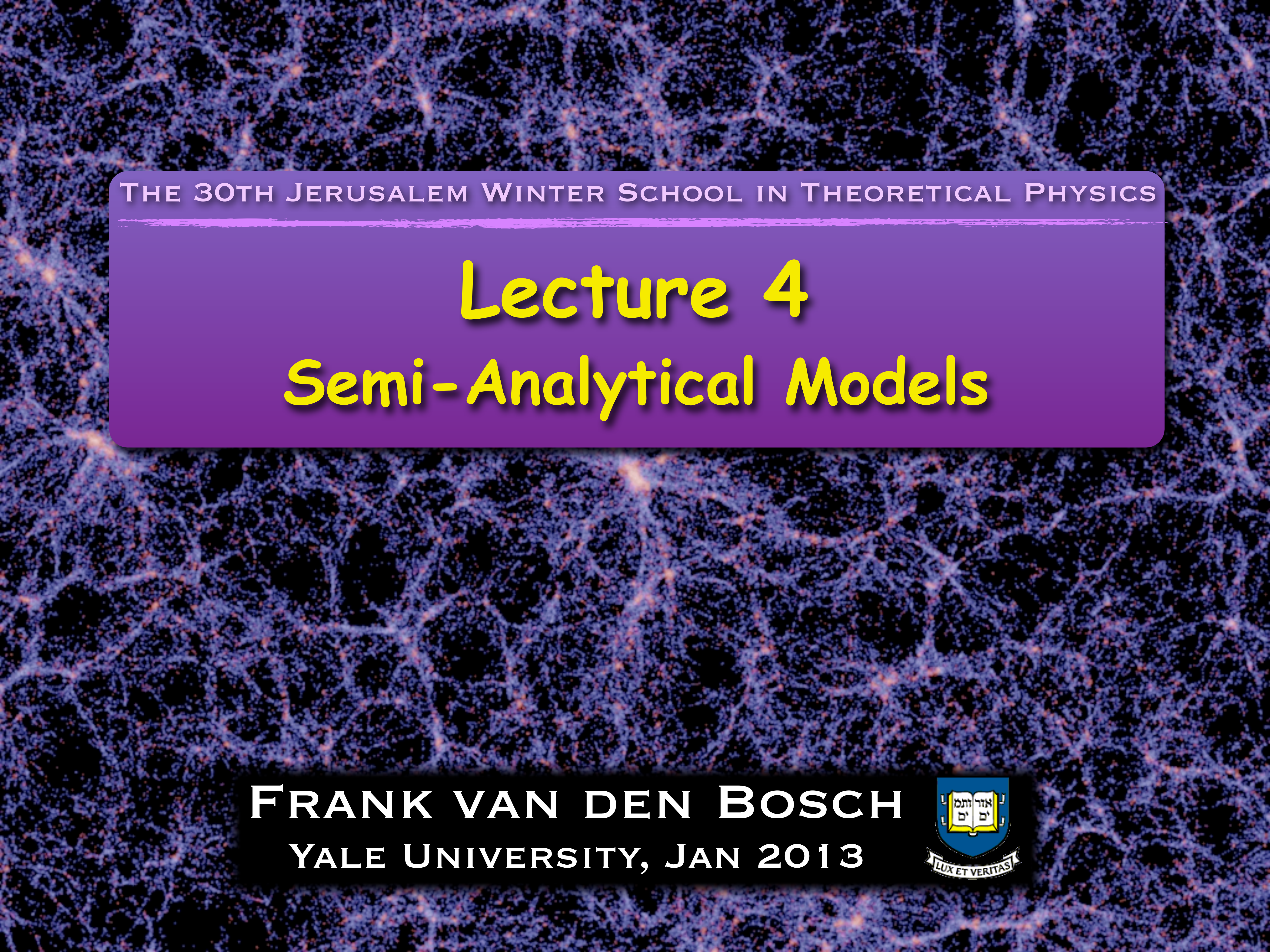


THE 30TH JERUSALEM WINTER SCHOOL IN THEORETICAL PHYSICS

Lecture 4

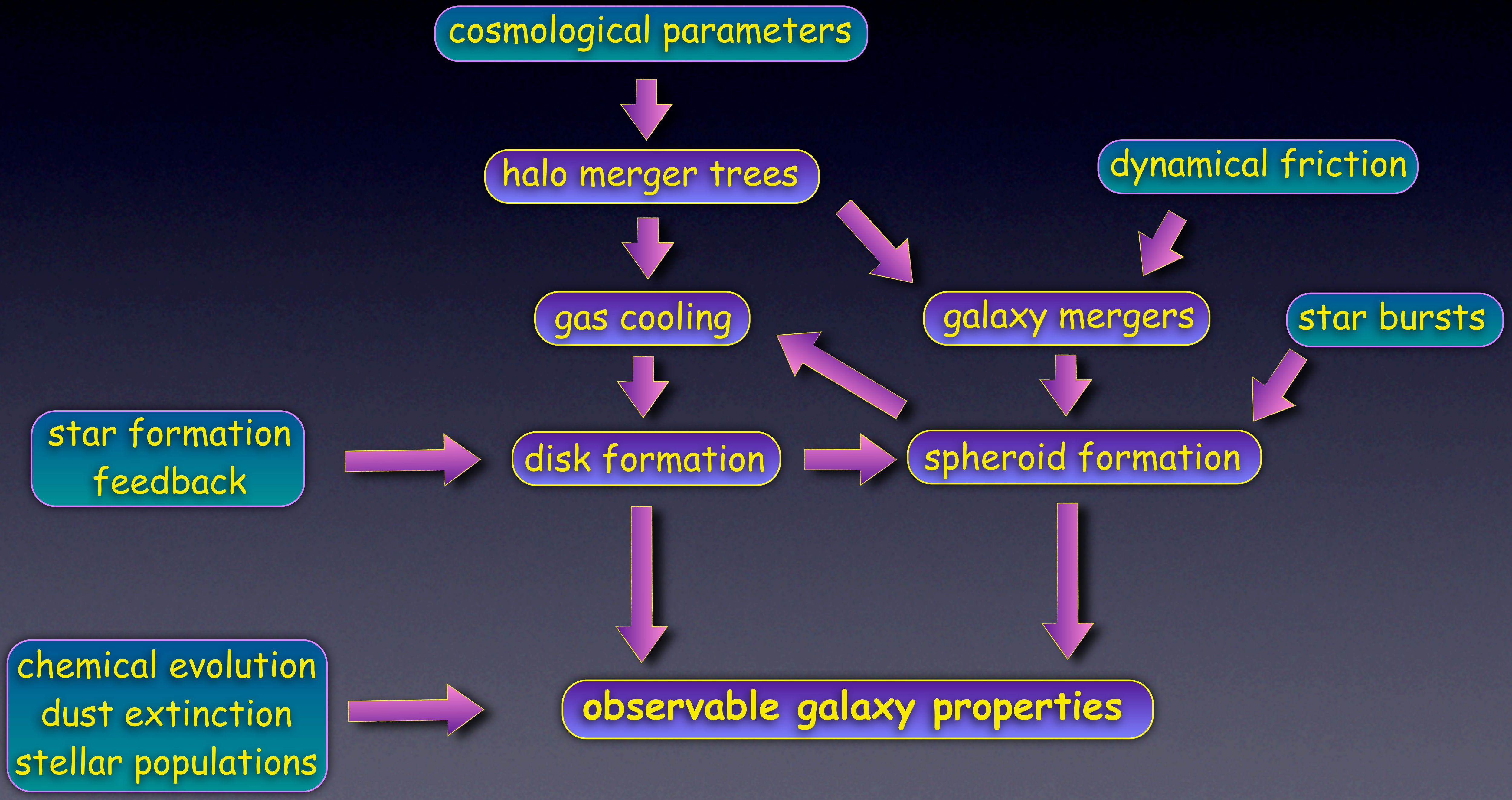
Semi-Analytical Models

FRANK VAN DEN BOSCH

YALE UNIVERSITY, JAN 2013

SAM Basics

Semi-Analytical Models (SAMs) for galaxy formation are phenomenological models that use approximate, analytical descriptions to describe the various processes relevant for galaxy formation in order to make predictions that can be compared to observations.



cosmological parameters

halo merger trees

dynamical friction

gas cooling

galaxy mergers

star bursts

star formation feedback

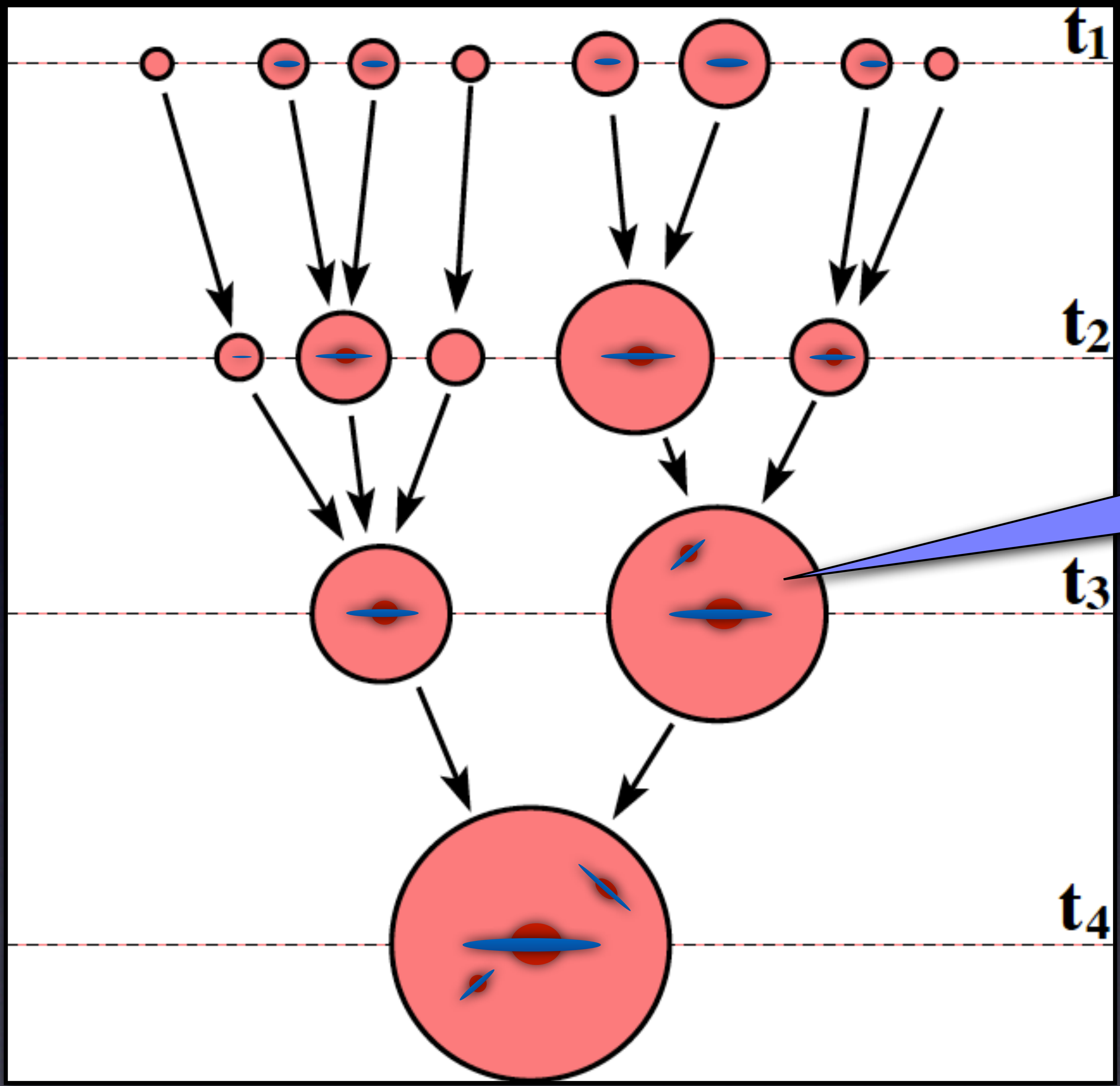
disk formation

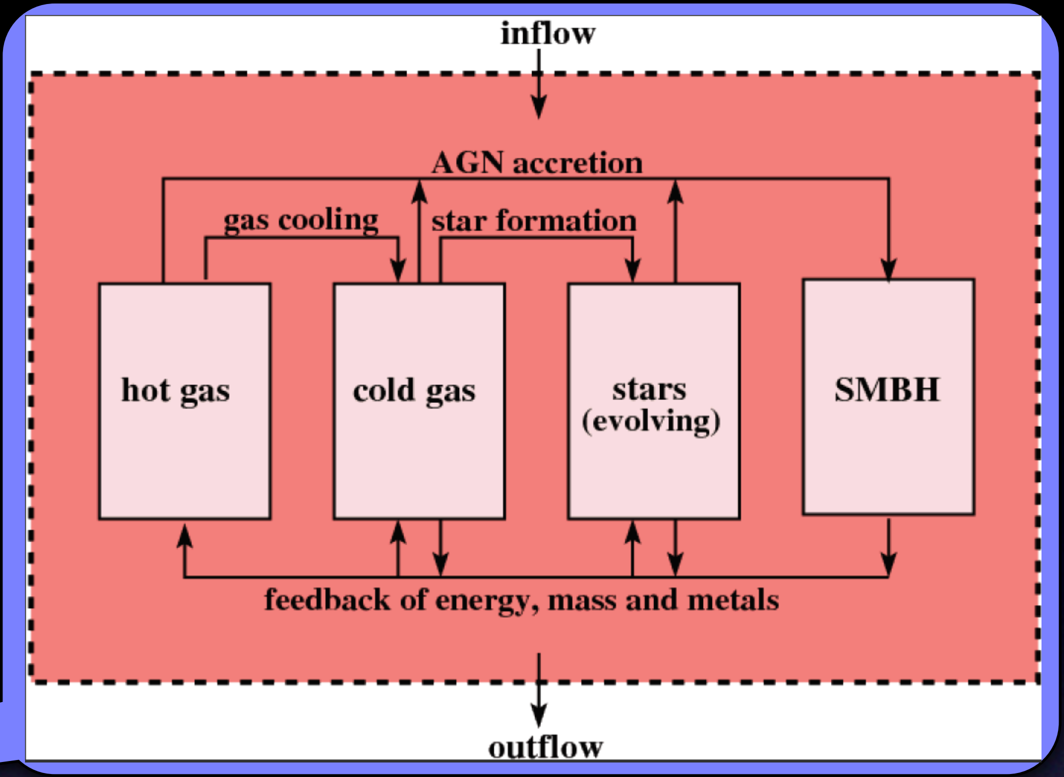
spheroid formation

chemical evolution
dust extinction
stellar populations

observable galaxy properties

SAM Basics




inflow
AGN accretion
gas cooling
star formation
hot gas
cold gas
stars (evolving)
SMBH
feedback of energy, mass and metals
outflow

$$M_{\rm vir} = M_{\rm DM} + M_* + M_{\rm cold} + M_{\rm hot}$$

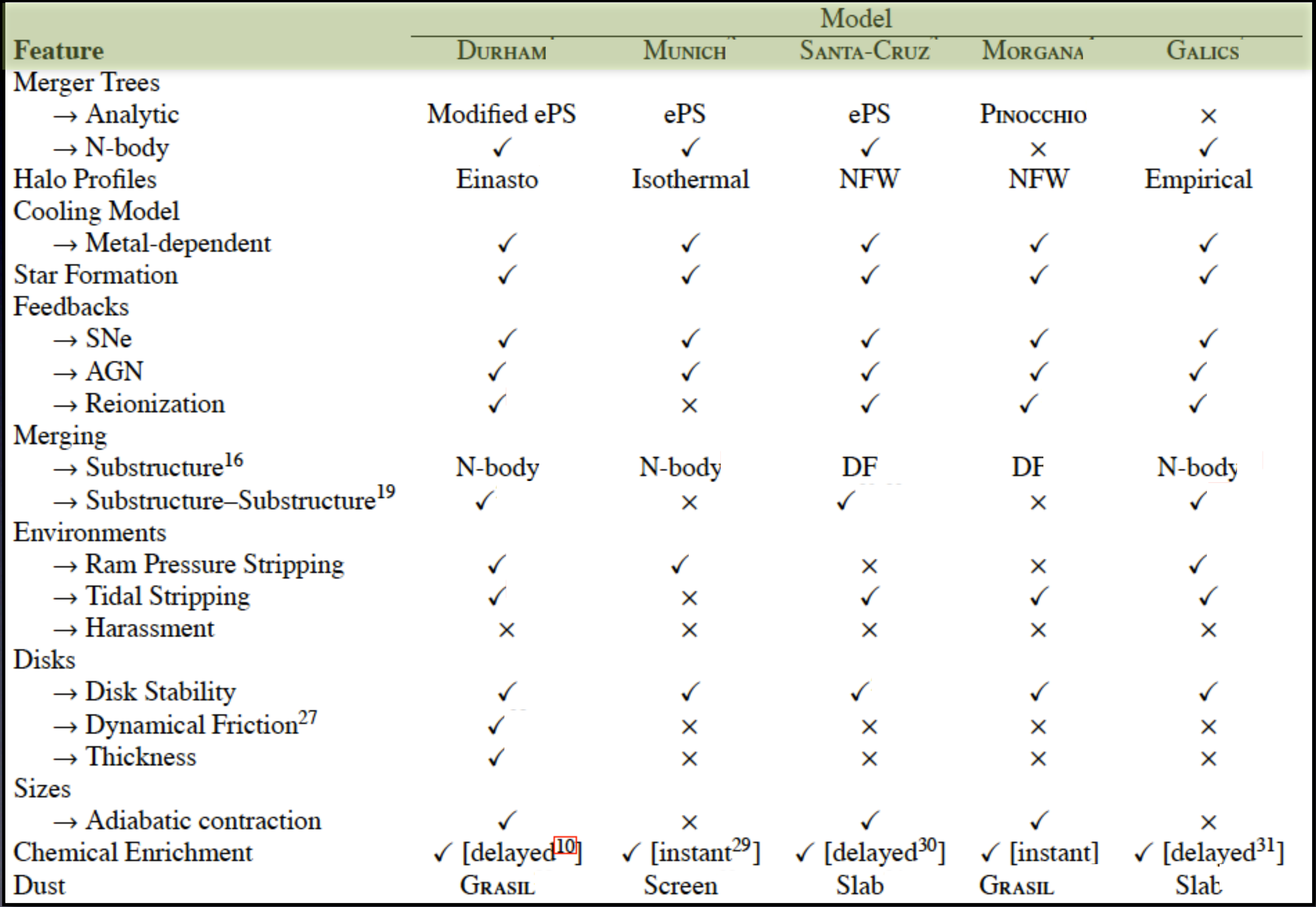
$$\dot{M}_{\rm hot} = -\dot{M}_{\rm cool} + \dot{M}_{\rm reheat} + f_{\rm bar}\,\dot{M}_{\rm vir}$$

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

$\dot{M}_{\rm vir}$	⟷	mass accretion
$\dot{M}_{\rm cool}$	⟷	cooling
$\dot{M}_*$	⟷	star formation
$\dot{M}_{\rm reheat}$	⟷	feedback

SAM Basics

Currently there are many different SAMs available in the literature, all with different treatments of different physical processes....



