV (1500 A)

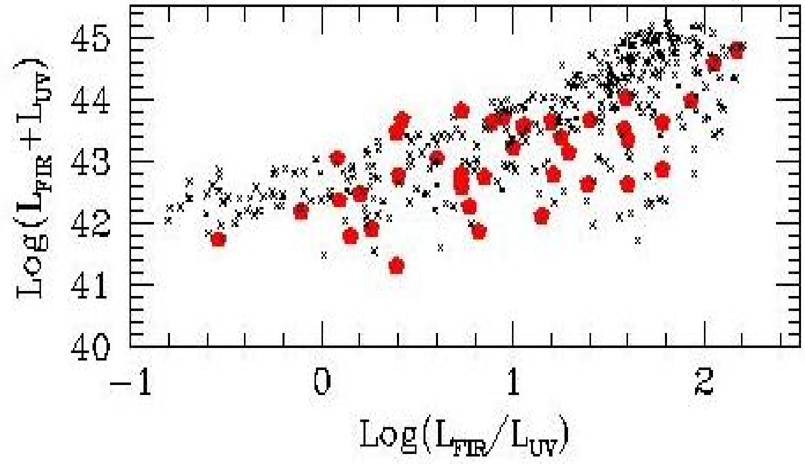


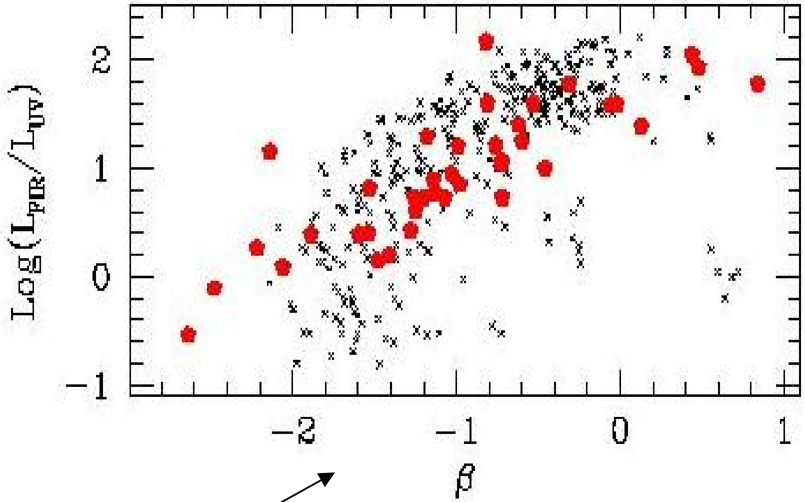
Far-IR (60 μ m)

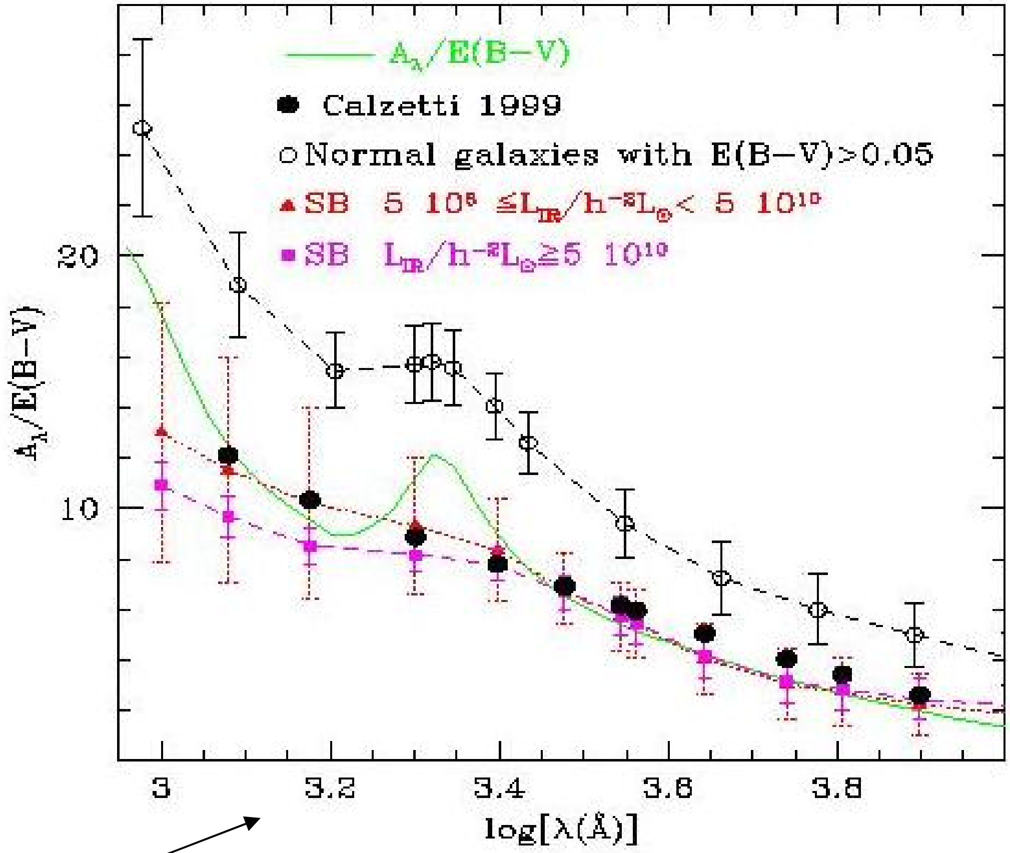


Granato, Lacey, Silva et al 2000, ApJ

UV bright starburst properties



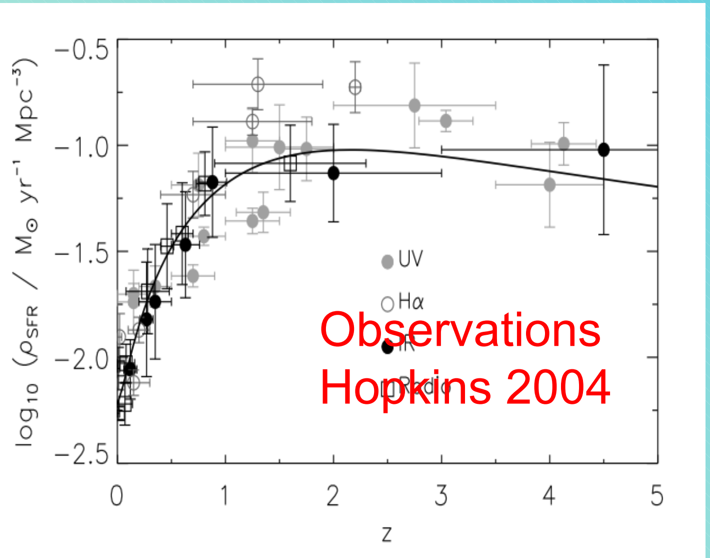


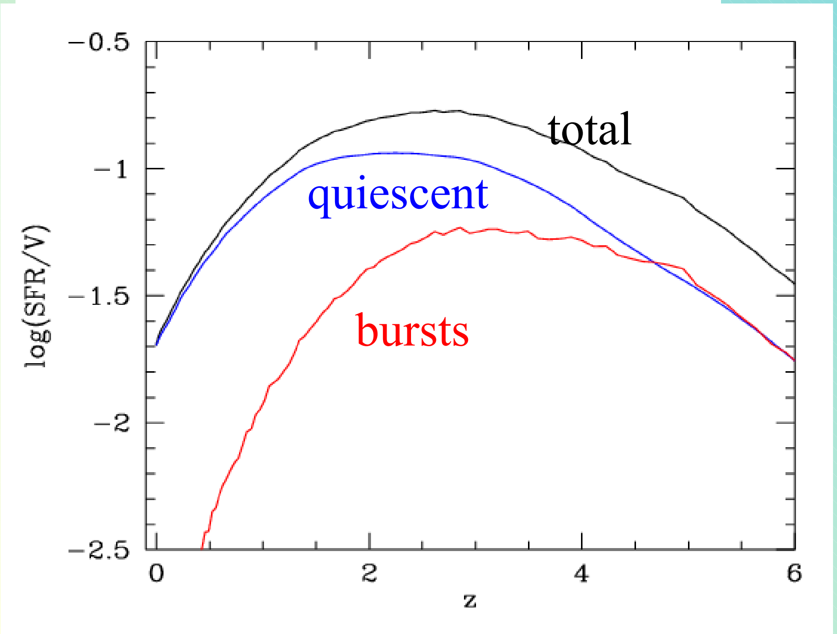


correlations
Red: data
black: models

The model reproduces the dust attenuation law of SB (Calzetti), as a consequence of age dependent geometry (differential extinction). Peculiar dust properties not required nor excluded.