	Model				
Feature	Durham	Munich	Santa-Cruz	Morgana	Galics
Merger Trees					
→ Analytic	Modified ePS	ePS	ePS	Pinocchio	×
→ N-body	✓	✓	✓	×	✓
Halo Profiles	Einasto	Isothermal	NFW	NFW	Empirical
Cooling Model					
→ Metal-dependent	✓	✓	✓	✓	✓
Star Formation	✓	✓	✓	✓	✓
Feedbacks					
→ SNe	✓	✓	✓	✓	✓
→ AGN	✓	✓	✓	✓	✓
→ Reionization	✓	×	✓	✓	✓
Merging					
→ Substructure[16]	N-body	N-body	DF	DF	N-body
→ Substructure–Substructure[19]	✓	×	✓	×	✓
Environments					
→ Ram Pressure Stripping	✓	✓	×	×	✓
→ Tidal Stripping	✓	×	✓	✓	✓
→ Harassment	×	×	×	×	×
Disks					
→ Disk Stability	✓	✓	✓	✓	✓
→ Dynamical Friction[27]	✓	×	×	×	×
→ Thickness	✓	×	×	×	×
Sizes					
→ Adiabatic contraction	✓	×	✓	✓	×
Chemical Enrichment	✓ [delayed[10]]	✓ [instant[29]]	✓ [delayed[30]]	✓ [instant]	✓ [delayed[31]]
Dust	Grasil	Screen	Slab	Grasil	Slab

Source: Benson 2010, Physical Reports, 495, 33

SAM Basics

Currently there are many different SAMs available in the literature, all with different treatments of different physical processes....

And each `individual' SAM is constantly being updated/expanded/modified...

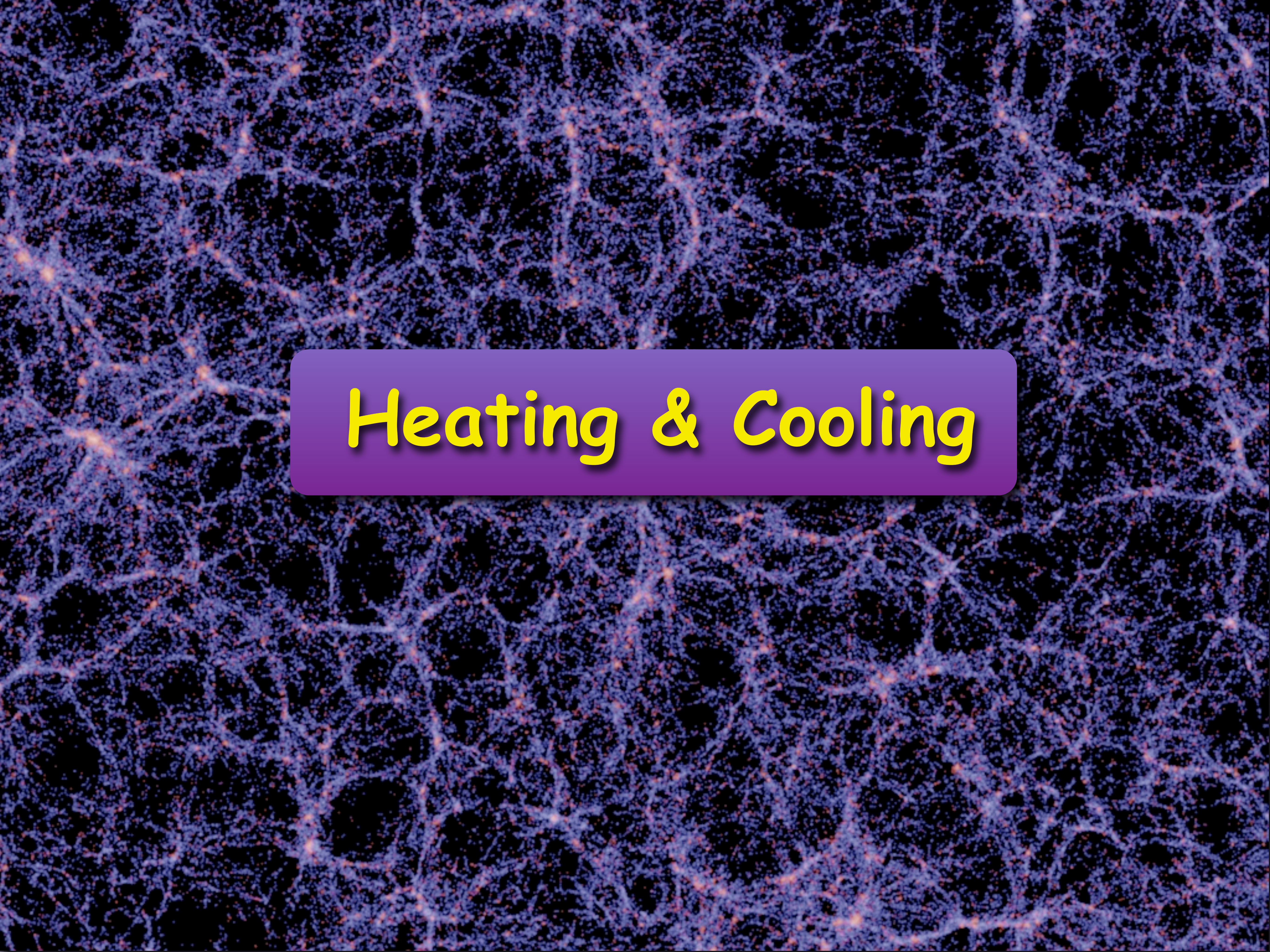
A 1.5hr lecture can't possibly do justice to the immense amount of work done on and with SAMs. Furthermore, a comparison among different models is largely meaningless...

Instead, we will focus on the basic structure of SAMs, discuss the treatment of some of the main ingredients (cooling, SF, feedback, dynamical friction), and examine the current status of some of the models, focussing only on

- luminosity/stellar mass functions at z=0
- galaxy-dark matter connection (clustering & TF relation)
- stellar mass assembly histories

Topics that are important but which I won't be able to cover include

- Stellar Population Synthesis Modeling
- Dust (absorption & emission)
- Chemical Evolution (metallicities & abundances)
- Galaxy Sizes
- BH-Bulge mass relation
- Disk Instabilities



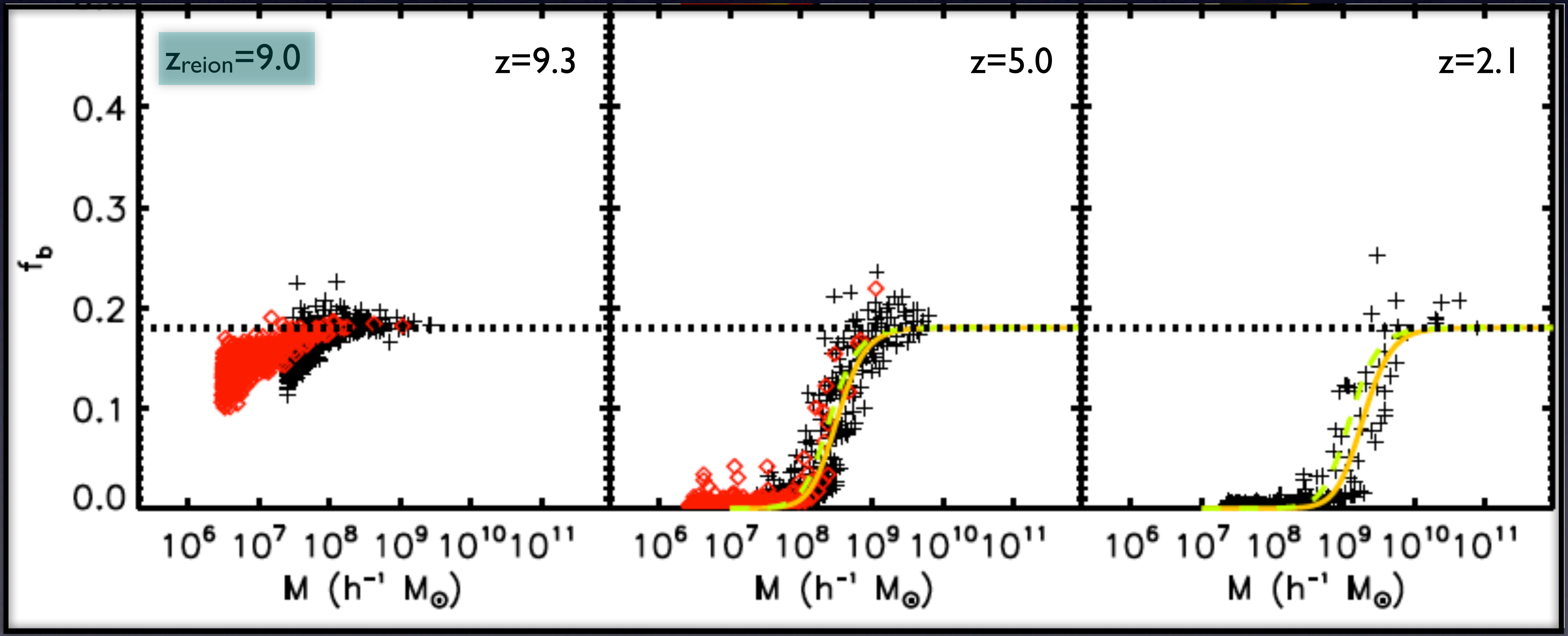
Heating & Cooling

Reionization

The fraction of baryonic matter that is bound to a dark matter halo is regulated by the potential depth and the pressure of the baryons.

When the Universe reionizes, the temperature and thus the pressure of the baryons increases. This results in a strongly reduced baryon fraction in low mass halos.

Okamoto et al. (2008) used hydro simulations to study the impact of reinization imposing a uniform, ionizing background as computed by Haardt & Madau (2001).



Source: Okamoto et al. 2008, MNRAS, 390. 920,

Reionization

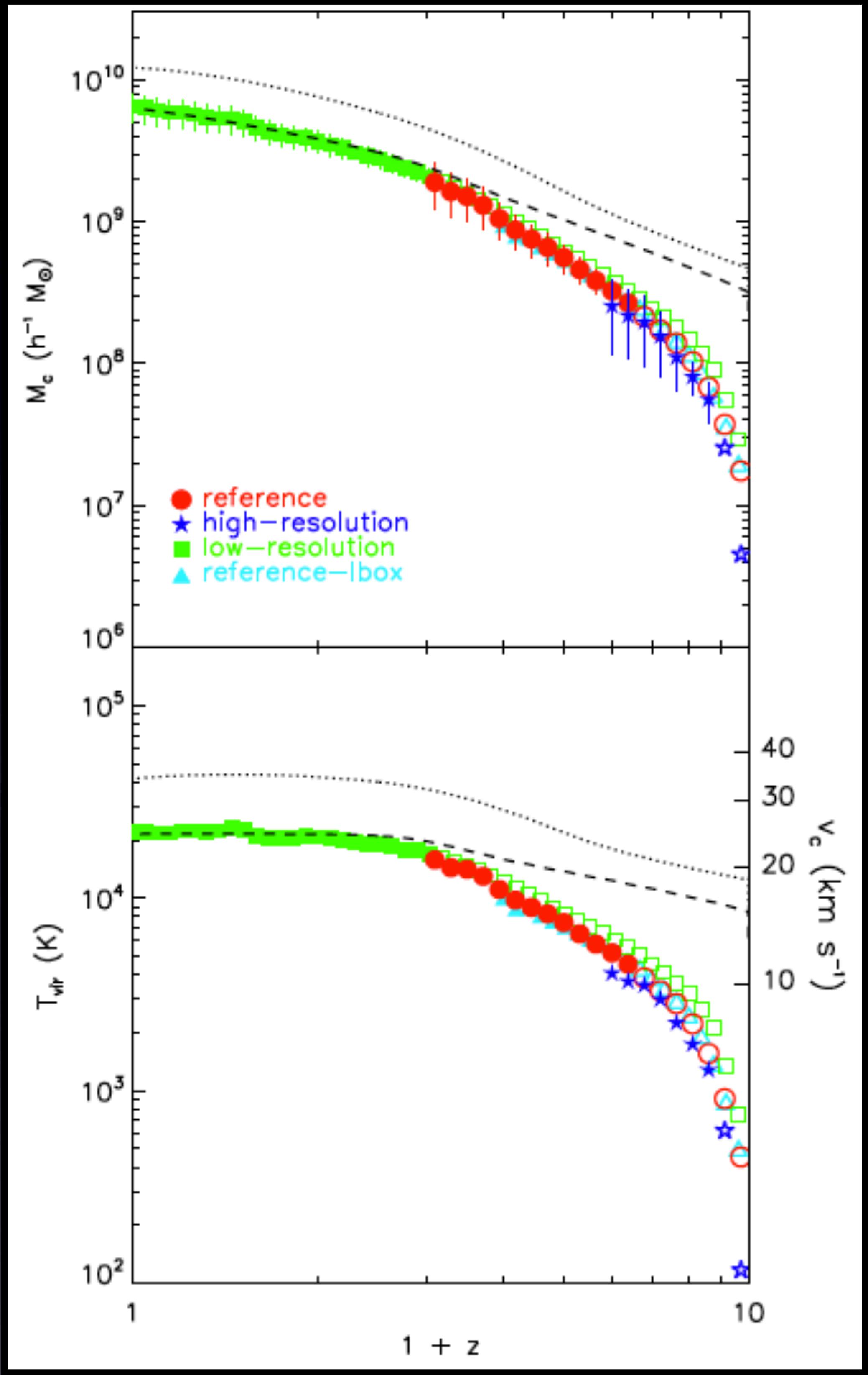
The fraction of baryons that remains bound to their halos is well fit by the following functional form, first introduced by Gnedin (2000):

$$f_{\rm b}(M,z) = \frac{\Omega_{\rm b}}{\Omega_{\rm m}} \left[1 + (2^{2/3} - 1)\left(\frac{M}{M_{\rm c}(z)}\right)^{-2}\right]^{-3/2}$$

Here $M_{\rm c}(z)$ is the (redshift-dependent), characteristic mass at which half the baryons are photo-evaporated from their host halo..

In the simulations of Okamoto et al. (2008), in which reionization occurs at z=9, the characteristic mass increases from ~10^8 $M_\odot$ shortly after reionization, to ~7×10^9 $M_\odot$ at z=0.

Latter corresponds to a virial velocity of ~25 km/s



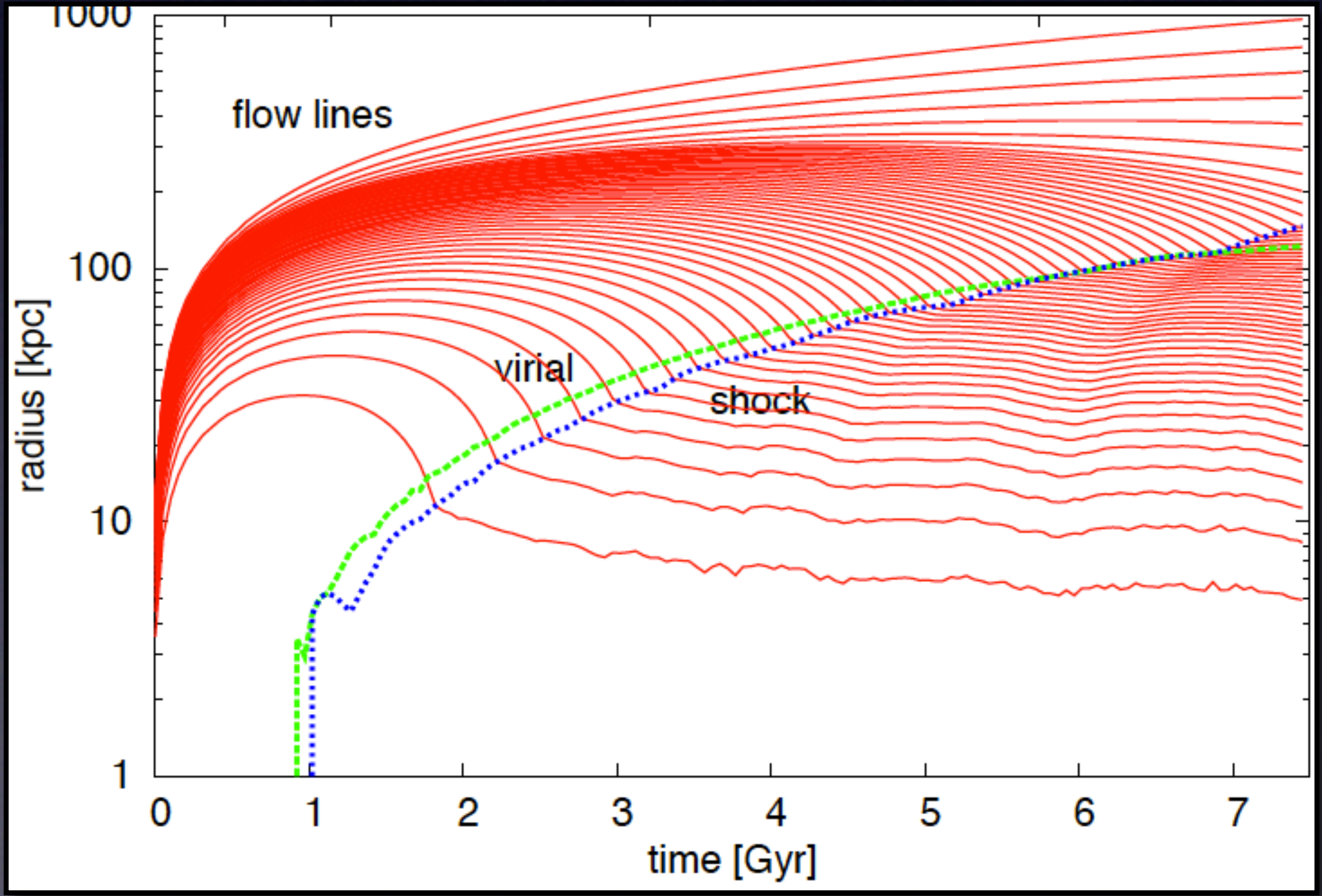
Source: Okamoto et al. 2008, MNRAS, 390. 920,

Accretion Shocks & Virial Temperature

When infalling gas passes the halo's accretion shock, its kinetic energy is thermalized, and therefore heated to of order the virial temperature

$$T_{\rm vir} = \frac{\mu m_{\rm p}}{2\,k_{\rm B}}\,V_{\rm vir}^2 \simeq 3.6 \times 10^5\,{\rm K}\left(\frac{V_{\rm vir}}{100\,{\rm km/s}}\right)^2$$

where we have assumed that dark matter halos are singular isothermal spheres, and we have adopted that $\mu = 0.59$, appropriate for a primordial gas (X,Y,Z) = ($\frac{3}{4}$,$\frac{1}{4}$,0)



Source: Birnboim & Dekel, 2003, MNRAS, 345, 349

CAUTION: in general, gas inside halo has T-profile, and cannot be described by single T. Nevertheless, concept of $T_{\rm vir}$ is useful for order of magnitude estimates, and is therefore frequently used.

The build-up of a virial shock (discontinuity in velocity) at around the virial radius in a collapsing structure. Based on 1D calculations in an expanding Universe...

Radiative Cooling

The primary cooling processes relevant for galaxy formation are two-body radiative processes in which gas loses energy through the emission of photons as a consequence of two-body interactions. The main processes are:

	type	reaction	name
1	free-free	$e^- + X^+ \rightarrow e^- + X^+ + \gamma$	bremsstrahlung
2	free-bound	$e^- + X^+ \rightarrow X + \gamma$	recombination
3	bound-free	$e^- + X \rightarrow X^+ + 2e^-$	collisional ionization
4	bound-bound	$e^- + X \rightarrow e^- + X'$ $\rightarrow e- + X + \gamma$	collisional excitation

In order to compute the cooling rate, one needs to know the above reaction rates, as well as the number densities of the various ionic species. In the case of a pure H/He mixture (the simplest, relevant case), these are $n_{\mathrm{e}}, n_{\mathrm{H}_0}, n_{\mathrm{H}^+}, n_{\mathrm{He}_0}, n_{\mathrm{He}^+}, n_{\mathrm{He}^{++}}$

One typically makes the following two assumptions:

- gas is optically thin (i.e., *Case A recombination*; all photons generated escape)
- collisional ionization equilibrium (CIE; number densities are in equilibrium)

Cooling Time

It is useful to define the **cooling function**:

$$\Lambda(T,Z) \equiv \frac{\mathcal{C}}{n_{\rm H}^2}$$

$$[\Lambda] = {\rm erg\,s^{-1}\,cm^{+3}}$$

which depends on the temperature, T, and composition (metallicity Z) of the gas, but not on its density, and where $\mathcal{C}$ is the volumetric cooling rate (in erg/s/cm^3)

The cooling time, the time it takes the gas to radiate away its internal energy, is given by

$$t_{\rm cool} \equiv \frac{\rho\,\varepsilon}{\mathcal{C}} = \frac{3n\,k_{\rm B}T}{2\,n_{\rm H}^2\,\Lambda(T)}$$

→ denser gas cools faster...

To assess impact of cooling, compare cooling time to two other timescales:

- the age of the Universe, which is roughly the Hubble time

$$t_{\rm H} = \frac{1}{H(z)} \propto \frac{1}{(G\bar{\rho})^{1/2}} \qquad \bar{\rho} = \Omega_{\rm m}\rho_{\rm crit}$$

- the dynamical time (or `free-fall time') of the system

$$t_{\rm ff} = \left(\frac{3\pi}{32G\bar{\rho}_{\rm sys}}\right)^{1/2} \propto \frac{1}{(G\,\bar{\rho}_{\rm sys})^{1/2}} \qquad \bar{\rho}_{\rm sys} = \bar{\rho}_{\rm gas} + \bar{\rho}_{\rm DM}$$

Cooling Time

$\bar{\rho}_{\rm sys} \sim 200\bar{\rho}$ ⇒ $t_{\rm ff} \sim t_{\rm H}/10$

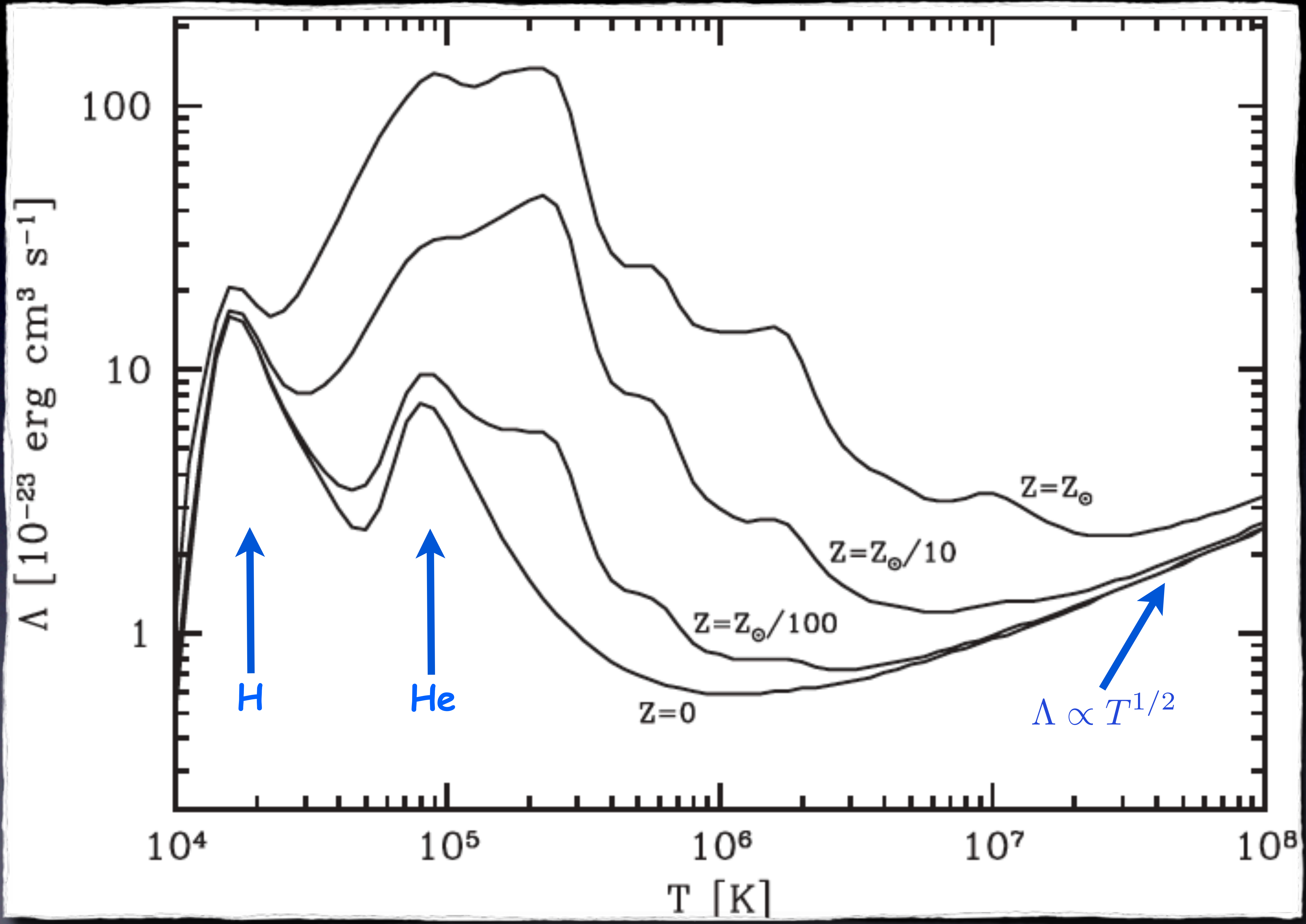
We distinguish three regimes:

$t_{\rm cool} > t_{\rm H}$	Cooling is not important. Gas is in hydrostatic equilibrium, unless it was recently disturbed
$t_{\rm ff} < t_{\rm cool} < t_{\rm H}$	System is in quasi-hydrostatic um. It evolves on cooling time scale. Gas contract cools, but system has sufficient time to cor re-establish hydrostatic equilibrium. **hot mode**
$t_{\rm cool} < t_{\rm ff}$	Cooling is catastropic. Gas cannot r nd fast enough to loss of pressure. Since cooling time with increasing density, cooling proceeds fast er (=catastrophic). Gas falls to center of dynamic sys on free-fall time... **cold mode**

$$t_{\rm cool} \propto \rho_{\rm gas}^{-1} \propto (1+z)^{-3}$$
$$t_{\rm ff} \propto \rho^{-1/2} \propto (1+z)^{-3/2}$$

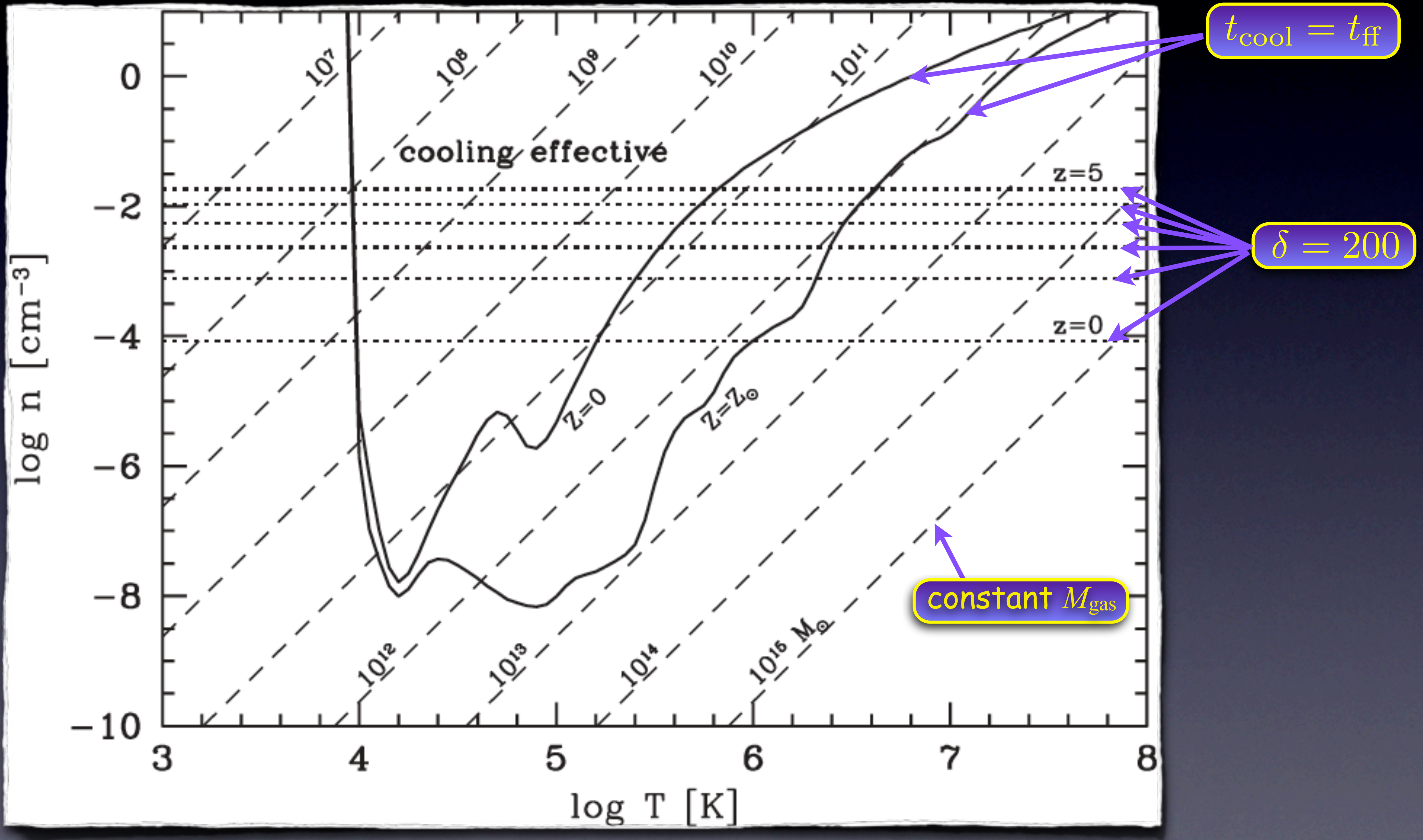
⇒ cooling is more efficient at higher redshifts ⇒ **overcooling problem**

The CIE Cooling Function

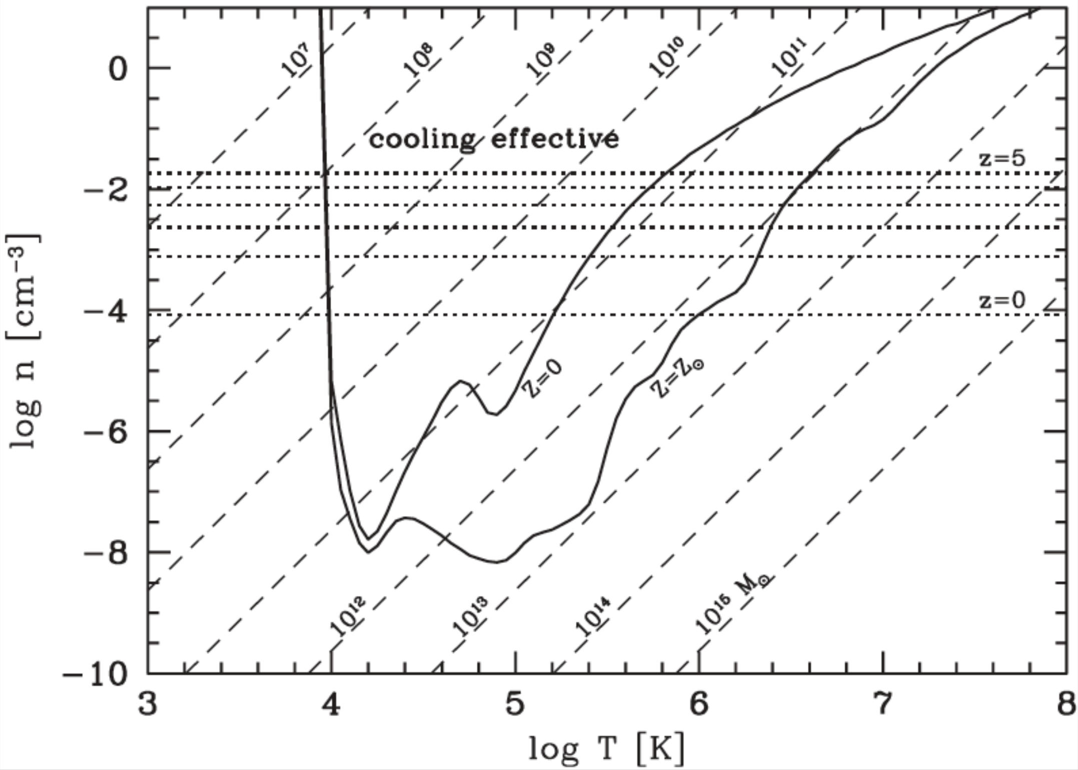


Cooling & Galaxy Formation

Under the assumption of CIE, we can compute $\frac{t_{\rm cool}}{t_{\rm ff}} = \frac{t_{\rm cool}}{t_{\rm ff}}(n, T, Z)$



Cooling & Galaxy Formation



$M_{\rm gas} = 0.15 M_{\rm vir}$ → $M_{\rm vir} \simeq 6.6 M_{\rm gas}$

- Haloes with $M_{\rm vir} < 10^7 M_\odot$ $(V_{\rm vir} < 20\,{\rm km/s})$ can't cool their gas (except by molecular cooling...)
- Haloes with $M_{\rm vir} > 10^{12} M_\odot (Z=0)$
  $M_{\rm vir} > 10^{13} M_\odot (Z = Z_\odot)$
  can't cool their gas efficiently either...

In early papers (and textbooks) on galaxy formation, this mass scale of $10^{12} - 10^{13} M_\odot$ was invoked to exlain the exponential cut-off in the luminosity/stellar mass function of galaxies; more massive galaxies can't form because they can't efficiently cool their gas...
(e.g., Binney, 1977, ApJ, 215, 483; Silk, 1977, ApJ, 211, 683; Ostriker & Rees, 1977, MNRAS, 179, 541)

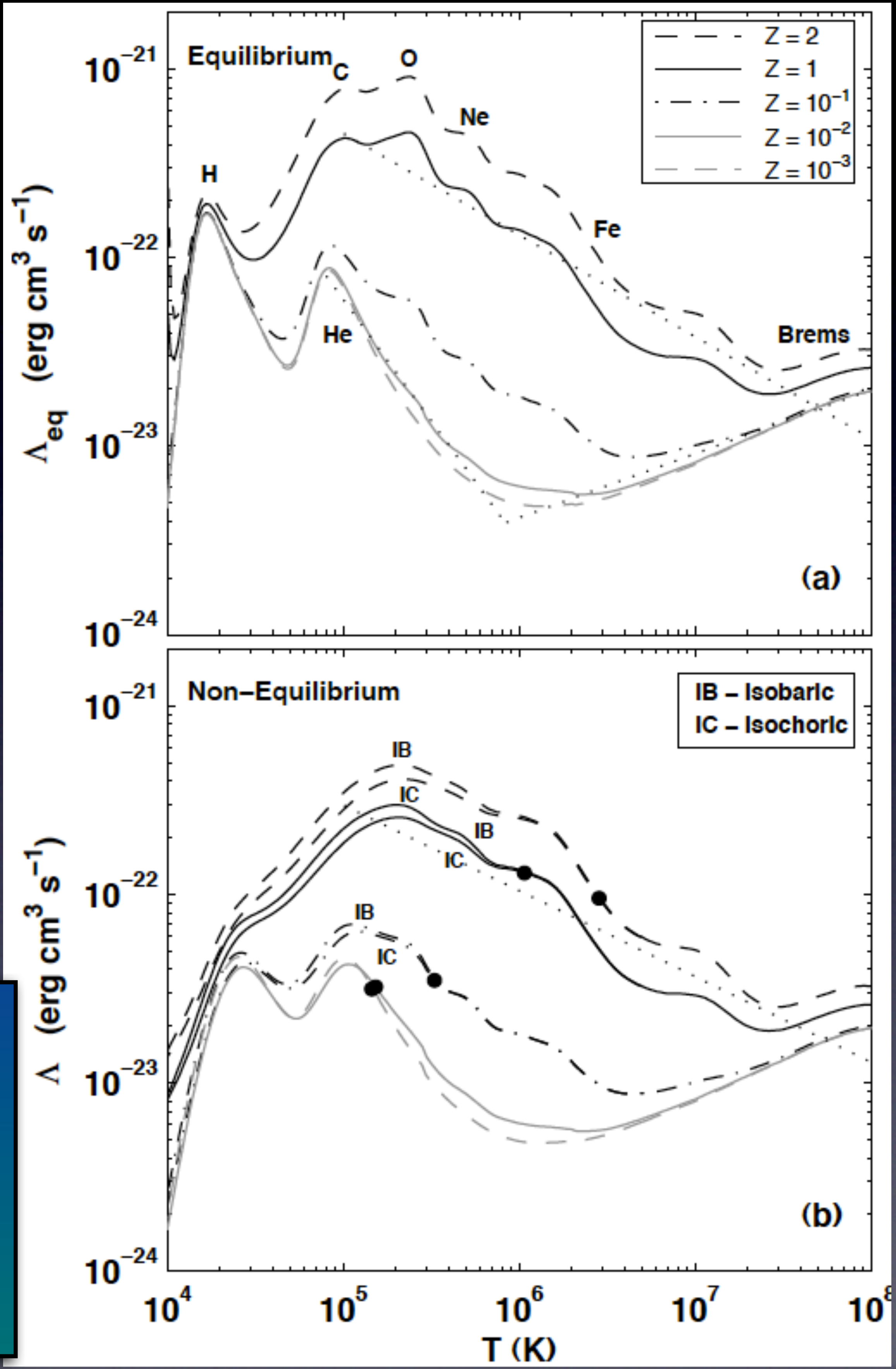
However, this argument is seriously flawed for two reasons:

- Haloes and galaxies form hierarchically → the progenitors of massive haloes can cool, especially at higher redshifts...
- The curve $t_{\rm cool} = t_{\rm ff}$ is calculated for an overdensity $\delta = 200$. The gas in a halo typically has a density profile, and can have $\delta \gg 200$ near the center.
  → at least some fraction of the gas should have cooled...