Interpretation of the cosmic star formation history with a standard SAM (GALFORM – Durham)





- rise in SFR/V at early times due to build-up of halo mass, allowing radiative cooling in halos with $T_{vir} > 10^4$ K
- further halo build-up allows more efficient star formation, due to weaker SN feedback as V_c increases
- decline at late times due to slower radiative cooling in very massive halos

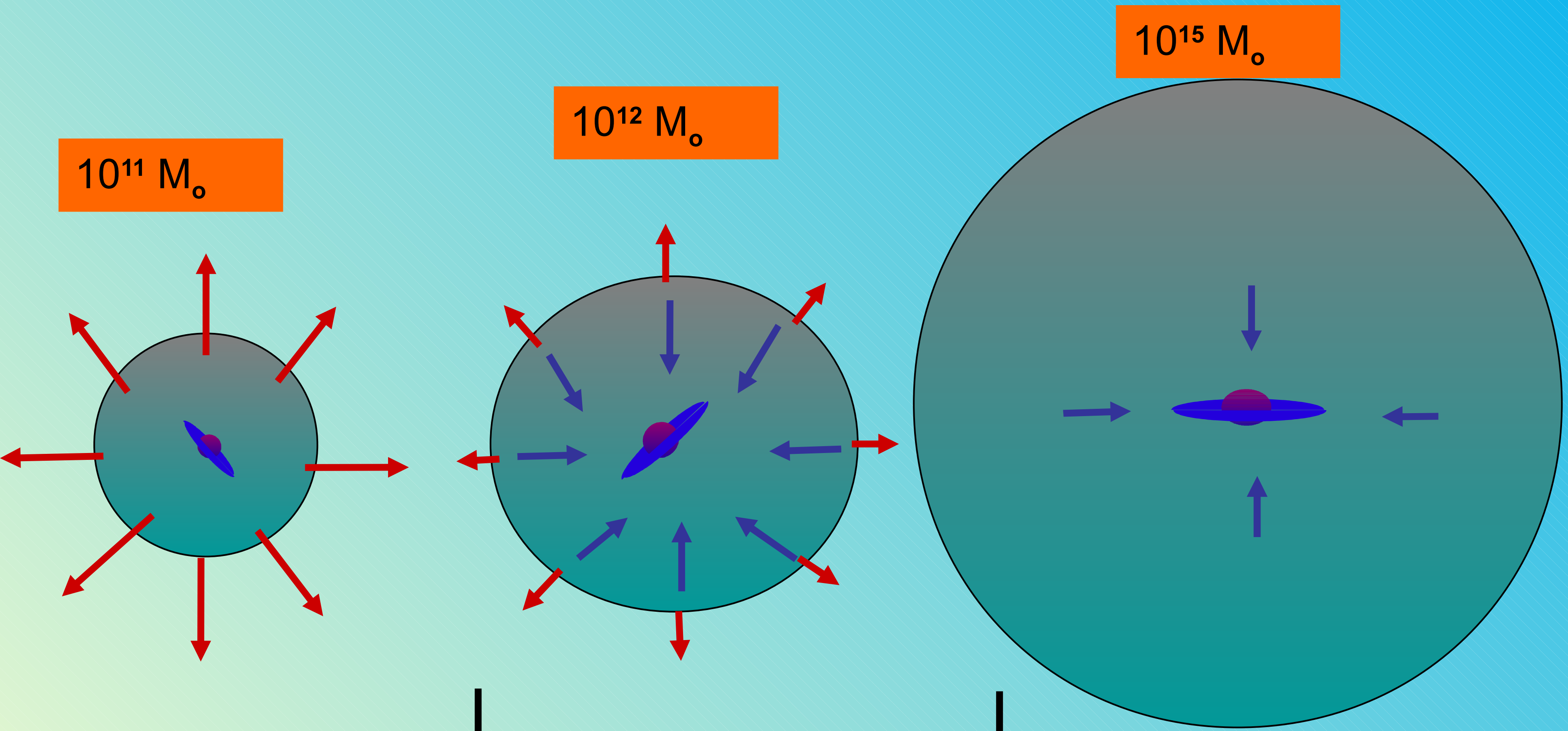
•However the agreement is not fantastic, in particular the decline going to high z in model is too fast. This despite years of struggle…

ON THE POWER OF SEMI ANALYTIC TECHNIQUE

The modularity (and moderate computer requirements) of SAMs allow studies the effects of individual physical mechanisms

As an interesting example let's see "deconstruction" of galaxy luminosity function made by Benson et al. 2003.

Efficiency of galaxy formation



Cooling: +++

Feedback: - - - -

Gal. Formation:

Inefficient

Cooling: ++

Feedback: - -

Gal. Formation:

Efficient

Cooling: +

Feedback: -

Gal. Formation:

Inefficient

Problems of standard SAM and simulations I

Besides some nontrivial successes, most calculations shows severe mismatches with several observations

	Name	reference model	Comment	Constraint	Equation/ Section
			mergers		
	$f_{\rm hmm}$	0.2	major merger condition for DM halos	N-body	eq. 10
	$f_{\rm gmm}$	0.3	major merger condition for galaxies	N-body	eq. 11
	f/c	2.0/4.0	bulge formation in mergers/disc instabilities	N-body	eq. 49
!	$f_{\rm scatter}$	0.1	fraction of stars scattered at a galaxy major merger	N-body	sect.4.3
			halo component		
!	γ_p	1.2	polytropic index of the hot gas	observ.	eq. 13
	$f_{\rm shock}$	1.2	shock heating factor	N-body	eqs. 20, 21
!	heat cold gas	YES	switch for heating cold halo gas at major mergers	N-body	sect. 5.2
!	$n_{\rm quench}$	1	no. of crossing times for quenching cooling	free	sect. 5.2
!	close hole	NO	switch for closing the cooling hole	free	eq. 29
!	infall on bulge	YES	switch for allowing infall on the bulge	free	eqs. 33, 35
!	$n_{\rm dyn}$	1.5	no. of dynamical times for infall	free	eqs. 31, 32
	$f_{\rm wind}$	2	energy factor to trigger a super-wind	free	eqs. 37, 41
	$f_{\rm back}$	0.5	fraction of super-wind mass that falls back	free	sect. 5.5
			disc structure		
	$\epsilon_{\rm limit}$	0.9	limit for bar instability	N-body	eq. 48
	$f_{\rm bar}$	0.5	fraction of disc that goes to bulge	free	sect. 6.2
	adiabatic contr.	NO	switch for adiabatic contraction	free	eq. 47
			stars and metals		
	$M_{\star,\rm SN}$	120 $M_\odot$	stellar mass per SN	IMF	eqs. 60, 67
	$f_{\rm rest}$	0.4	fraction of restored mass	IMF	eqs. 59, 63
!	Y	0.03	metal yield per generation	observ.	eqs. 77, 78
	$Z_{\rm pre}$	10^{-6}	metallicity due to pre-enrichment	free	sect. 8
!	$f_{Z\rm ej}$	0.5	fraction of metals ejected to halo	free	eqs. 77, 78
			star formation and feedback		
!	$f_{\rm th,D}$	0.7	thermal efficiency of feedback in thin systems	free	eq. 60
!	$f_{\rm th,B}$	0.5	thermal efficiency of feedback in thick systems	free	eq. 67
!	$f_{\rm kin}$	0	kinetic energy from hot winds	free	eqs. 61, 67
!	σ_0	60 km s^{-1}	turbulent velocity of clouds	free	eqs. 55, 56
!	$\Sigma_{\rm thr}$	0 $M_\odot$ pc^{-2}	gas surface density threshold for star formation	observ.	section 7.2
!	$\Sigma_{\rm limit}$	∞ $M_\odot$ pc^{-2}	critical gas surface density for discs	free	eq. 50
!	hot kin. fb	YES	switch for heating cold gas by kinetic feedback	free	eqs. 71-75
			AGN		
	$M_{\rm seed}$	1000 $M_\odot$	seed black hole mass	theory	sect. 9.1
!	$f_{\rm low J}$	0.003	rate of loss of angular momentum	free	eq. 79
!	$f_{\rm jet,0}$	1	efficiency of jet feedback for a 1000 km s^{-1} halo	N-body	eq. 82
			numerical parameters		
!	$M_{\rm part}$	$10^9 M_\odot$	particle mass	—	—
	Δ_t	0.1 Gyr	numerical interval for the integration	free	sect. 2.2

Table 4. Model parameters, with their value adopted in the reference model, brief description, available constraints (independently of the model) and reference in the text. Parameters highlighted by a mark are of primary importance. Cosmological parameters are not included.

This is a quite surprising given the large (if not ridicule) number of free parameters involved.

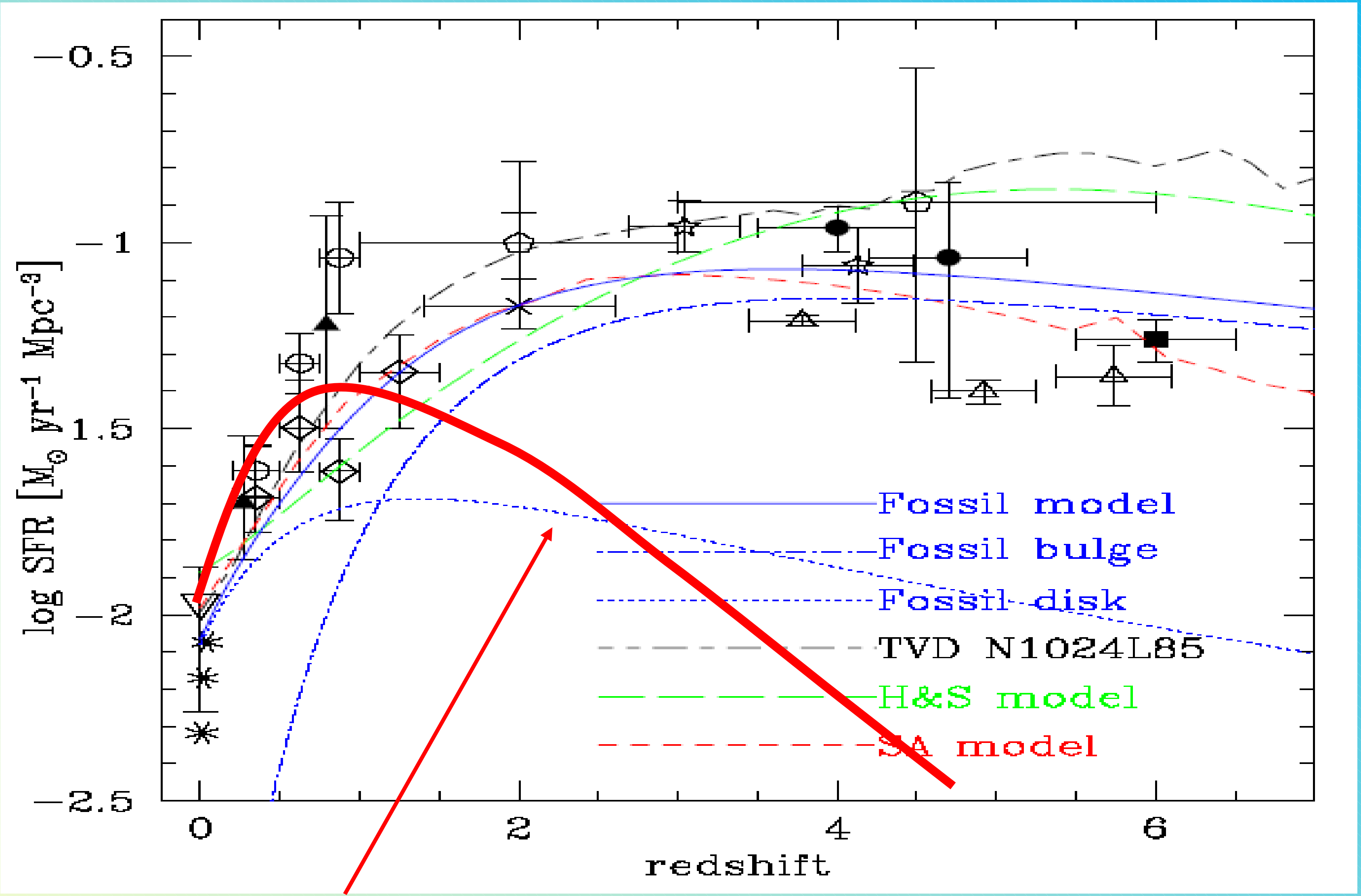
Are there fundamental flaws in the picture?

Parameters in the Monaco et al 2006 model

On predictive power of models: Standard SAMs

- In retrospect, models based on this general scheme performed quite poorly in anticipating observational breakthroughs in the last 10-15 years
- Let’s see the main examples..

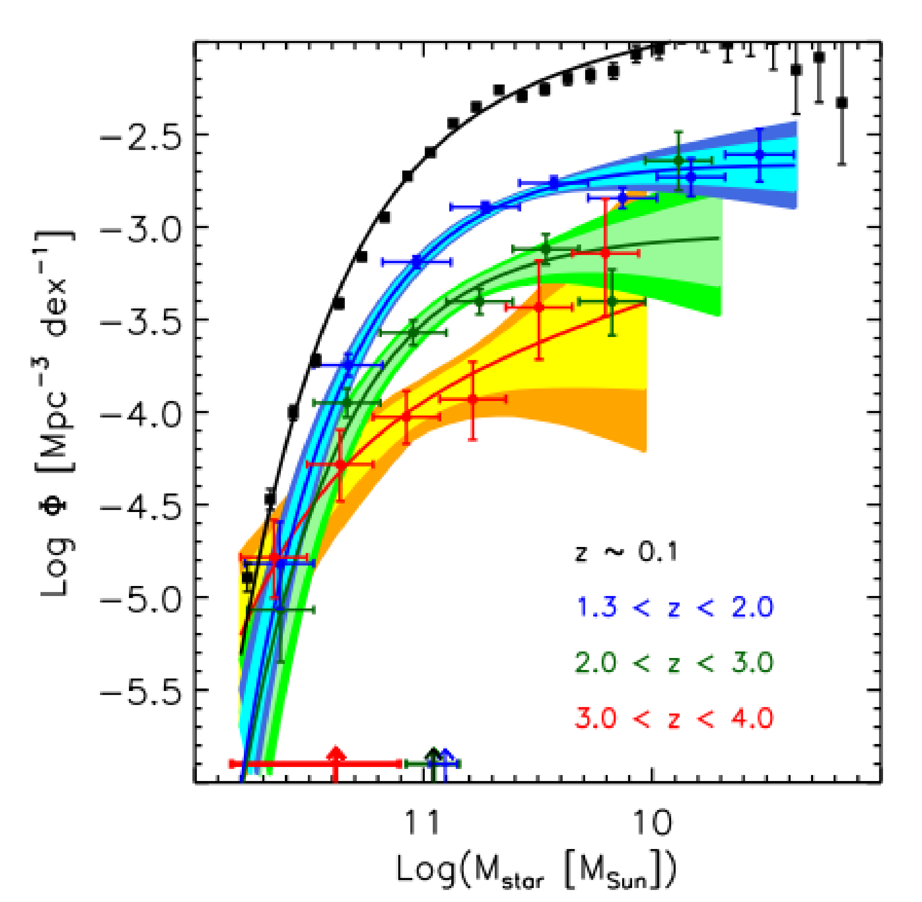
On predictive power of standard SAM 1: SFR(z)

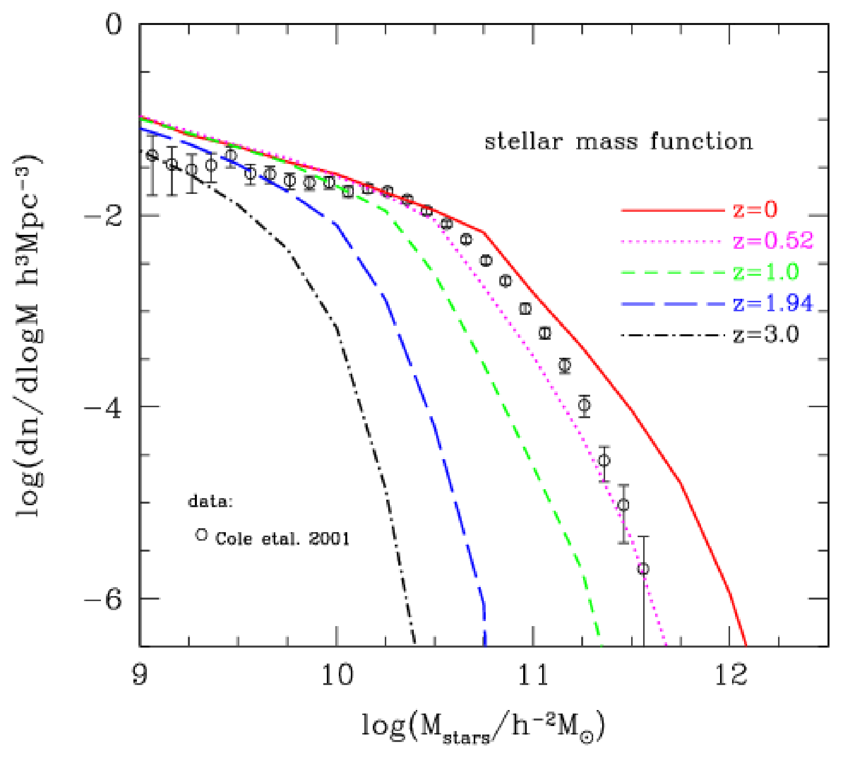


Forecast by Cole et al 1994 (before Madau at al 1996)