Deviations from CIE

In particular for low T gas, the recombination time-scale can exceed the cooling time scale of the gas, so that, while it cools, it starts to drift away from CIE. This can result in significantly different cooling times (by up to factors of two to three).

To my knowledge this is not accounted for in any SAM...



Source: Gnat & Sternberg, 2007, ApJS, 168, 213

Comparison of cooling functions for gas in CIE (upper panel) with the effective cooling functions for gas in which time-dependent ionization states are computed throughout the cooling (lower panel). IB and IC correspond to cases in which gas is assumed to be cooling isobarically or isochorically...

Cooling in SAMs

Almost all SAMs treat cooling by defining a cooling radius, $r_{\rm cool}$, as the radius where

$$t_{\rm cool}(r) = \frac{3\, n(r)\, k_{\rm B}\, T(r)}{2\, n_{\rm H}^2(r)\, \Lambda(T)} = \tau_{\rm cool}$$

Here $\tau_{\rm cool}$ is some `characteristic' time, which varies from SAM to SAM.

Some SAMs set it equal to the Hubble time, others to the time since the last major merger, or the free-fall time....Computing this cooling radius also requires assumptions about the density and temperature profiles of the hot gas, which also differ from SAM to SAM....

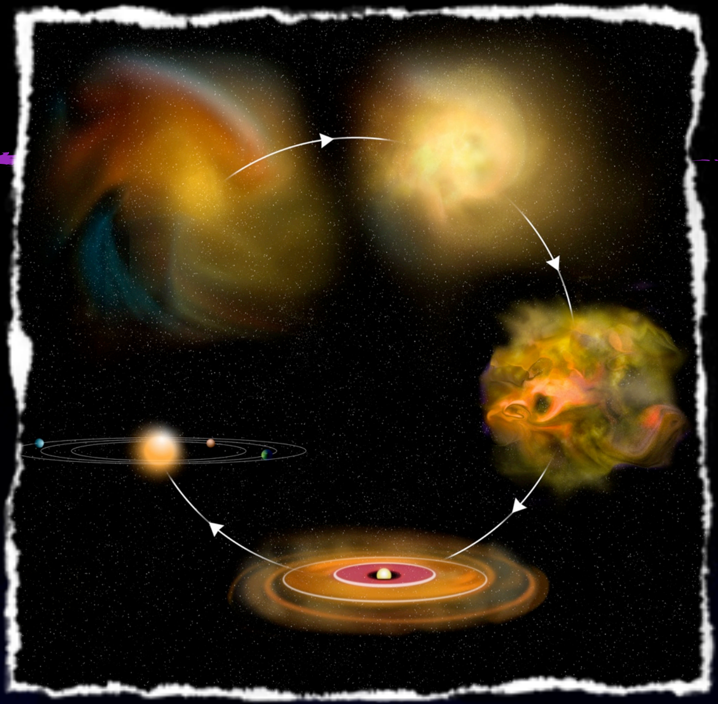
Once the cooling radius has been computed, the cooling rate is computed as follows:

$$\dot{M}_{\rm cool} = \begin{cases} \frac{M_{\rm hot}}{t_{\rm ff}} & \text{if } r_{\rm cool} > r_{\rm vir} \\ \frac{M_{\rm hot}(<r_{\rm cool})}{\tau_{\rm cool}} & \text{if } r_{\rm cool} \leq r_{\rm vir} \end{cases}$$

Typically cooling rates differ by factors two to three from SAM to SAM,
but no SAM includes free parameter(s) to scale their cooling rates...

Star Formation

Star Formation in SAMs



Lacking a detailed theory for star formation, the treatment of SF in SAMs has thus far been very pragmatic:

Old SAMs: no internal structure of galaxies

$$\dot{M}_* = \varepsilon_{\rm SF} \frac{M_{\rm cold}}{\tau_{\rm SF}} \qquad \tau_{\rm SF} = t_{\rm dyn} \left(\frac{V_{\rm vir}}{200\,{\rm km/s}} \right)^{\alpha_{\rm SF}}$$

free parameters $(\varepsilon_{\rm SF}, \alpha_{\rm SF})$ vary from SAM to SAM with $-2.5 \leq \alpha_{\rm SF} \leq 0$

Modern SAMs: disk galaxies are assumed to be exponential

$$\Sigma_{\rm cold}(R) = \left(\frac{M_{\rm cold}}{2\pi R_{\rm d}^2} \right) {\rm e}^{-R/R_{\rm d}} \qquad R_{\rm d} \propto \lambda_{\rm DM}\, r_{\rm vir}$$

(Mo, Mao & White 1998)

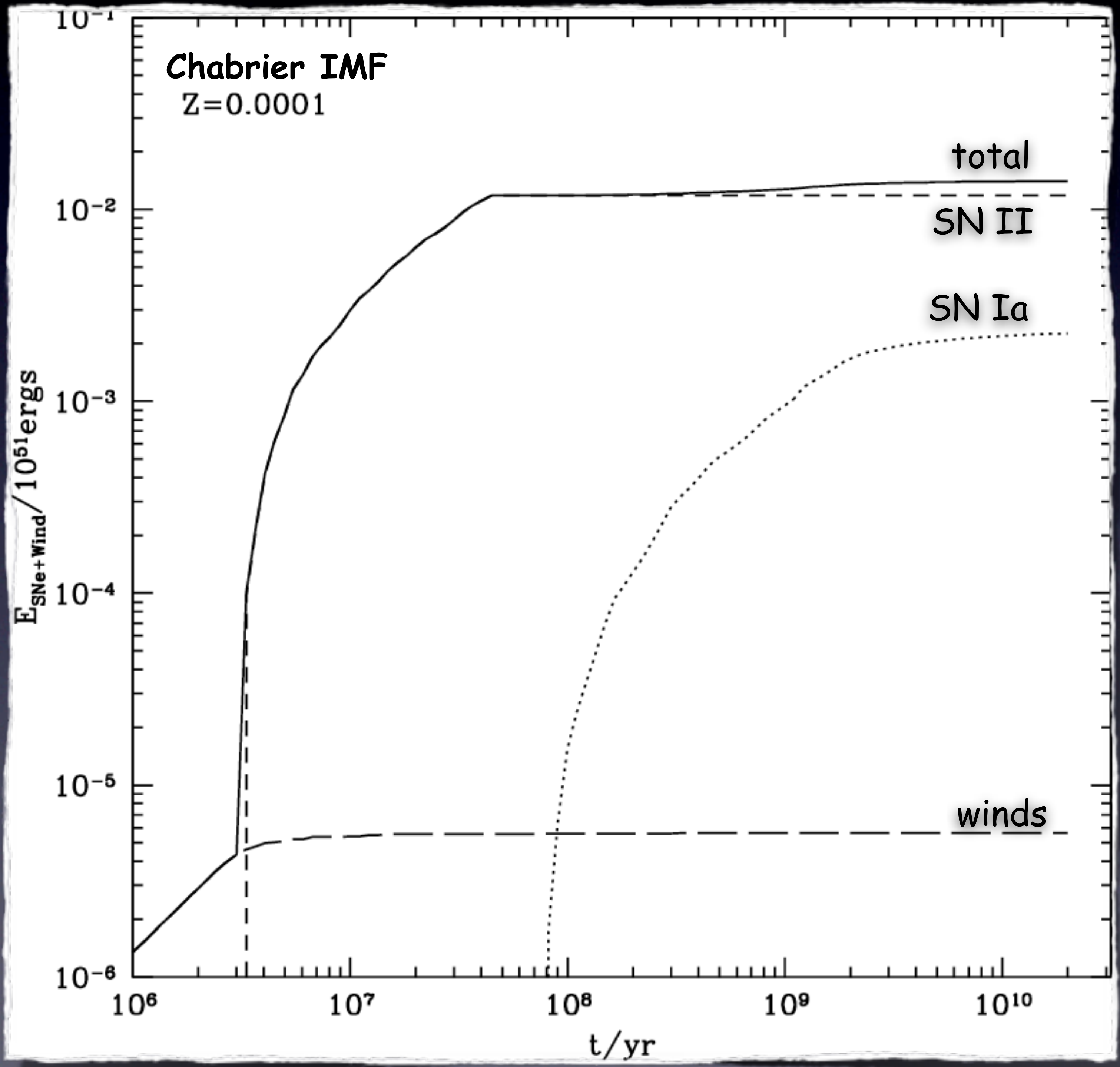
$$\dot{M}_* = 2\pi \int \dot{\Sigma}_*(R)\, R\, {\rm d}R \qquad \dot{\Sigma}_*(R) = \begin{cases} \varepsilon_{\rm SF} \Sigma_{\rm cold}^{1.4}(R) & \text{if } \Sigma_{\rm cold} > \Sigma_{\rm crit} \\ 0 & \text{otherwise} \end{cases}$$

Hence, the SFR follows the Kennicutt-Schmidt law (with `free' parameter for normalization), and takes account of a critical surface density for star formation in agreement with empirical findings (e.g., Kennicutt 1998)

Some Feedback Please...

Stellar Feedback

Stellar evolution injects energy & momentum back into ICM, mainly in form of stellar winds, supernovae and radiation pressure. For a Chabrier IMF, the effective number of SN ($E_{inject}/10^{51}$ erg) per solar mass of stars formed is ~0.01 and dominated by SNII



Effective number of SN per solar mass of stars formed

Source: Benson, 2010, Phys. Rep. 495, 33

SuperNova Feedback (ejection)

To get a feel for whether the energy input from SN can be relevant for galaxy formation, imagine ejecting a mass $M_{\rm ej}$ from the center of a NFW dark matter halo.

This requires an energy injection of $E_{\rm ej} = \frac{1}{2}\, M_{\rm ej}\, V_{\rm esc}^2$. Using that, to a good approximation, the escape velocity from the center of a NFW halo is $V_{\rm esc} \simeq \sqrt{2\,c}\,V_{\rm vir}$ where c is the halo concentration parameter, we have that $E_{\rm ej} \simeq c\, M_{\rm ej}\, V_{\rm vir}^2$

The energy available from SN is $\boxed{E_{\rm fb} = \varepsilon_{\rm SN}\, \zeta\, M_*\, E_{\rm SN}}$

$\varepsilon_{\rm SN} \leq 1$ = fraction of SN energy available for feedback (not just radiated away)

$\zeta \simeq 0.01\, M_\odot^{-1}$ = number of SN produced per Solar mass of stars formed (IMF dependent)

$E_{\rm SN} \simeq 10^{51}\, {\rm erg}$ = energy supplied per SN

Equating $E_{\rm fb}$ to $E_{\rm ej}$ we obtain that

$$\boxed{\frac{M_{\rm ej}}{M_*} \simeq \varepsilon_{\rm SN} \left(\frac{c}{12}\right)^{-1} \left(\frac{V_{\rm vir}}{200\,{\rm km/s}}\right)^{-2}}$$

Hence, even if 100% of SN energy can be converted into kinetic energy of a galactic wind, SN can only eject about the equivalent of the stellar mass from a MW-sized halo.

This efficiency increases with decreasing halo mass; for $V_{\rm vir} = 50\,{\rm km/s}$ we have that $M_{\rm ej} \leq 10\, M_*$.

SuperNova Feedback (reheating)

Rather than ejecting gas from the halo, SN energy can also be used to reheat gas. The internal energy of gas is $E_{\rm int} = \frac{3}{2} M_{\rm gas} \frac{k_{\rm B}\, T}{\mu m_{\rm p}}$.

Imagine we want to reheat this gas from in initial temperature of $T_{\rm init} = 10^4\, K$ to the virial temperature of the halo, $T_{\rm vir} = \frac{\mu m_{\rm p}}{2\, k_{\rm B}}\, V_{\rm vir}^2$

This requires an energy

$$\boxed{E_{\rm reheat} = \frac{3}{2} M_{\rm gas} \frac{k_{\rm B}\,(T_{\rm vir} - T_{\rm init})}{\mu m_{\rm p}} = \frac{3}{4} M_{\rm gas} V_{\rm vir}^2 \left(1 - \frac{T_{\rm init}}{T_{\rm vir}}\right)}$$

Equating $E_{\rm reheat}$ to $E_{\rm ej}$ yields

$$\boxed{\frac{M_{\rm gas}}{M_*} \simeq 17\, \varepsilon_{\rm SN} \left(\frac{V_{\rm vir}}{200\,{\rm km/s}}\right)^{-2} \left(1 - \frac{T_{\rm init}}{T_{\rm vir}}\right)^{-1}}$$

Hence, in a MW halo, SN can reheat up to $17\, M_\odot$ for every Solar mass of stars formed.

Reheating is more efficient than ejecting gas, by roughly factor $(V_{\rm esc}/V_{\rm vir})^2 \simeq 2\, c$

The all important question for gauging the potential impact of SN feedback is

what is the SN feedback efficiency parameter $\varepsilon_{\rm SN}$

Depending on the ISM and SF conditions, one can have $0.01 < \varepsilon_{\rm SN} \leq 1$

AGN Feedback

What about feedback from AGN?

The energy output from an AGN (over its lifetime) is

$$\boxed{E_{\rm AGN} \sim 0.1\, M_{\rm BH}\, c^2}$$

where we have assumed that roughly 10% of the rest-mass energy is radiated away. If we assume that a fraction $\varepsilon_{\rm AGN}$ of this radiation is used to reheat gas or to eject it from the halo, and we use that $M_{\rm BH} \simeq 0.002\, M_{\rm bulge}$ we obtain that

$$\boxed{\frac{E_{\rm AGN}}{E_{\rm SN}} \sim 36\, \frac{\varepsilon_{\rm AGN}}{\varepsilon_{\rm SN}} \left(\frac{M_{\rm bulge}}{M_*}\right)}$$

If $\varepsilon_{\rm AGN} \sim \varepsilon_{\rm SN}$ AGN feedback can be order of magnitude more efficient than SN feedback

Key question: what is a realistic value for $\varepsilon_{\rm AGN}$

NOTE: fact that we can see AGN implies that $\varepsilon_{\rm AGN}$ has to be significantly smaller than unity

SN Feedback in SAMs

- SAMs use the SN energy to either eject or reheat cold gas from the disk.
- Fraction of SN energy used for ejection vs. reheating varies from SAM to SAM.
- Ejected gas is often stored in \`reservoir' for later re-accretion into the halo.

Majority of SAMs adopt

$$\dot{M}_{\rm fb} \equiv \dot{M}_{\rm eject} + \dot{M}_{\rm reheat} = \varepsilon_{\rm fb}\,\dot{M}_* \quad \text{with} \quad \varepsilon_{\rm fb} = \varepsilon_{\rm fb,0}\left(\frac{V_{\rm vir}}{200\,{\rm km/s}}\right)^{-\alpha_{\rm fb}}$$

Energy driven winds have $\alpha_{\rm fb} = 2$ (e.g., Dekel & Silk 1986)

Momentum driven winds have $\alpha_{\rm fb} = 1$ (e.g., Murray et al. 2005)

SAMs typically \`require' $\alpha_{\rm fb} > 2$ in order to match faint-end slope of LF.

Extreme example: $\alpha_{\rm fb} = 5.5$ (Cole et al. 1994)

The recent \`Munich' SAM of Guo et al. (2011) adopt $\alpha_{\rm fb} = 3.5$, but with a modified functional form:

for reheating

$$\varepsilon_{\rm fb} = \varepsilon_{\rm fb,0}\left[0.5 + \left(\frac{V_{\rm vir}}{70\,{\rm km/s}}\right)^{-3.5}\right]$$

\`arbitrary' characteristic scale...

AGN Feedback in SAMs

- AGN feedback is `required' in order to quench star formation in massive haloes
- Two different modes of AGN feedback are included: Quasar mode & Radio mode
- AGN feedback has become a `standard ingredient' in SAMs since 2006

"Quasar Mode" associated with mergers or disk instabilities when large amounts of gas are funneled to galaxy's center. Main channel of BH growth

$$\boxed{\dot{M}_{\rm BH} = M_{\rm BH,min} + f(M_{\rm cold,1}, M_{\rm cold,2}, M_{\rm vir,1}, M_{\rm vir,2})}$$

Typically the function f is tuned to match the empirical $M_{\rm BH}$-$M_{\rm bulge}$ relation.

Treatment of feedback from Quasar Mode varies strongly from SAM to SAM

'Santa Cruz' SAMS: momentum driven outflow $\dot{M}_{\rm out} = \varepsilon_{\rm quasar}\,\dot{M}_{\rm BH}\,c/V_{\rm esc}$

'Munich' SAMS: no explicit treatment of quasar mode feedback,
but starburst associated with merger typically
ejects all cold gas via `regular' SN feedback

AGN Feedback in SAMs

- AGN feedback is `required' in order to quench star formation in massive haloes
- Two different modes of AGN feedback are included: Quasar mode & Radio mode
- AGN feedback has become a `standard ingredient' in SAMs since 2006

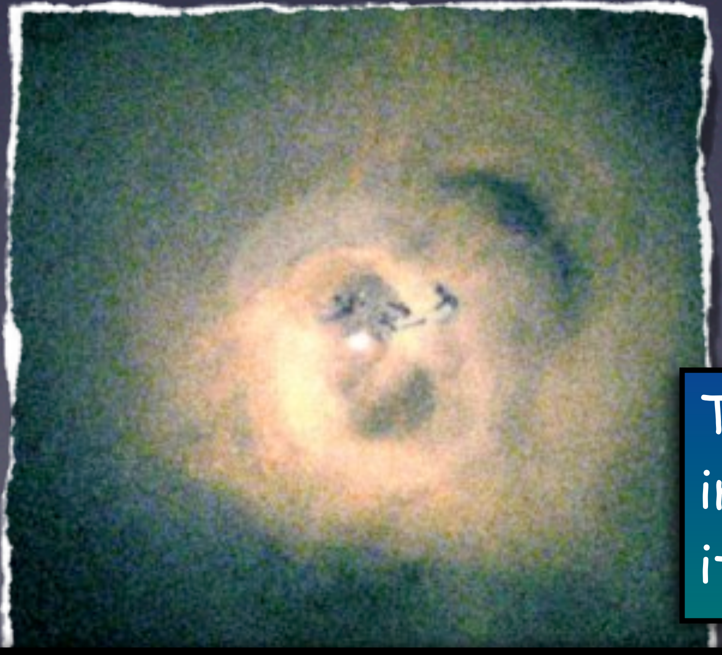
"Radio Mode"

associated with low-Eddington Bondi accretion of hot gas in massive haloes. Little BH growth, but main mode to offset cooling

$$\dot{M}_{\rm BH} = \varepsilon_{\rm radio} \left(\frac{M_{\rm BH}}{10^8\, M_\odot} \right) f(M_{\rm hot}, T_{\rm hot})$$

Motivation is that low-Eddington ratio accretion results in powerful jets. Most SAMs assume that this energy output heats the hot gas, thereby `quenching' star formation

$$\dot{M}_{\rm cool,eff} = \dot{M}_{\rm cool} - \frac{0.1\, \dot{\rm M}_{\rm BH}\, c^2}{V_{\rm vir}^2}$$



There is ample observational evidence to suggest that radio-mode AGN impact the ICM of clusters, such as here in the Perseus cluster. However, it is unclear whether the AGN energy is actually thermalized...