On predictive power of standard SAM 2: downsizing

Observations

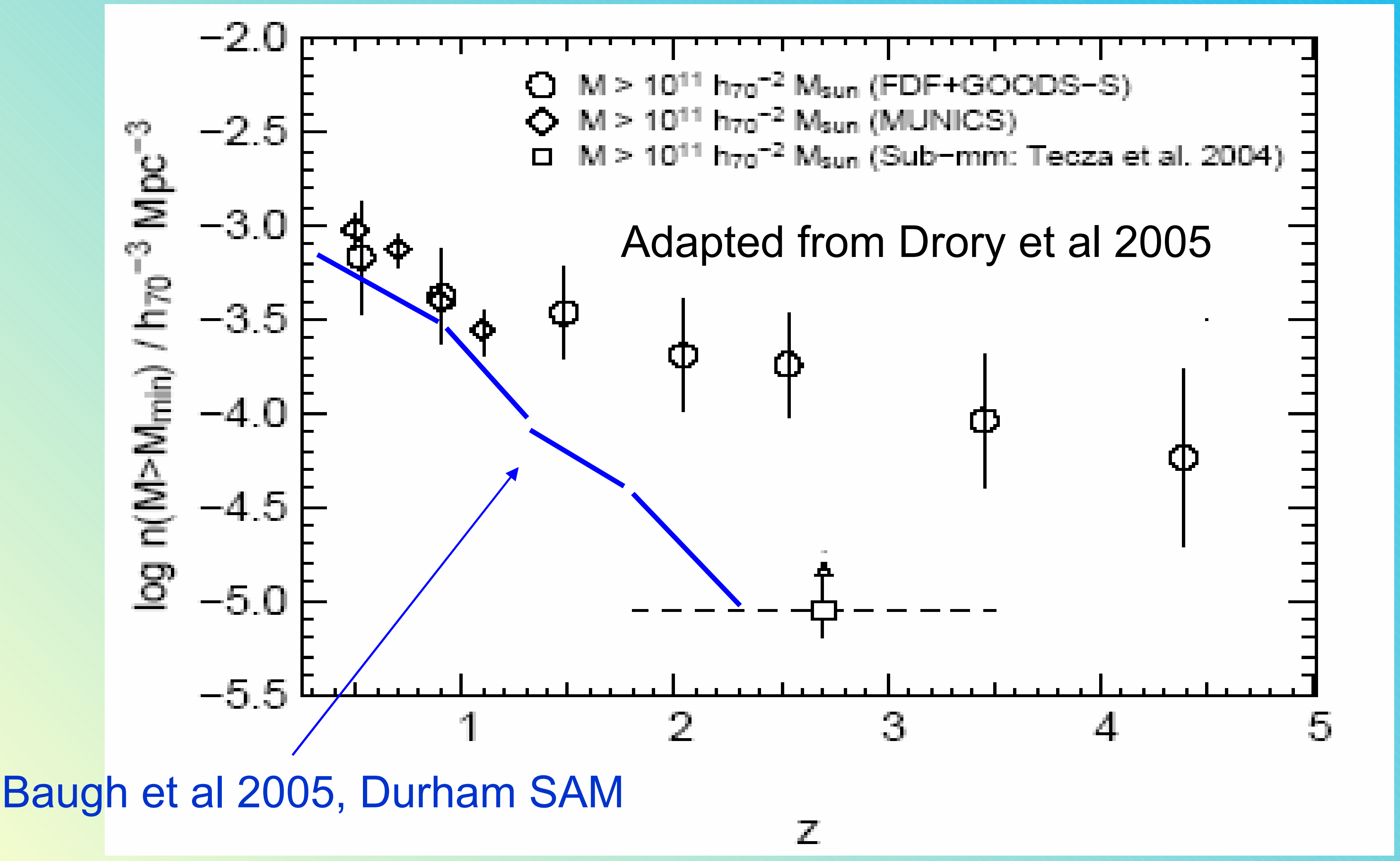




MODEL Cole et al. 2000

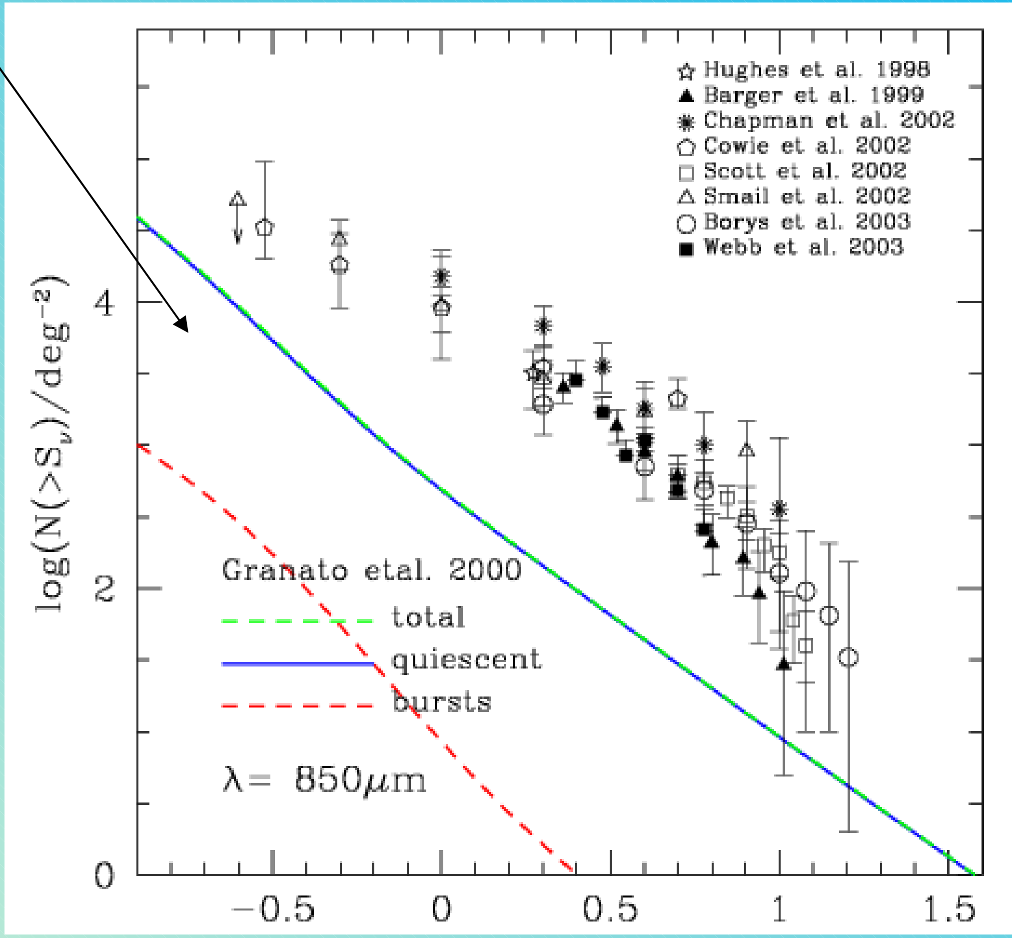
Mass increases (funny plot)

On predictive power of models
3: massive galaxies at high-z



On predictive power of standard SAM
4: sub-mm galaxies at high-z should not exist!

GALFORM (Durham) reference model 2000 +GRASIL



No SAM anticipated the existence of high-z sub-mm population

Even now most models are heavily challenged by the high number counts and (even more) by z-distribution.

Typically model sources at z<0.1, and far too few and/or too faint!!

CONCLUSION

(why to spend time with models?)

- Many of the observational breakthroughs of the last 10-15 years were surprises for existing models;
- models at present are not first principles but tools to see if general physical ideas may explain what is largely already seen; for instance to assess importance of various processes.