Dynamical Friction, Mergers & Morphological Transformations

Galaxy Mergers in SAMs

- When two haloes merge, most massive galaxy becomes central galaxy in new halo.
- All other galaxies become satellites that orbit central.
- Satellites experience dynamical friction, causing them to “sink” to center of halo.
- Satellites will merge with central after dynamical friction time, $t_{\rm df}$ (cannibalism)
- Some SAMs also allow for mergers between satellite galaxies...

When two galaxies merge:

$M_1/M_2 \leq \mathcal{R}_{\rm maj}$ → major merger → disks are transformed into spheroid

$M_1/M_2 > \mathcal{R}_{\rm maj}$ → minor merger → disk masses added together

$\mathcal{R}_{\rm maj}$ is free parameter, tuned to match observed morphologies (typically $\mathcal{R}_{\rm maj} \sim 3-4$)

Dynamical friction time is based off Chandrasekhar’s derivation

$$\boxed{t_{\rm df} = \varepsilon_{\rm df}\, t_{\rm dyn} \left(\frac{M_{\rm vir}}{M_{\rm sat}}\right) \frac{f_{\rm orbit}}{\ln \Lambda}}$$

free fudge parameter

`Coulomb’ logarithm

function describing orbital dependence

Numerical simulations and analytical calculations suggest $\varepsilon_{\rm df} \sim 3$ (see MBW §12.3)

Starbursts

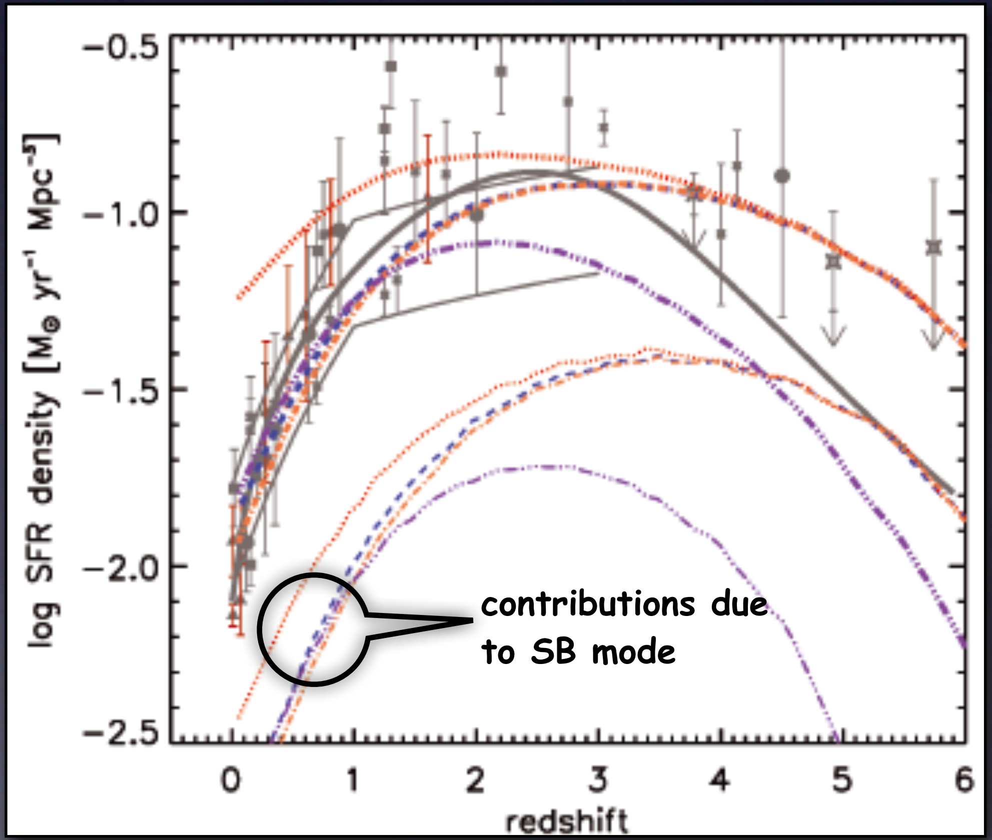
Most SAMs assume that during (major) mergers a (large) fraction of the cold gas is consumed in a burst of star formation:

$$\dot{M}_{*,\mathrm{burst}} = \frac{e_{\mathrm{SB}}\, M_{\mathrm{cold}}}{\tau_{\mathrm{SB}}} \qquad e_{\mathrm{SB}} = \varepsilon_{\mathrm{SB}} \left(\frac{M_{\mathrm{min}}}{M_{\mathrm{maj}}}\right)^{\gamma_{\mathrm{SB}}}$$

'Santa Cruz' SAMS: $(\varepsilon_{\mathrm{SB}}, \gamma_{\mathrm{SB}}, \tau_{\mathrm{SB}})$ calibrated against hydro sims (Cox et al. 2008)

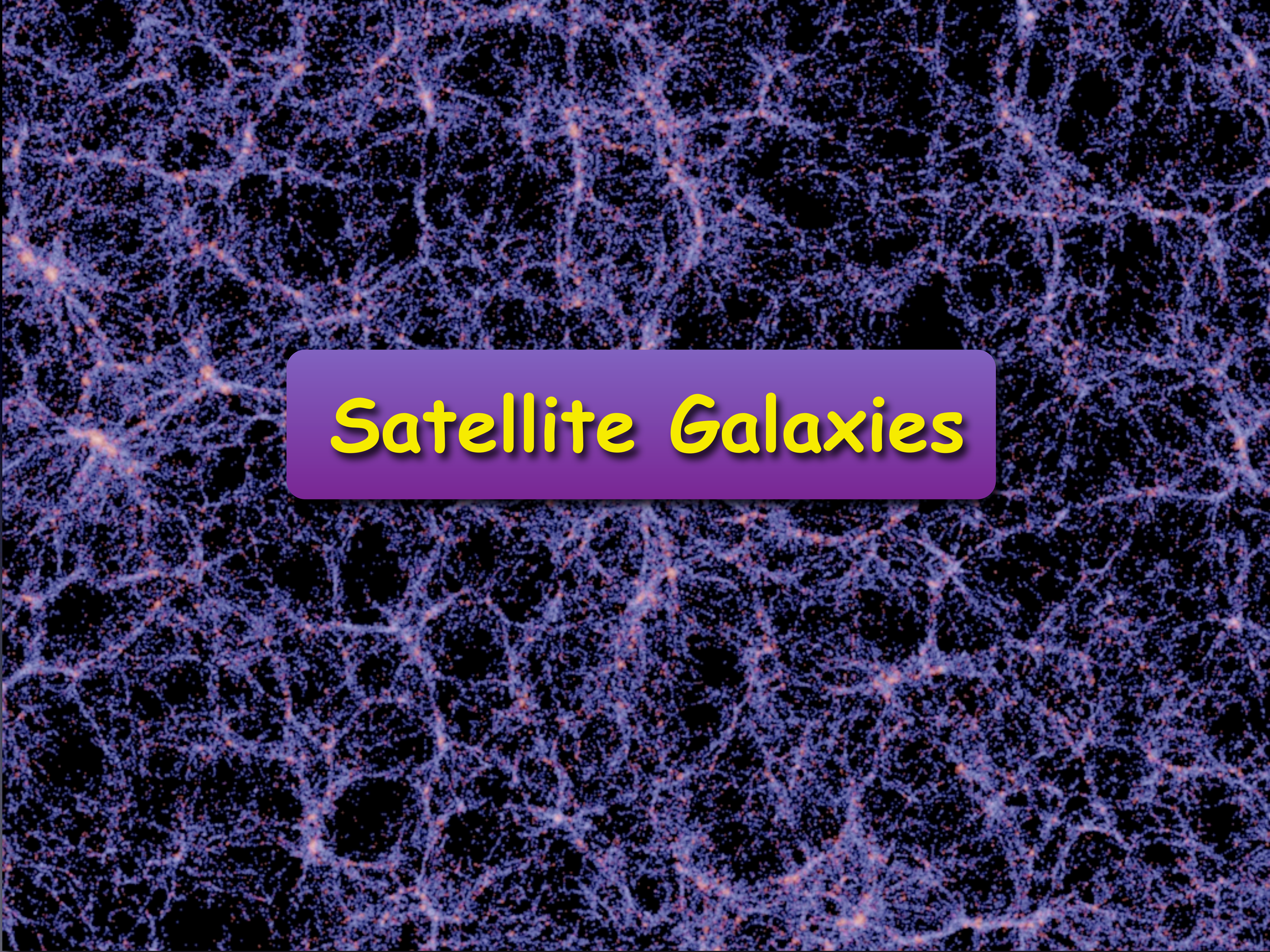
'Munich' SAMS: $\varepsilon_{\mathrm{SB}} = 0.56,\ \gamma_{\mathrm{SB}} = 0.7$ and $\tau_{\mathrm{SB}} = 0$ (major+minor mergers)

'Durham' SAMS: $\varepsilon_{\mathrm{SB}} = 1,\ \gamma_{\mathrm{SB}} = 0$ and $\tau_{\mathrm{SB}} = t_{\mathrm{ff,sph}}$ (major mergers only)



Source: Somerville et al. 2008, MNRAS, 391, 481

Cosmic star formation histories in the SAMs of Somerville et al. (2008). Different colors correspond to different models. The thin lines show the contributions due to star bursts, which typically are less than 10 percent...



Satellite Galaxies

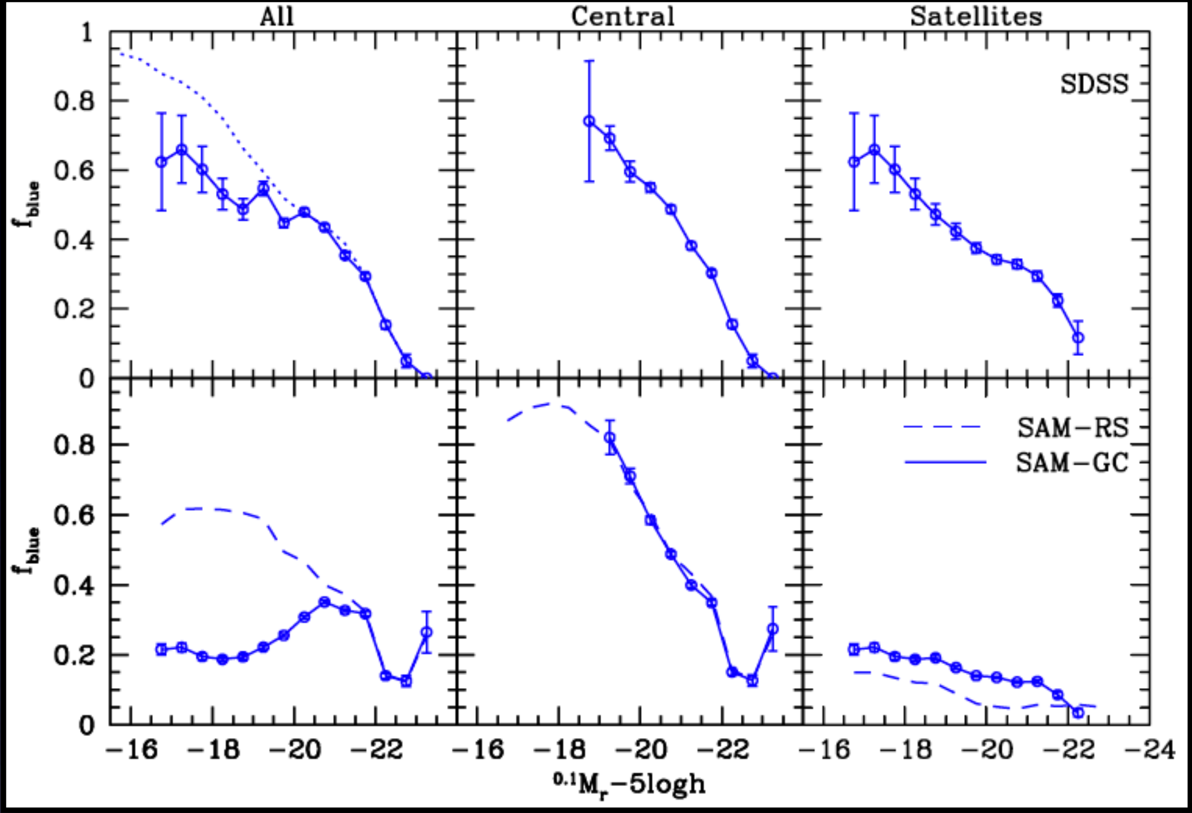
Satellite Quenching

Prior to ~2005 most SAMs instantaneously stripped satellite galaxies of their hot gas reservoir at infall, resulting in rapid quenching of their star formation (strangulation)

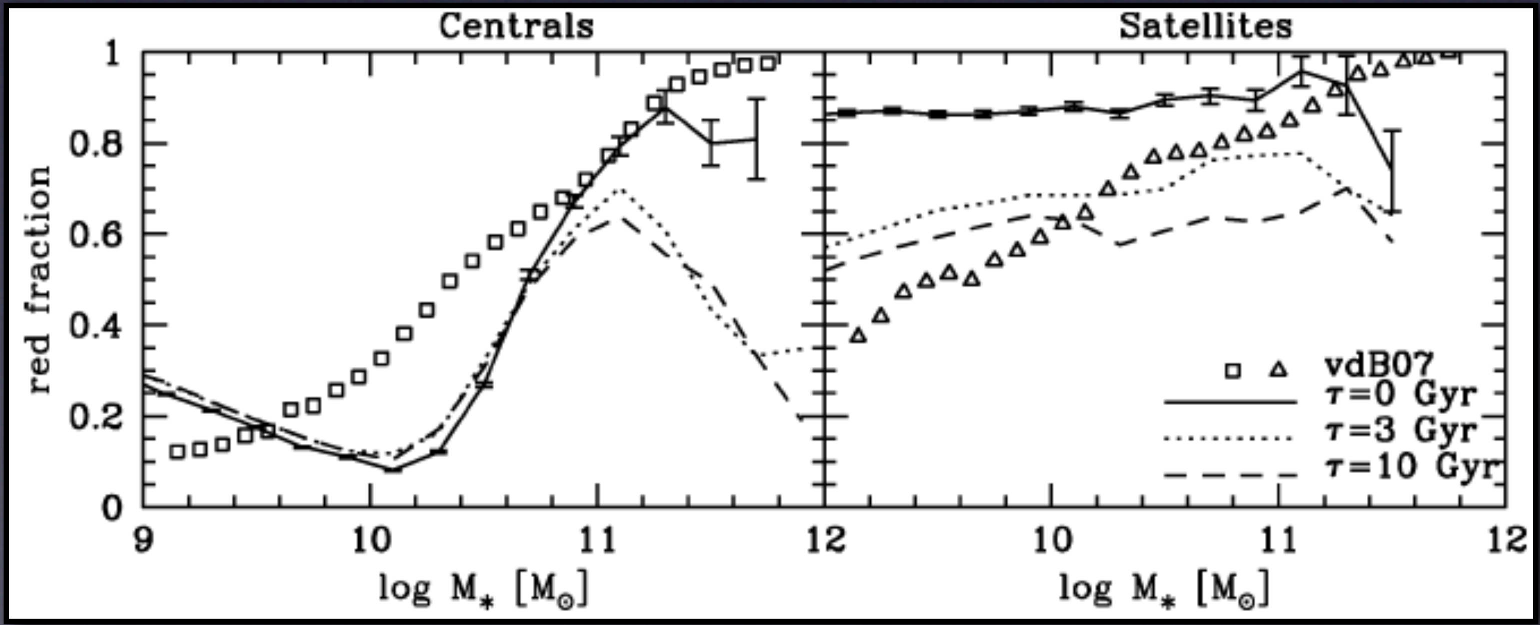
However, as first shown in Weinmann et al. (2006), this results in blue fractions of satellite galaxies that are much too low compared to observations.

Instantaneous strangulation is too rapid.

As shown by Kang & vdB (2008), simply slowing down rate at which hot gas is stripped results in problems with the centrals: they become blue (despite AGN feedback) because they accrete blue satellites..



Weinmann, et al. 2006, MNRAS, 372, 1161



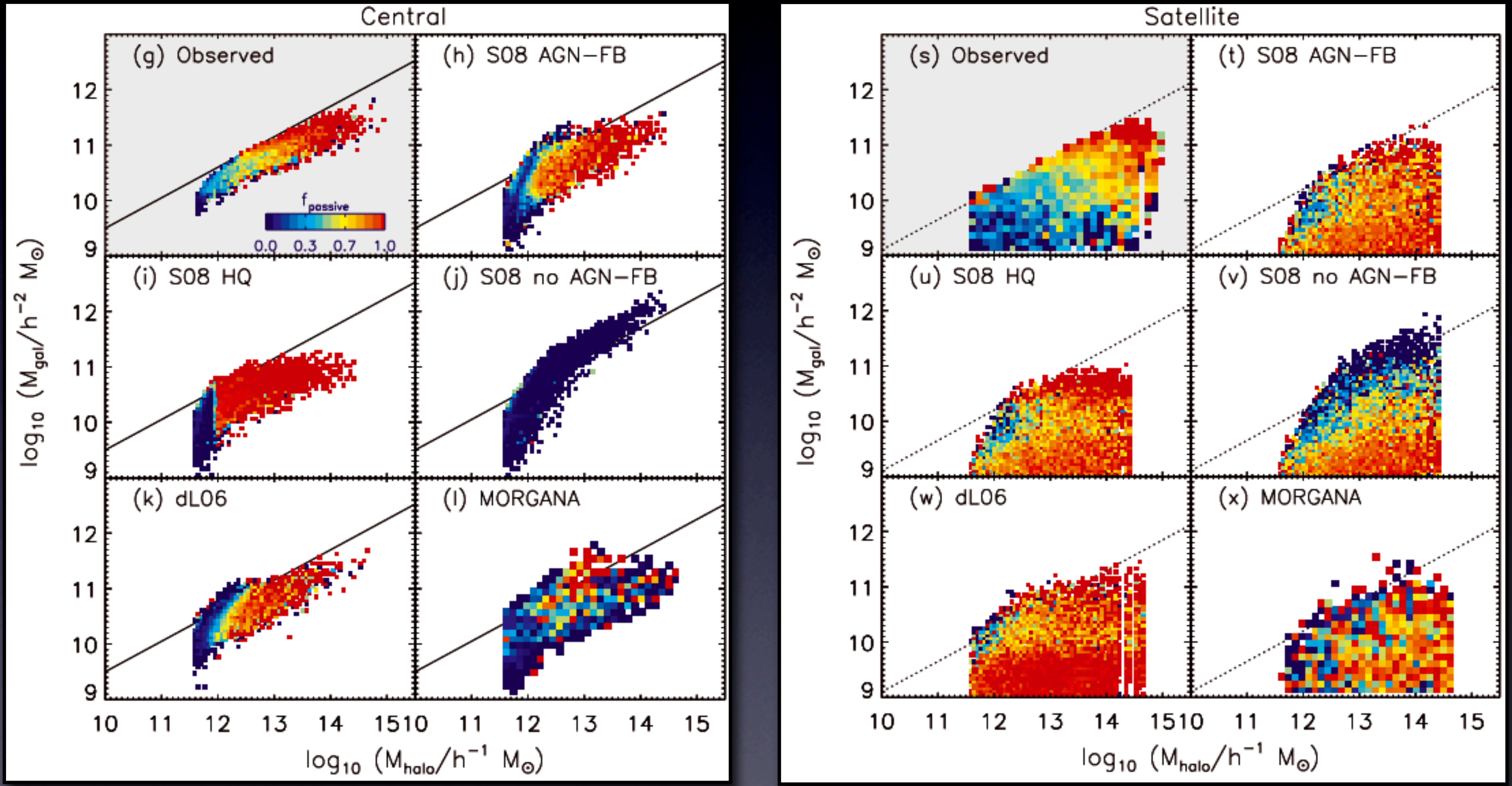
Kang & vdB, 2008, ApJ, 676, 101

Satellite galaxies mainly need to be disrupted, rather than cannibalized

IntraClusterLight & Stellar Halos

Quenching in SAMs

Kimm et al. (2009) performed a detailed comparison of the passive fractions of central and satellite galaxies in SDSS (based on galaxy group catalog of Yang et al. 2005) with those in different SAMs reveals severe problems for the latter....

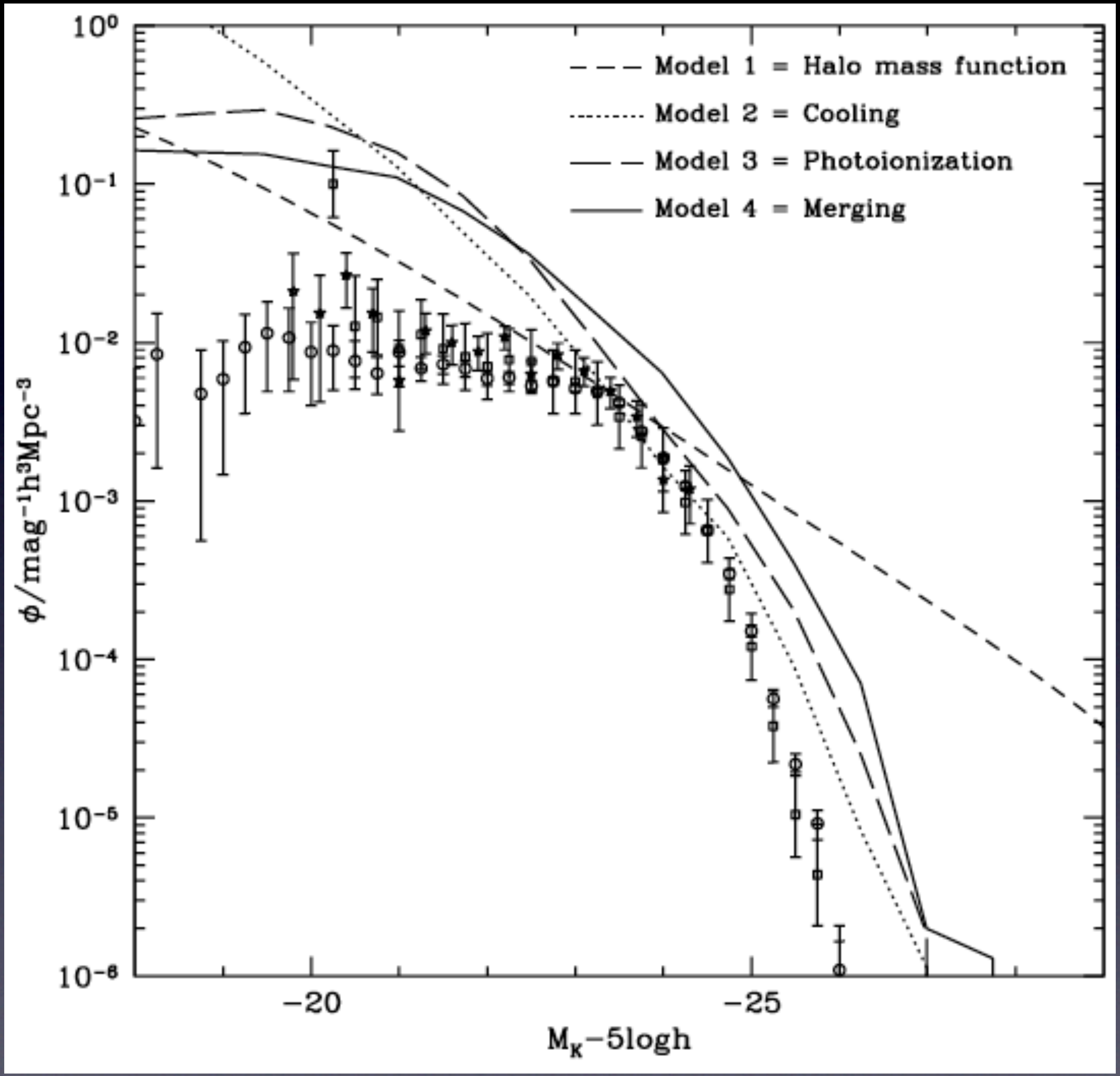


Source: Kimm et al. 2009, MNRAS, 394, 1131

Quenching is not yet treated properly in SAMs

Galaxy Abundances

Fitting the Galaxy Luminosity Function



Source: Benson et al. 2003, ApJ, 599, 38

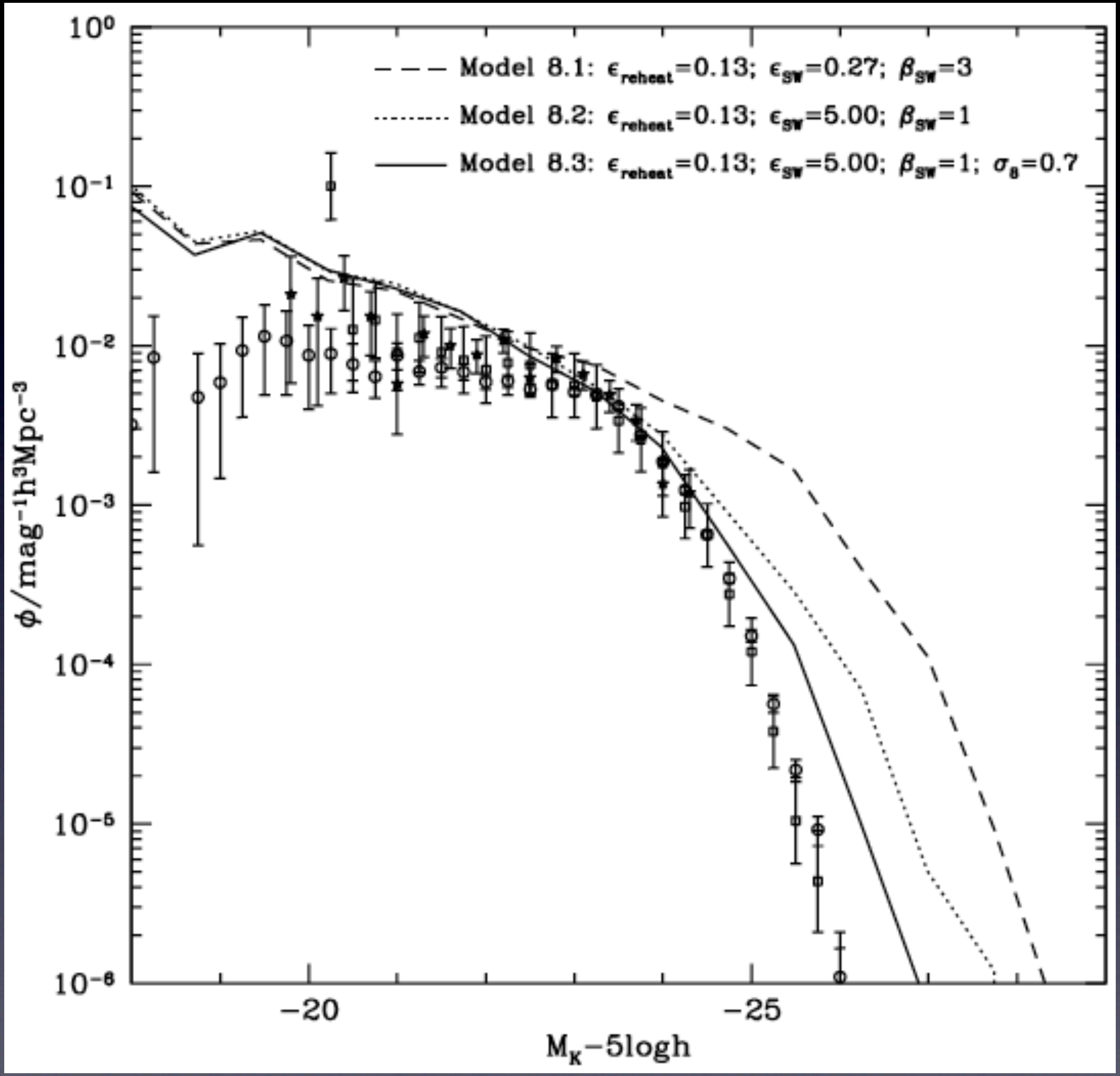
If galaxy formation would be equally efficient in halos of all masses, the shape of the LF would be very different from what is observed.

Since cooling is less efficient in more massive haloes, this was originally believed to explain the bright end of the LF.

However, hierarchical formation results in too much cooling at early times...

Photoionization by UV background is not sufficient to solve overcooling problem

Fitting the Galaxy Luminosity Function



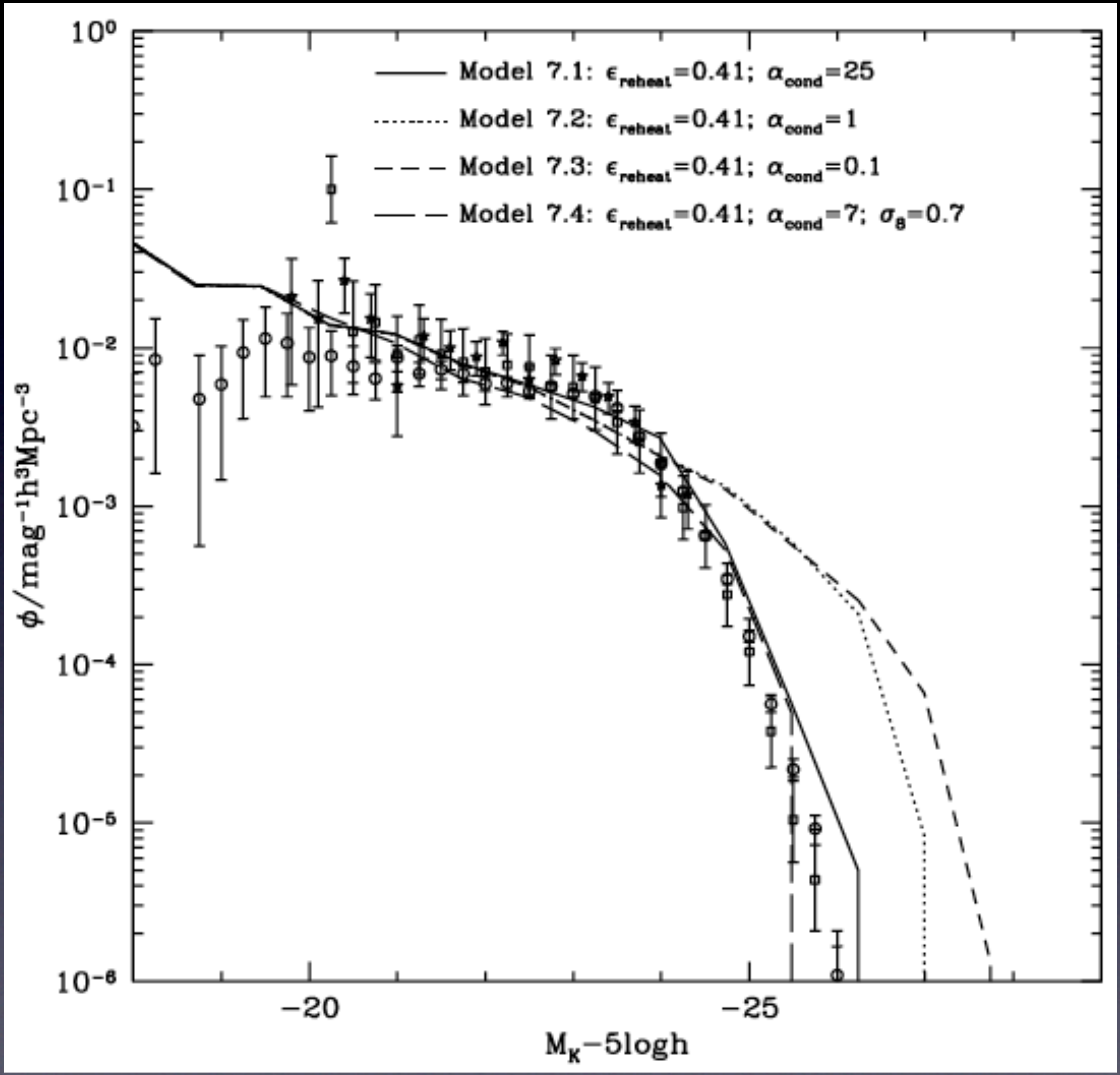
Source: Benson et al. 2003, ApJ, 599, 38

Adding SN feedback helps in suppressing galaxy formation in low mass haloes, resulting in a reduced abundance of faint galaxies...

Matching the bright-end of the LF requires at least 5x as much energy as what is available from stellar evolution...

SN feedback by itself cannot solve the overcooling problem

Fitting the Galaxy Luminosity Function



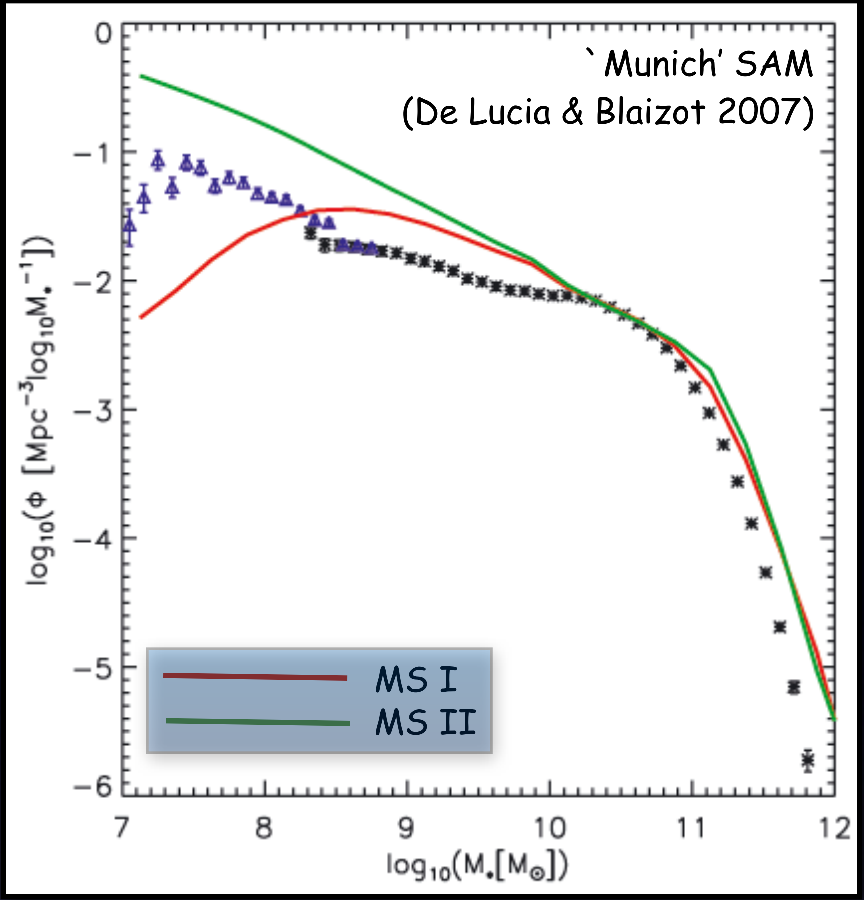
Source: Benson et al. 2003, ApJ, 599, 38

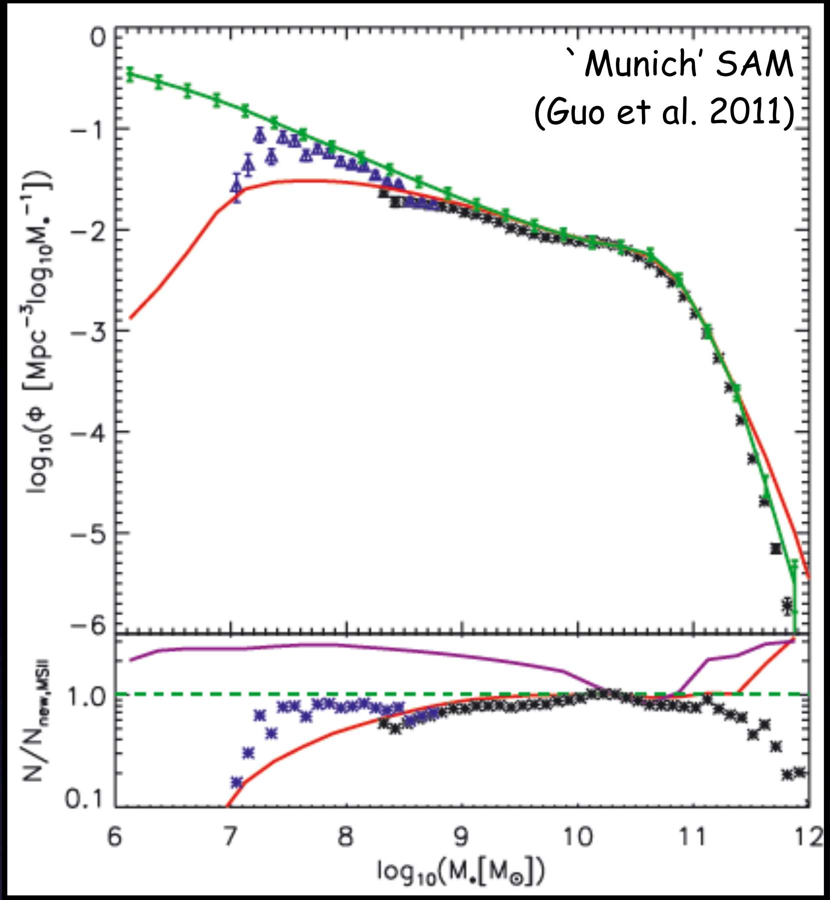
Trying to solve the problems at the bright end of the LF with conductivity requires efficiencies that are 25x the maximum allowed Spitzer conductivity.