- Galaxy formation theory is led by observations
- Corollary: keep your mind open to alternatives, and promptly use data to refine or rethink ingredients;

Properties of spheroids

Standard semi-analytic models have troubles with a set of observations broadly speaking related to properties of (large) E galaxies:

Large E pop are observed already in place (and old) at z>1-1.5 (eg. Im et al 2001, Cohen 2001, Van Dokkum & Ellis 2003)

The color-magnitude and/or the $[\alpha/Fe]$-Mass relations (Cole et al 2002, Thomas et al 2002, Nagashima et al 2005)

Sub-mm high z population is under-produced (SMG)

Deep NIR-band selected samples are under-produced

Cole et al. 2000

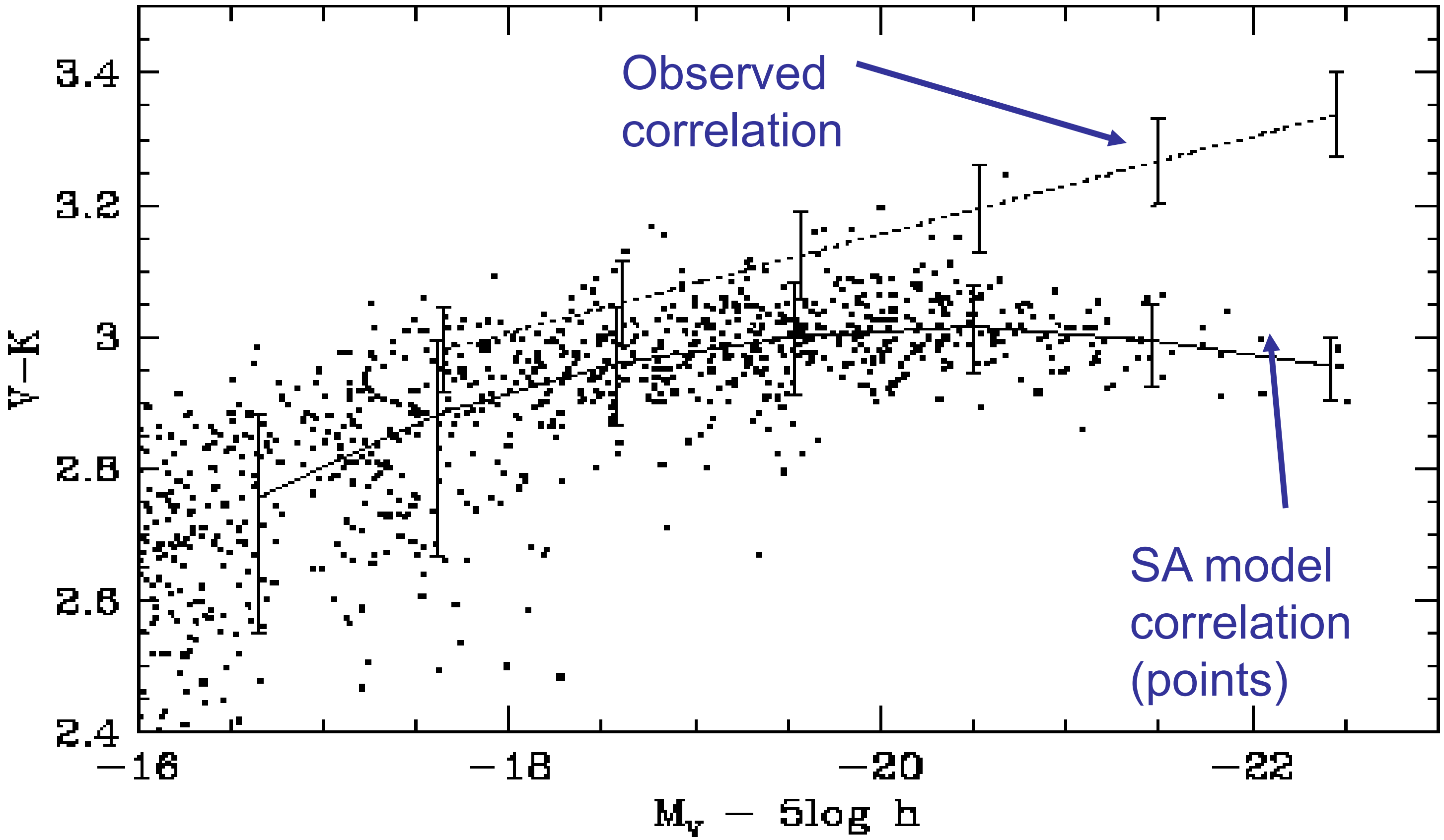
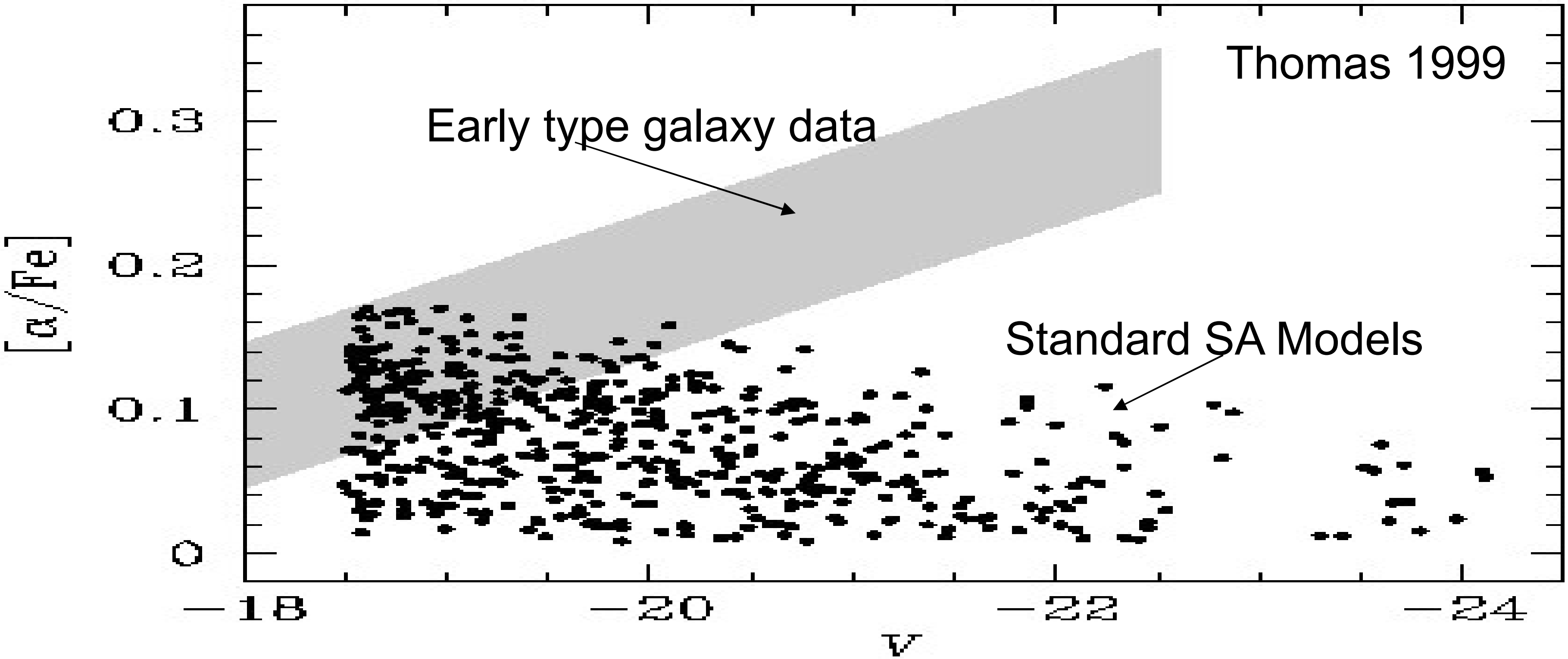
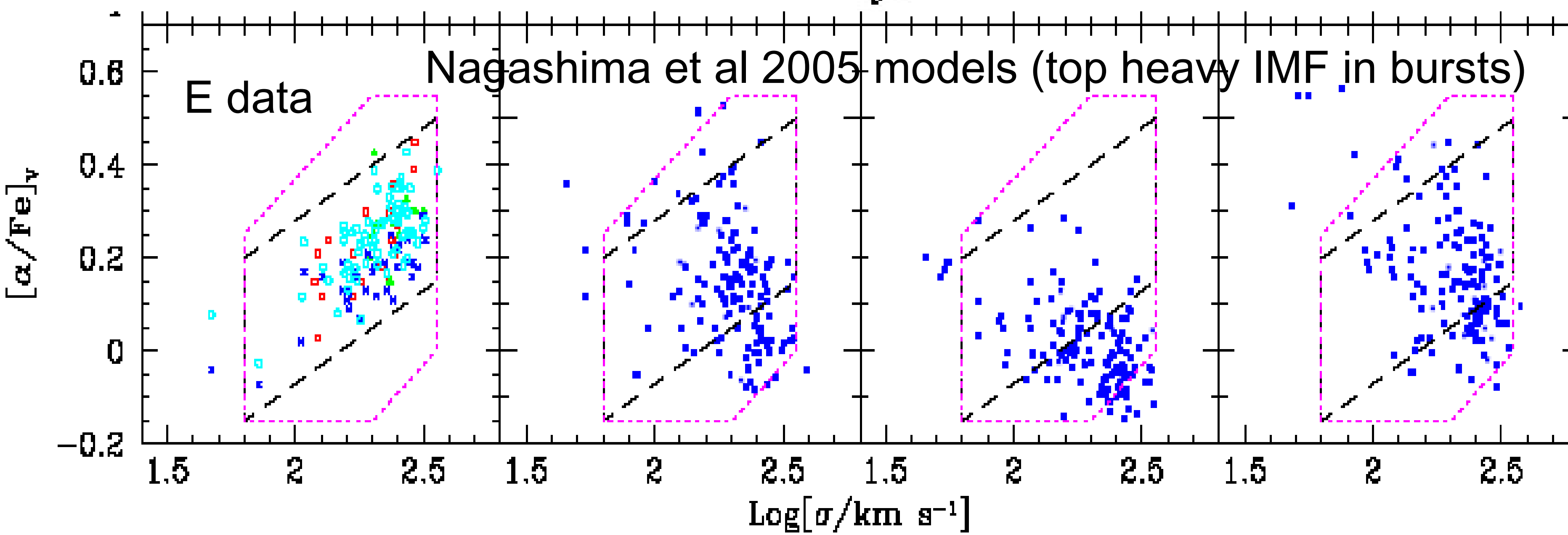
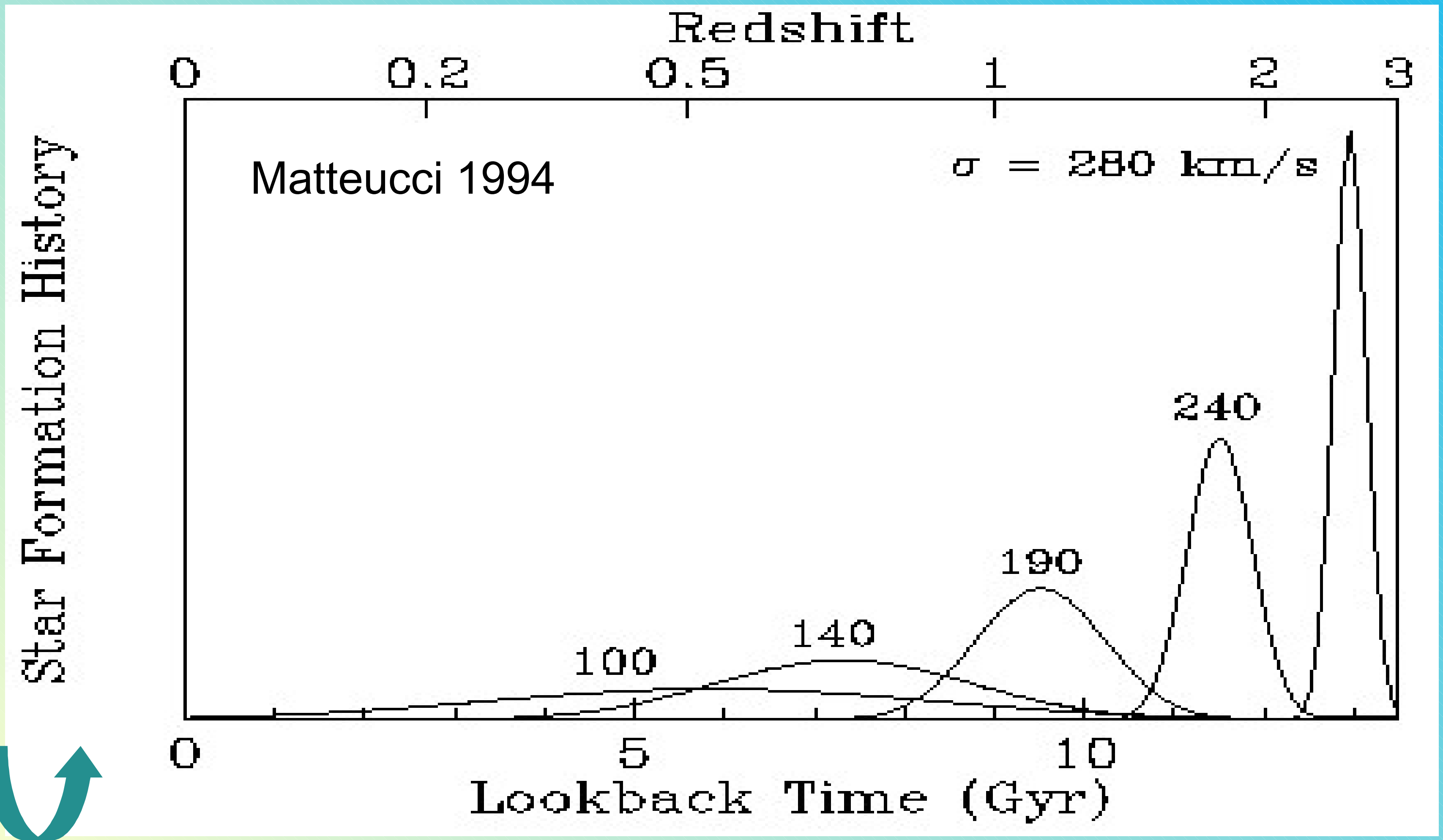