In presence of magnetic fields, conductivity is expected to be significantly less than Spitzer...

Conductivity is not a solution to the overcooling problem

The Current State-of-Affair

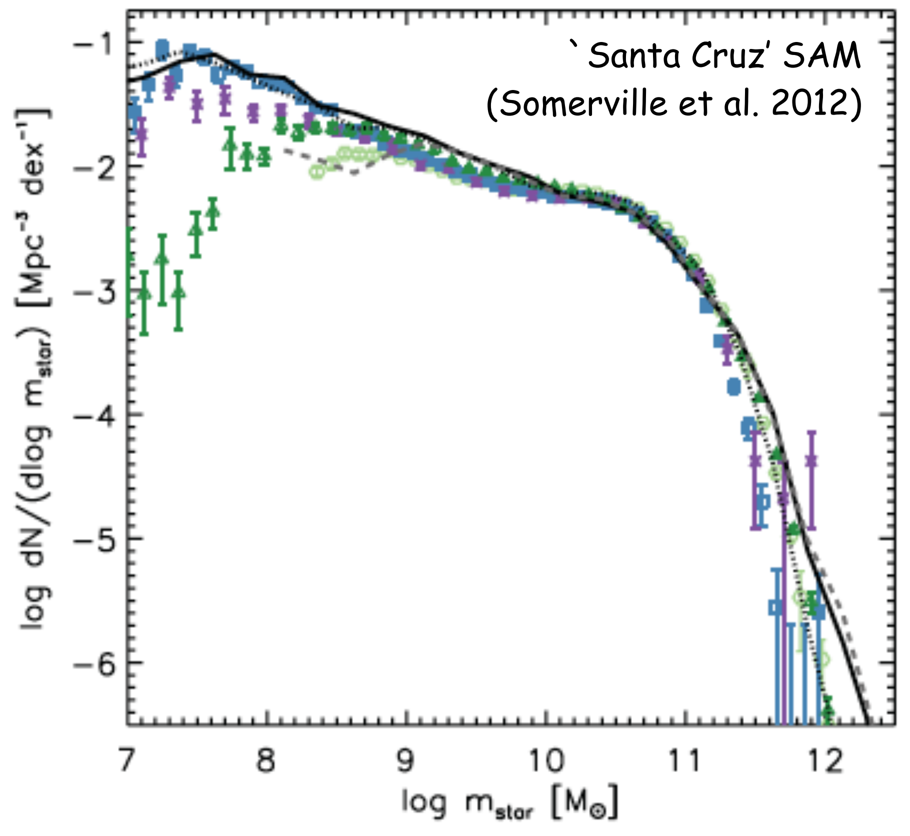




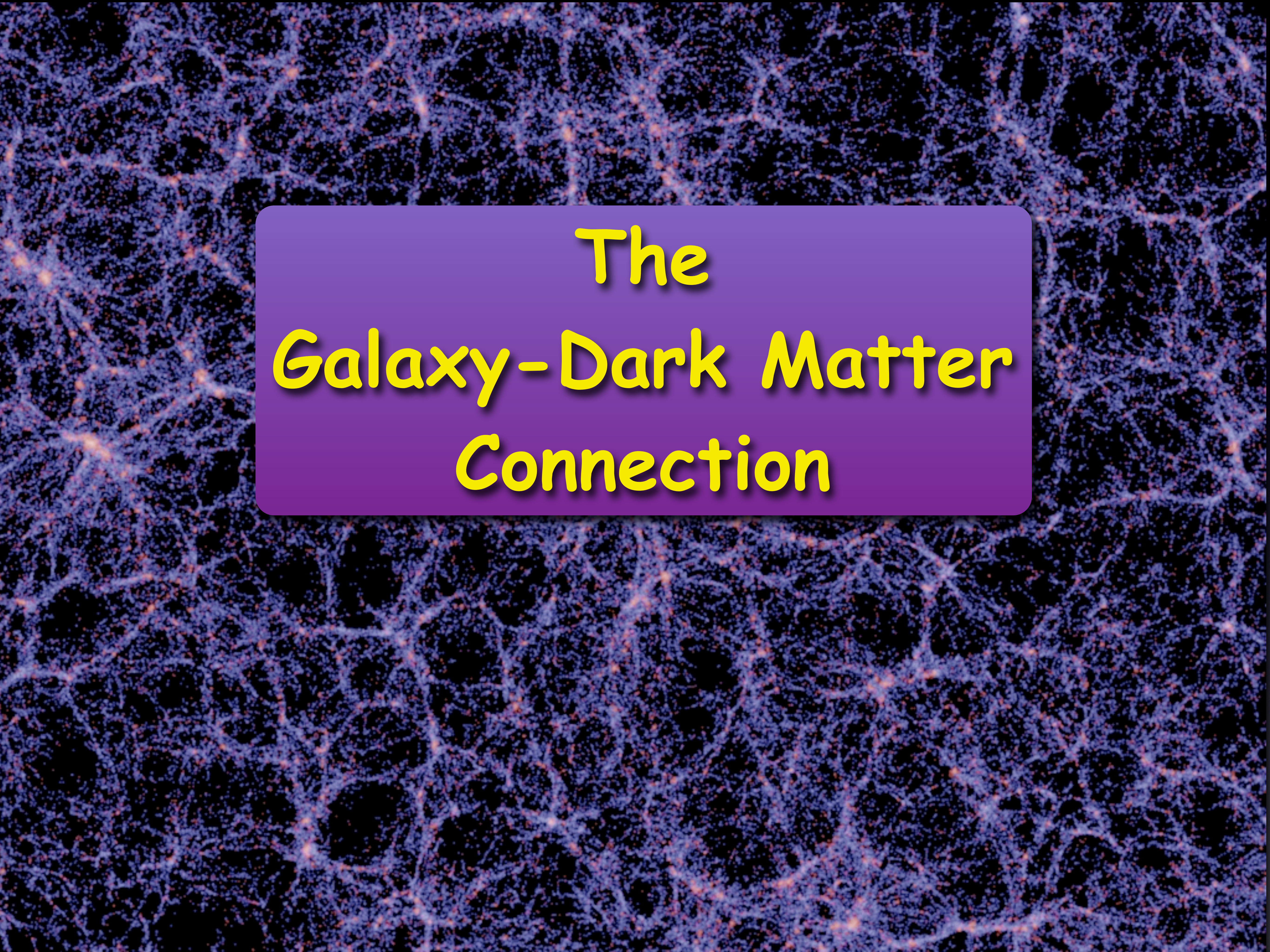
Main difference is in treatment of SN feedback; the Guo et al. model has introduced additional degrees of freedom.

This results in a much improved fit of stellar mass function at low-mass end.

With AGN feedback (radio mode) SAMs nowadays have little difficulty matching high-mass end of stellar mass function.



The latest `Santa Cruz' SAM of Somerville et al. (2012) also matches low-mass end of stellar mass function, but using virtually same SN feedback model as De Lucia & Blaizot (2007). They do include Quasar Mode Feedback though....

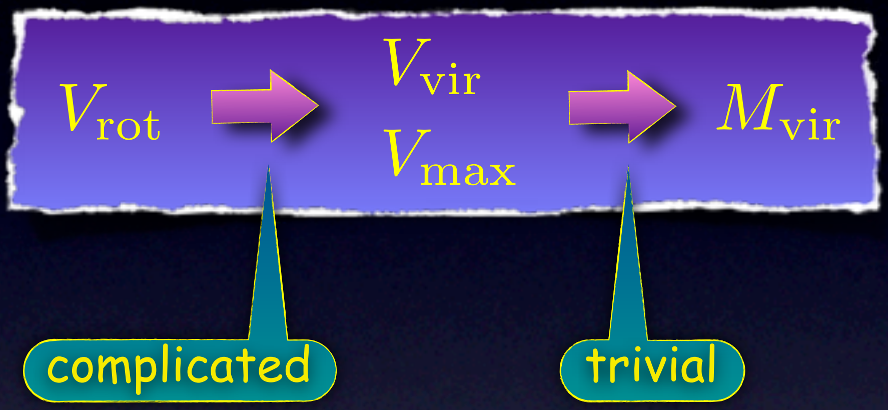


The Galaxy-Dark Matter Connection

The Tully-Fisher Zero-Point

Reproducing the galaxy luminosity or stellar mass function(s) does not yet guarantee that these galaxies reside in the correct dark matter haloes.

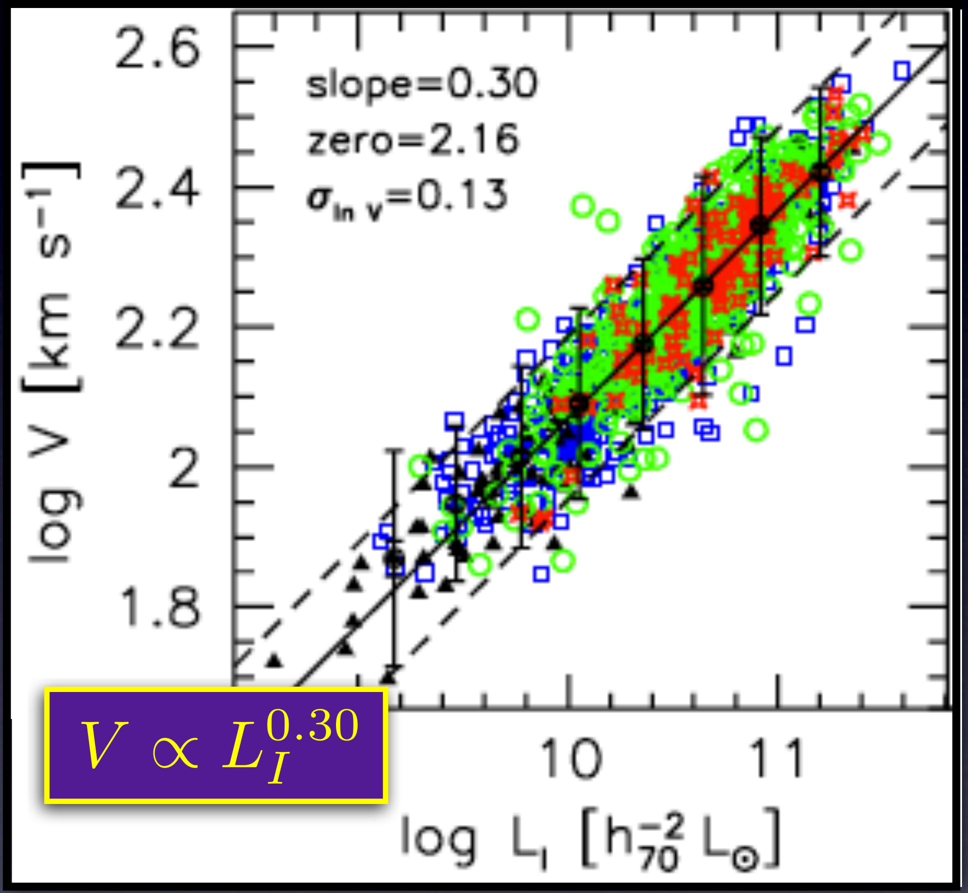
Historically, the main \`test' that was used to check this was the Tully-Fisher relation.



$$V_{\rm rot} \rightarrow \begin{matrix} V_{\rm vir} \\ V_{\rm max} \end{matrix} \rightarrow M_{\rm vir}$$

complicated

trivial



$V \propto L_I^{0.30}$

Source: Courteau et al. 2007, ApJ, 671, 203

Prior to ~2005 SAMS failed to match zero-point of TF relation, overpredicting $V_{\rm rot}$ at fixed L. Only when SAMs started to adopt ΛCDM concordance cosmology did problem \`disappear', but ONLY under the \`naive' assumption that $V_{\rm rot} = V_{\rm max}$ or $V_{\rm rot} = V_{\rm vir}$

That assumption, however, ignores

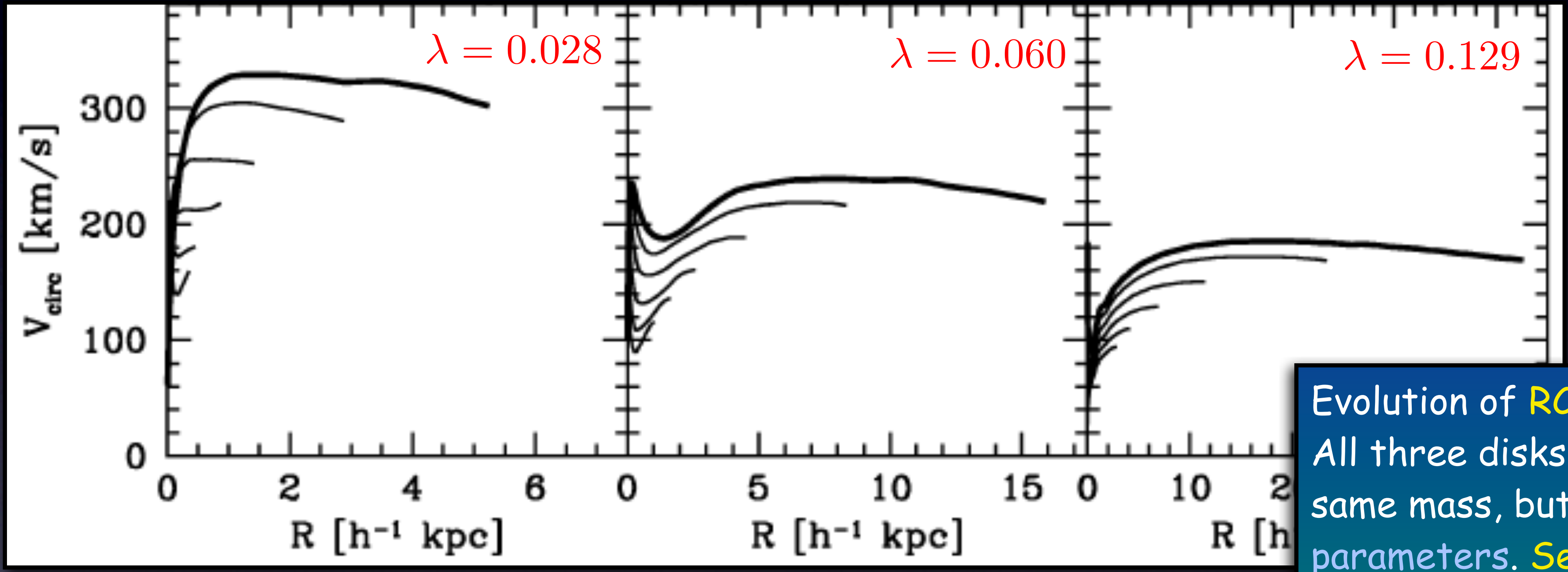
- self-gravity of the disk galaxy
- adiabatic contraction of the dark matter halo

Adiabatic Contraction

Adiabatic Contraction: contraction of the dark matter halo as a gravitational response to baryons accumulating at center of potential well.

Following Blumenthal et al. (1986), adiabatic contraction is typically modeled assuming that r M(r) is an adiabatic invariant during the formation of a disk galaxy.

van den Bosch, 2002, MNRAS, 332, 456



Evolution of RCs in disk formation models. All three disks reside in haloes of the same mass, but with different spin parameters. Self-gravity of disk and AC are both taken into account

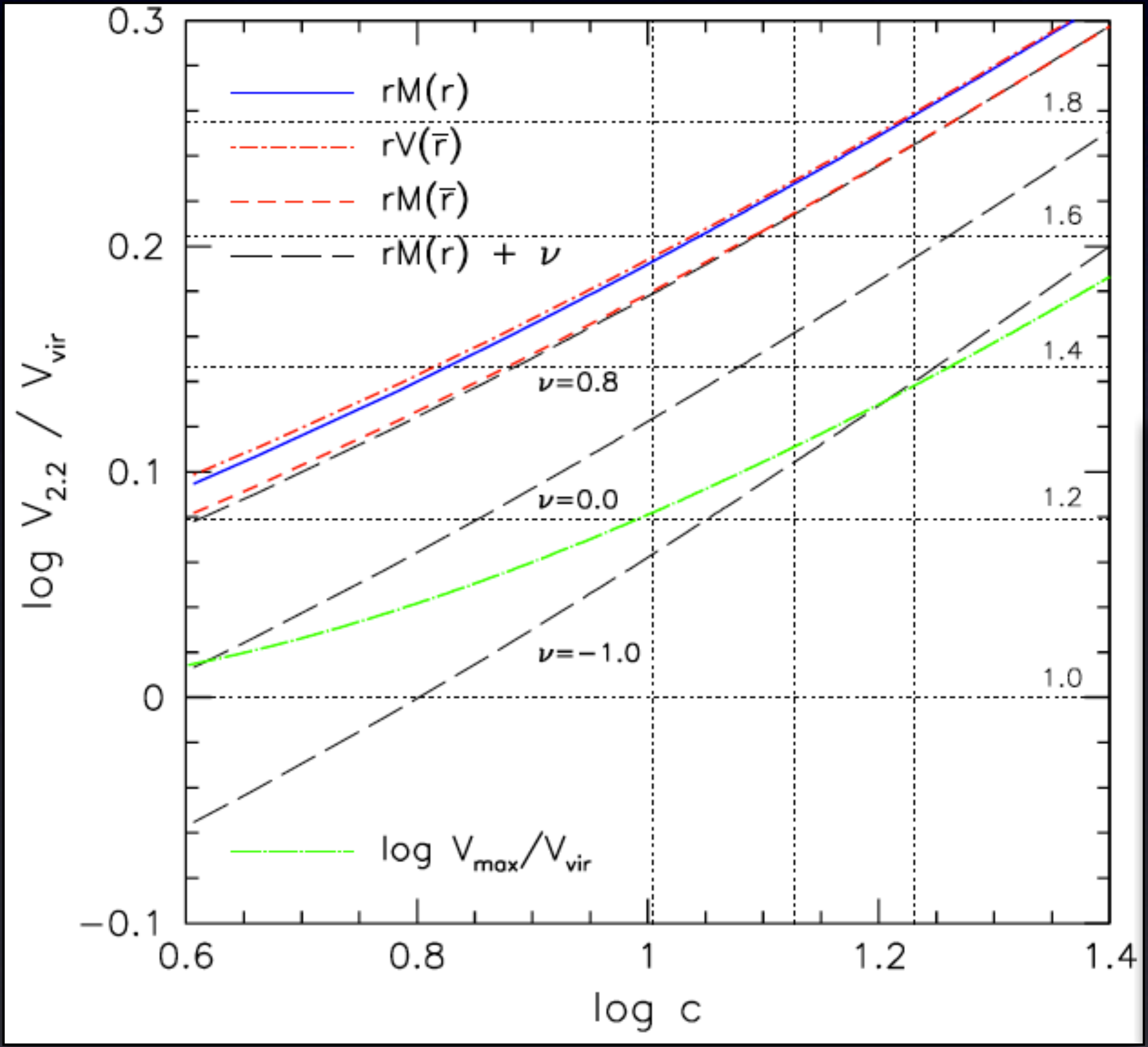
Realistic, detailed models of disk formation that include self-gravity of disk and adiabatic contraction of halo typically predict that $V_{rot}/V_{vir} \sim 1.4\text{-}1.8$, whereas NFW haloes typically have $V_{max}/V_{vir} \sim 1.2$...