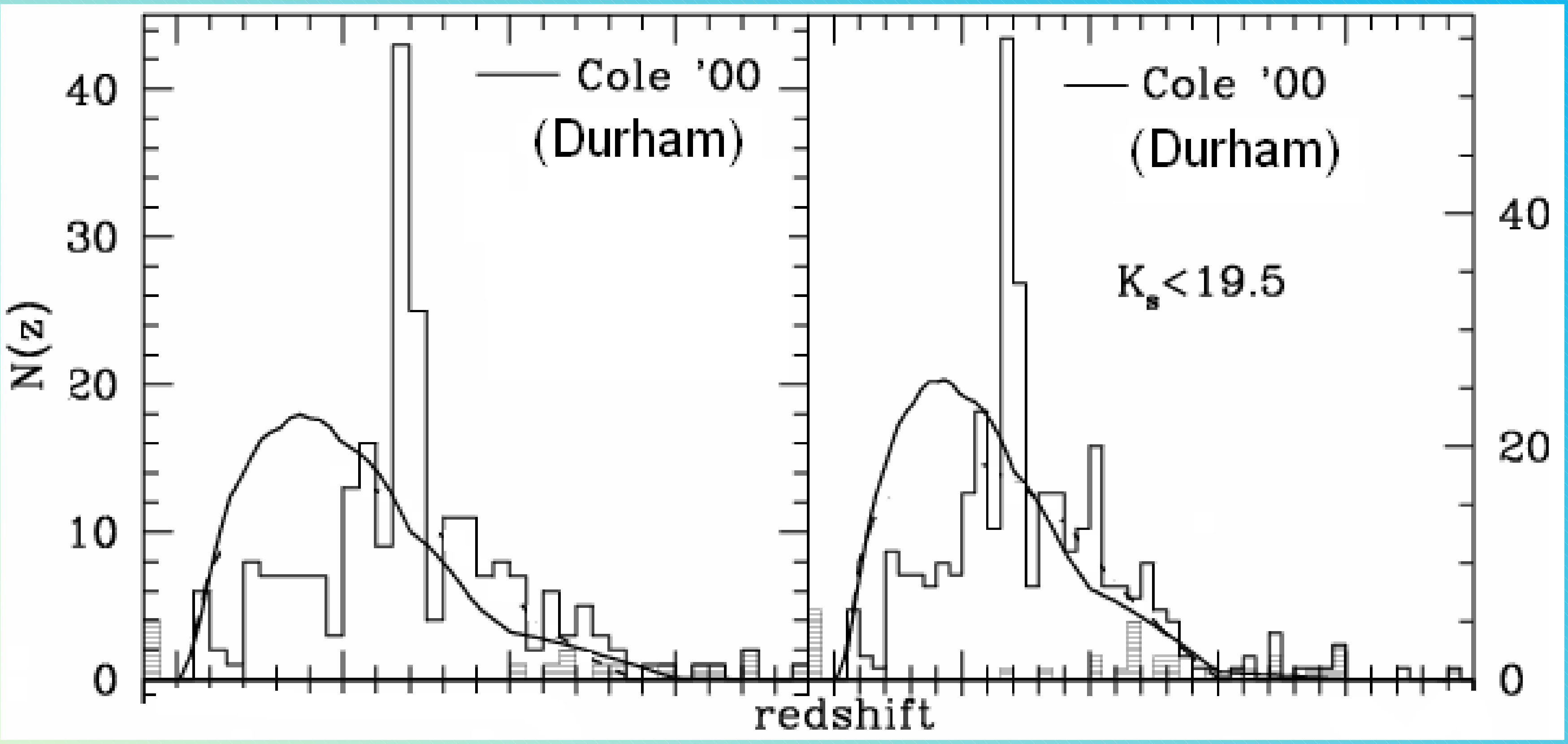
Figure 13. The colour-magnitude relation for cluster elliptical galaxies in the reference model, compared to observations. The points give the predicted distribution of V-K colour versus V-band magnitude for elliptical galaxies in clusters with circular velocity greater than $1000\,\mathrm{km\,s^{-1}}$. The heavy line and error bars indicate the median and the 20 and 80 percentiles of this distribution. The observed correlation and scatter, from Bower, Lucey & Ellis (1992), are indicated by the dotted line and associated error bars.


Thomas 1999
Early type galaxy data
Standard SA Models
[α/Fe]
0.3
0.2
0.1
0
−18
−20
−22
−24
V


Nagashima et al 2005 models (top heavy IMF in bursts)
E data
[α/Fe]v
0.6
0.4
0.2
0
−0.2
1.5
2
2.5
Log[σ/km s⁻¹]

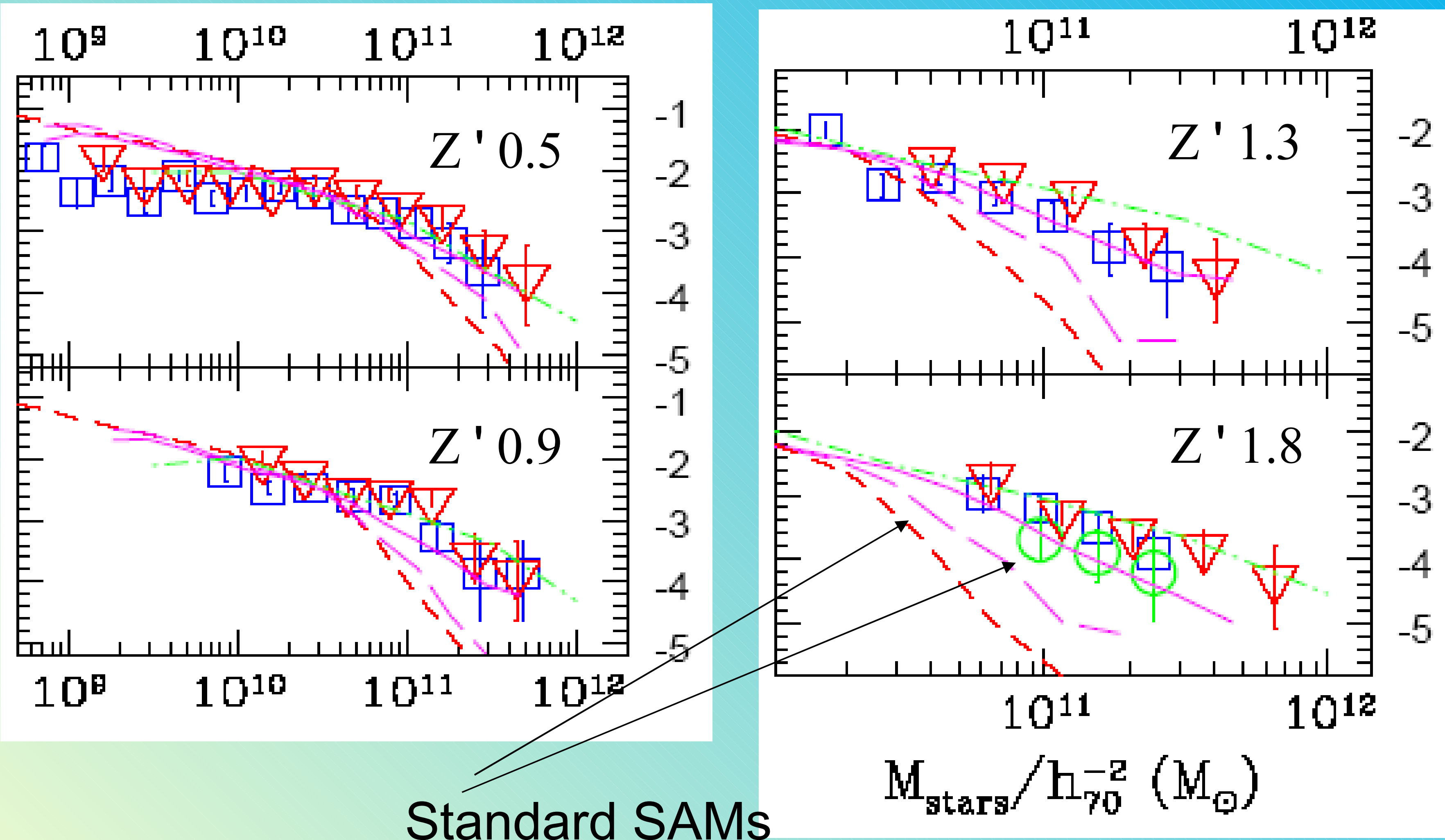
The ‘natural’ explanation would be that E are old and almost coeval, with SF time scale shorter for more massive galaxies. This is impossible to obtain in standard SAM.



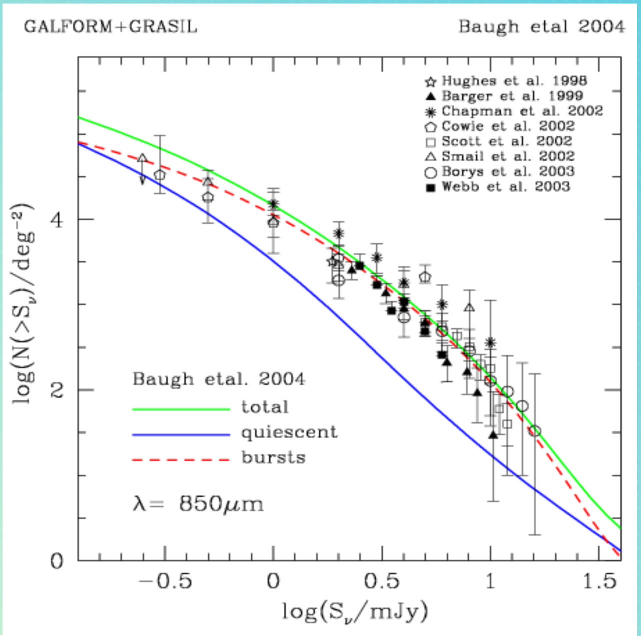


Redshift distributions of deep K-band selected samples show more high z and less low z objects than predicted by standard SAMs

Fontana et al 2004: galaxy stellar mass function in K20 sample



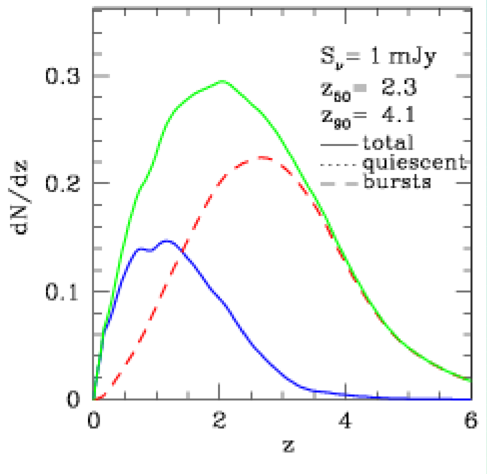
Standard SAMs under produce massive galaxy, by a fraction increasing with z



One may try to fix some problems of standard SAMs pushing parameters within the same general scenario.

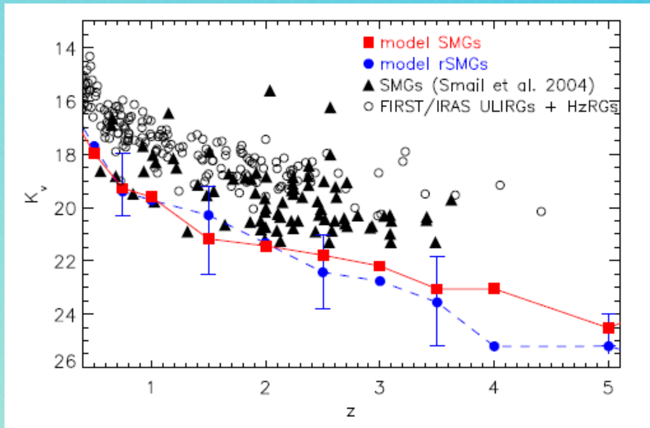
E.g. Baugh et al. (2005, Durham) reproduced SCUBA counts with severe modifications:

1. SF in disks at high redshift suppressed
2. Very top heavy IMF in bursts (*dn/dln m = m*$^{-x}$, *x=0*, rather *than x>1, gives a factor » 5*)
3. Shallower than standard dust emissivity (*gives a factor » 2*)
4. Added another burst mode



But this does not help with other problems, such as chemistry of E or massive high-z galaxies….

….and actually the predicted K band magnitudes suggest that the model galaxies are a factor ten too small (Swinbank et al 2008)



A more radical view?

These observations suggest an assembly of baryons in spheroids mimicking the traditional monolithic scenario, with downsizing.

To get this within hierarchical assembly of dark matter we proposed (Granato et al 2001, 2004) a revision of SAM based on:

1. Reduced role of gas disk formation at high z: cool and collapsing gas in big halos start vigorous SF without setting in a quiescent disk.

2. Keep into account the mutual feed-back between formation of high-z QSO and their host galaxies largely ignored by simulations…..

A more radical view? Our two-phase galaxy formation model: original motivations and general scheme

Problems in standard SAMs, plus evidences of mutual link between SF and AGN activity:

•M_{BH}-spheroid relations (L_{sph}, M_{sph}, σ_{sph})

•Similarity of cosmic SFR(z) and $_{QSO}$}(z)

•High z QSO seem to shine in an evolved environmet

•Simulated galaxy mergers drive gas to the centre

) We proposed (Granato et al 2001, 2004) a revision of SAMs focussed on high-z massive galaxies- Anti-hierarchical Baryonic Collapse –ABC model:

1) Reduced role of gas disk formation at high z: cool collapsing gas in big halos at high-z start vigorous SF due to quick DM halos assembly (Zhao et al. 2003a,b; Diemand et al. 2007)

2) Large SFR promotes the development of SMBH from a seed, which after ' 0.5-1 Gyr powers an high-z QSO.

3) Keep into account the feed-back of the QSO on the ISM that ultimately quenches further SF, neglected by any previous model

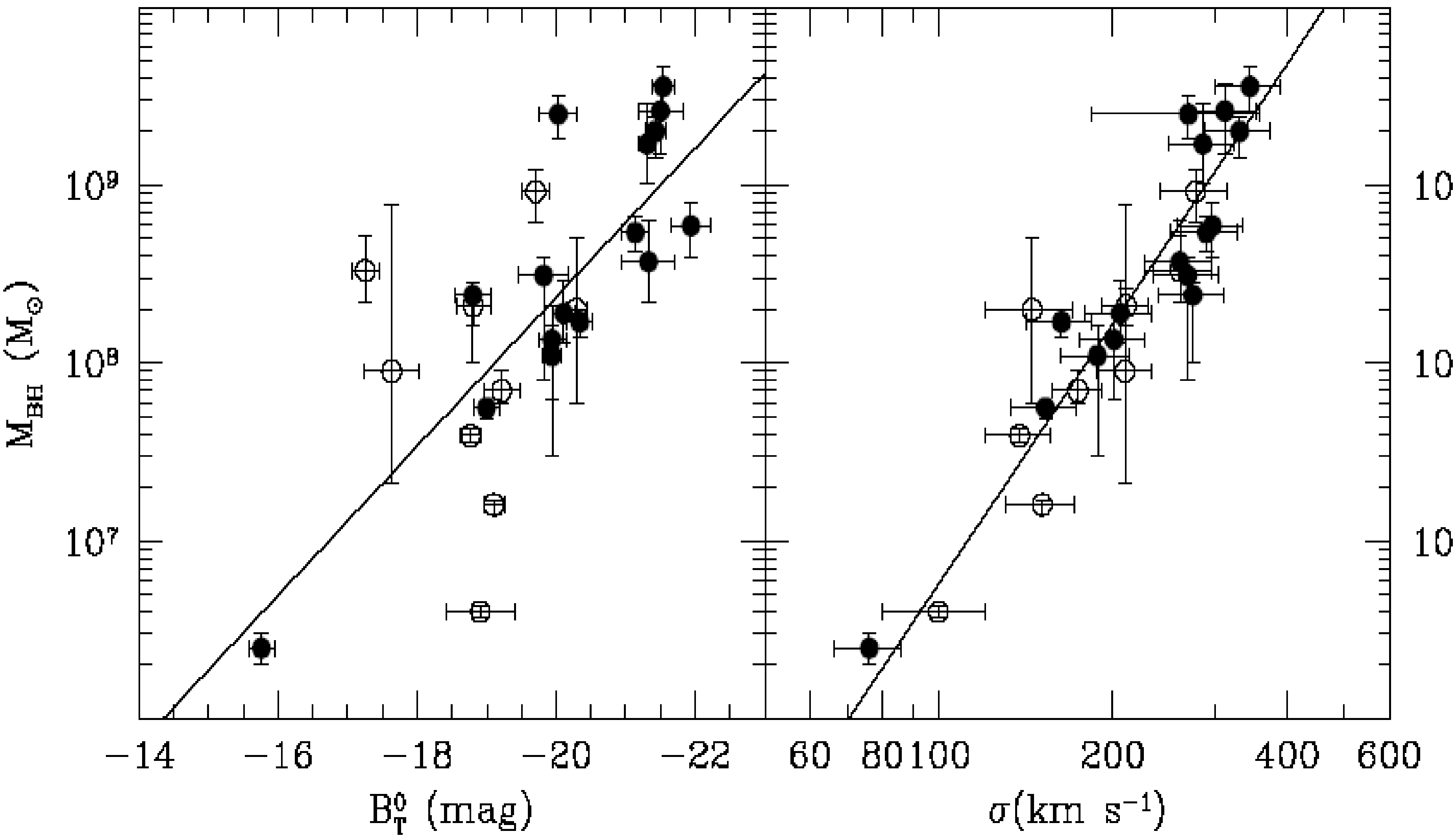
……but hinted by several facts:

• Local spheroids contain a central MDO (SMBH), with $M=10^6$-3×10^9, whose mass function matches that accreted onto BH during QSO activity;

• SMBH mass correlates with properties of the spheroidal component, in particular $M_{BH} \propto M_*$ and $M_{BH} \propto \sigma^{4-5}$;

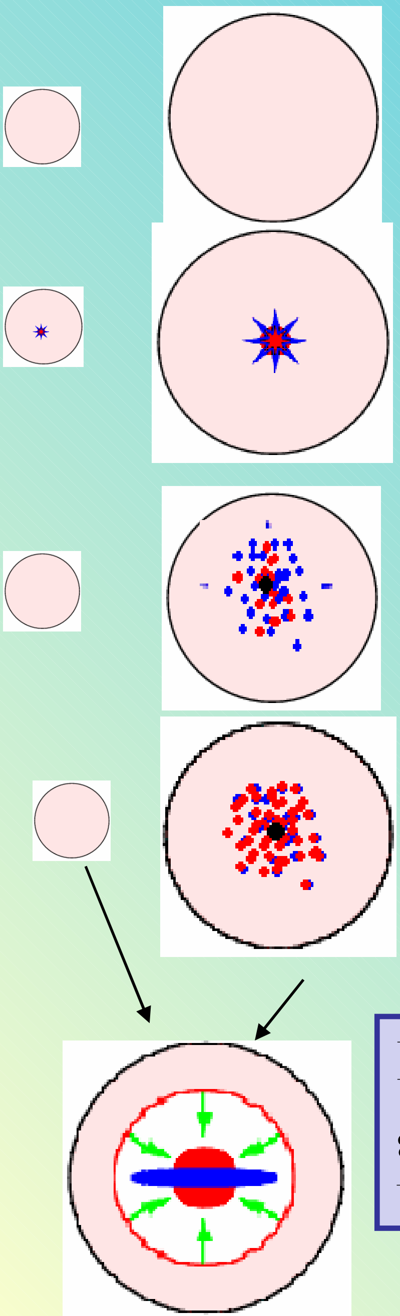
• Spheroidal galaxies are the most common hosts of bright QSOs;

• QSOs at high z are associated to high Z, dusty environments (Hamann & Ferland 1999; Freudling et al 2003, Andreani et al, 1999, Maiolino et al 2003);



From Ferrarese & Ford (2004)

ABC scenario



at high-z halos form quickly, gas is heated to virial T

At high z gas cools, collapse and forms stars directly, in small halos SNae quench SF, in big ones a huge burst of dusty SF (' 1000 M-/yr over 0.5 Gyr), SubMmGalaxies phase...

..with SMBH growth promoted by SF eventually powering high z QSO after » 0.5 Gyr, which cleans ISM and quenchs further SF and then itself. QSO phase followed by...

... ' passive evolution of stellar population. Red and dead massive galaxies at high z (ERO) with dormant SMBH

Possibly a disk form from accretion-minor merger of high J gas around a spheroid. The smaller the halo, the more likely this happens (Cook et al 09)

QSO mode vs Radio mode AGN feedback in SAM

In general, AGN feedback introduced only in the last few years in SAMs (and hydro sim.), quickly becoming very popular.

Two well distinct flavours, with different aims:

- FB associated with the main phase of BH growth, related to the bright high-z QSOs, to sterilize massive high-z galaxies, little affected by SNae (Granato et al 2001, 2004, Monaco & Fontanot 2005; Menci et al 2006)

- FB associated with lower redshift, low accretion rate phase of AGN, optically irrelevant , to halt cooling flows and avoid overproduction of local bright galaxies (Bower et al 2006, Croton et al 2006, see also Cattaneo et al 2006)

A few models now include both (e.g. Somerville et al 2008)

Summary

- First principle models are a myth
- Standard SAM, "galaxy merger driven" have many troubles. Prescriptions for evolution of baryons require substantial revision.
- Evidence is now compelling that in order to better understand the evolution of luminous matter, the mutual influence of evolution of galaxies and quasar has to be incorporated into models. In particular QSO feedback could play a major role in the thermodynamics of clusters and in the evolution of spheroids
- To do this, several aspects need to be considered: central engine, seeds, fuelling, growth, effects on ISM