Adiabatic Contraction

Note: r M(r) is only an adiabatic invariant under oversimplified assumptions. Under realistic conditions, halo contraction can be modelled using $r_{\rm f} = \Gamma^{\nu} r_{\rm i}$, where $\Gamma = r_{\rm f}/r_{\rm i}$ is the simplified AC model of Blumenthal et al. (1986) and ν is a `free parameter'.

Simulations suggest that $\nu \simeq 0.8$

$\nu = 1$ standard AC
$\nu = 0$ no adiabatic contraction
$\nu < 0$ halo expansion

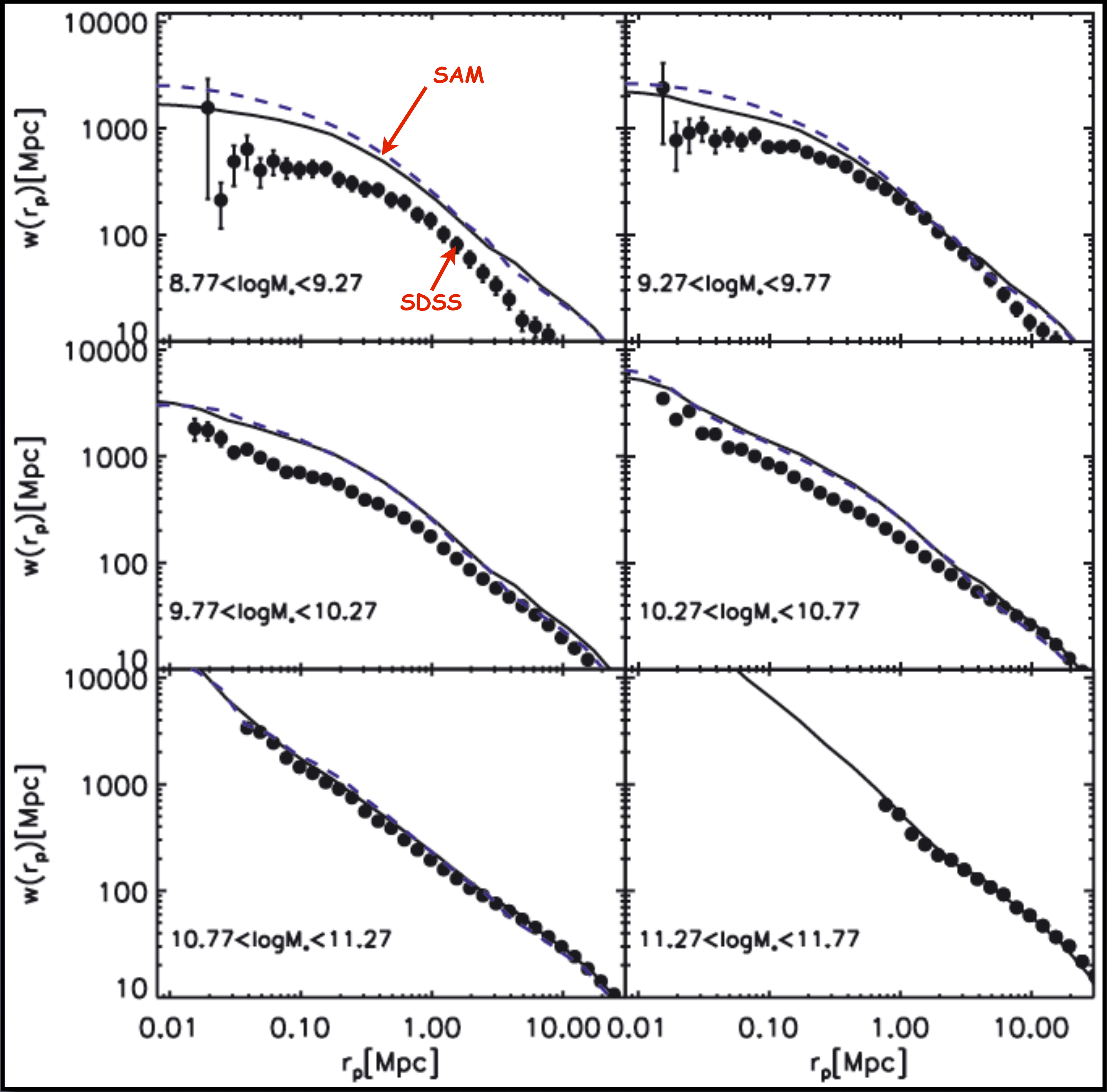


Dutton, vdBosch, Dekel & Courteau, 2007, ApJ, 654, 27

Note how even in the case without AC ($\nu = 0$) $V_{2.2} > V_{\rm max}$, which is simply due to the self-gravity of the disk

Model predictions for the ratio $V_{2.2}/V_{\rm vir}$ for disk galaxies embedded in NFW haloes. Results are shown as function of the halo concentration parameter, c. Here $V_{2.2}$ is the rotation velocity of the disk measured at 2.2 disk scale-lengths. Results are shown for different `forms' of adiabatic contraction (different values of ν). For comparison, the green curve shows the ratio $V_{\rm max}/V_{\rm vir}$.

Galaxy Clustering



Source: Guo et al. 2011, MNRAS, 413, 101

CDM predicts that more massive haloes are more strongly clustered. Hence, only if galaxies reside in the correct dark matter haloes will the models match clustering data.

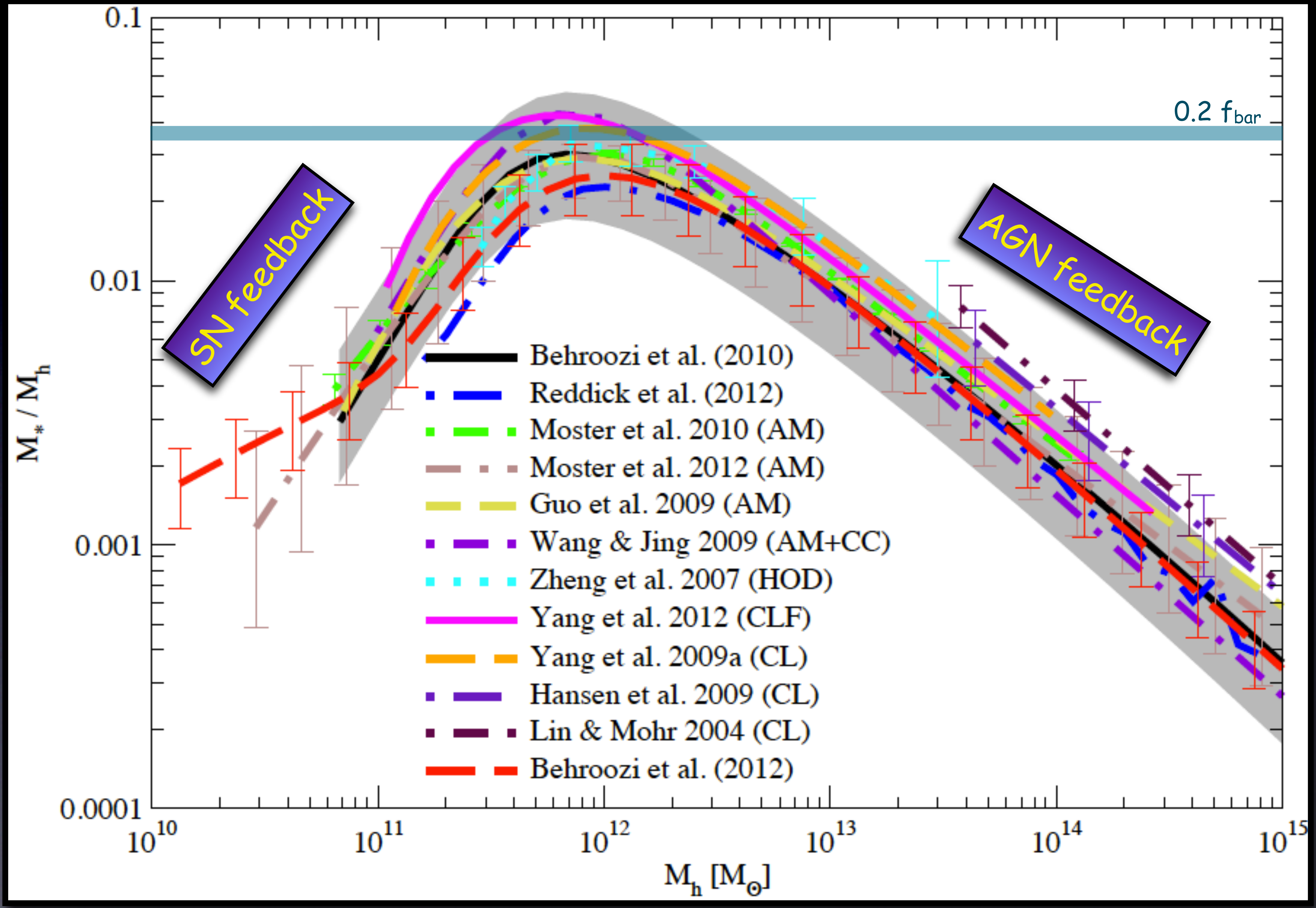
"Munich" SAM of Guo et al. (2011) fails in that respect, despite good match to stellar mass function....

Part of problem is likely to be that Millennium Simulation adopts wrong cosmology....

Several studies have shown that the WMAP1 cosmology cannot simultaneously match galaxy abundances & galaxy clustering.
(e.g., van den Bosch et al 2004; Cacciato et al. 2009, 2012)

Halo Occupation Modeling

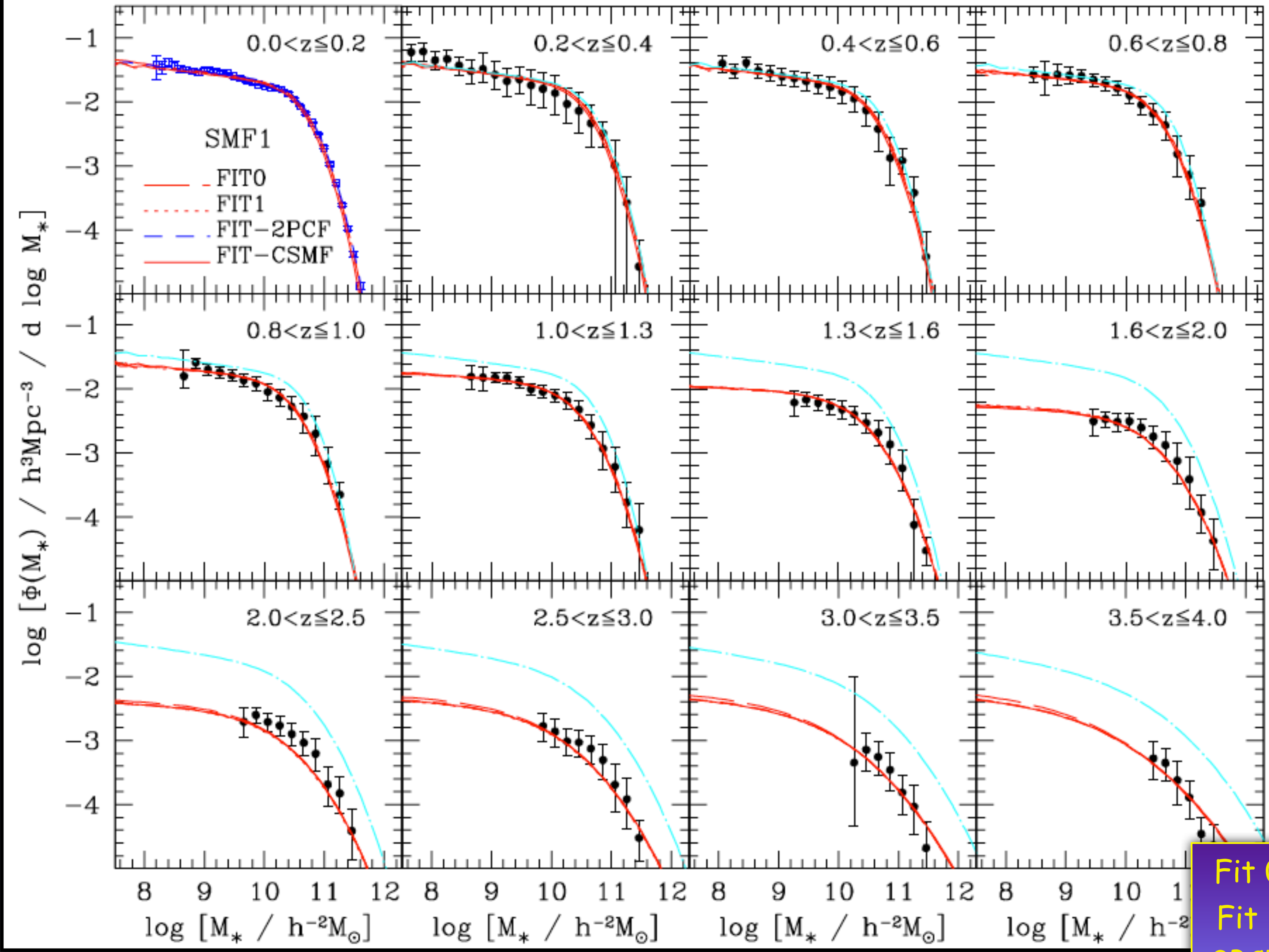
The Galaxy - Dark Matter Connection



Source: Behroozi, Wechsler & Conroy, 2012, arXiv:1207.6105

The Galaxy - Dark Matter Connection

Data: Yang et al (2009; z~0.1)
Perez-Gonzales et al. (2008)



Source: Yang et al. 2012, ApJ, 752, 41

Fit 0 : α=0 --> no sats

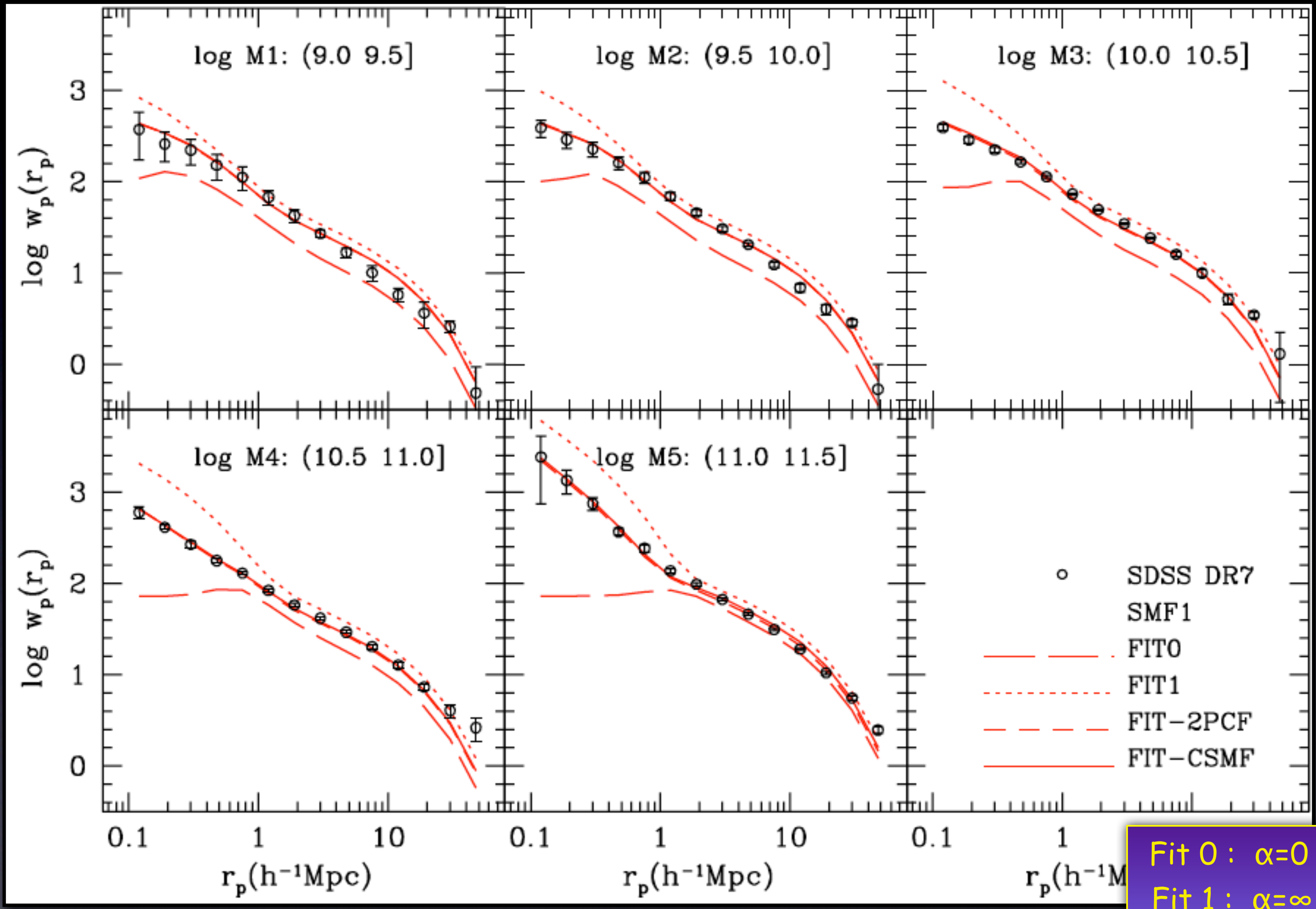
Fit 1 : α=∞ --> no evolution

2PCF : fit to $\Phi(M_\star)$ + 2PCF

CSMF : fit to $\Phi(M_\star) + \Phi(M_\star|M,z=0)$

The Galaxy - Dark Matter Connection

Data: SDSS DR7
(Yang et al. 2012)

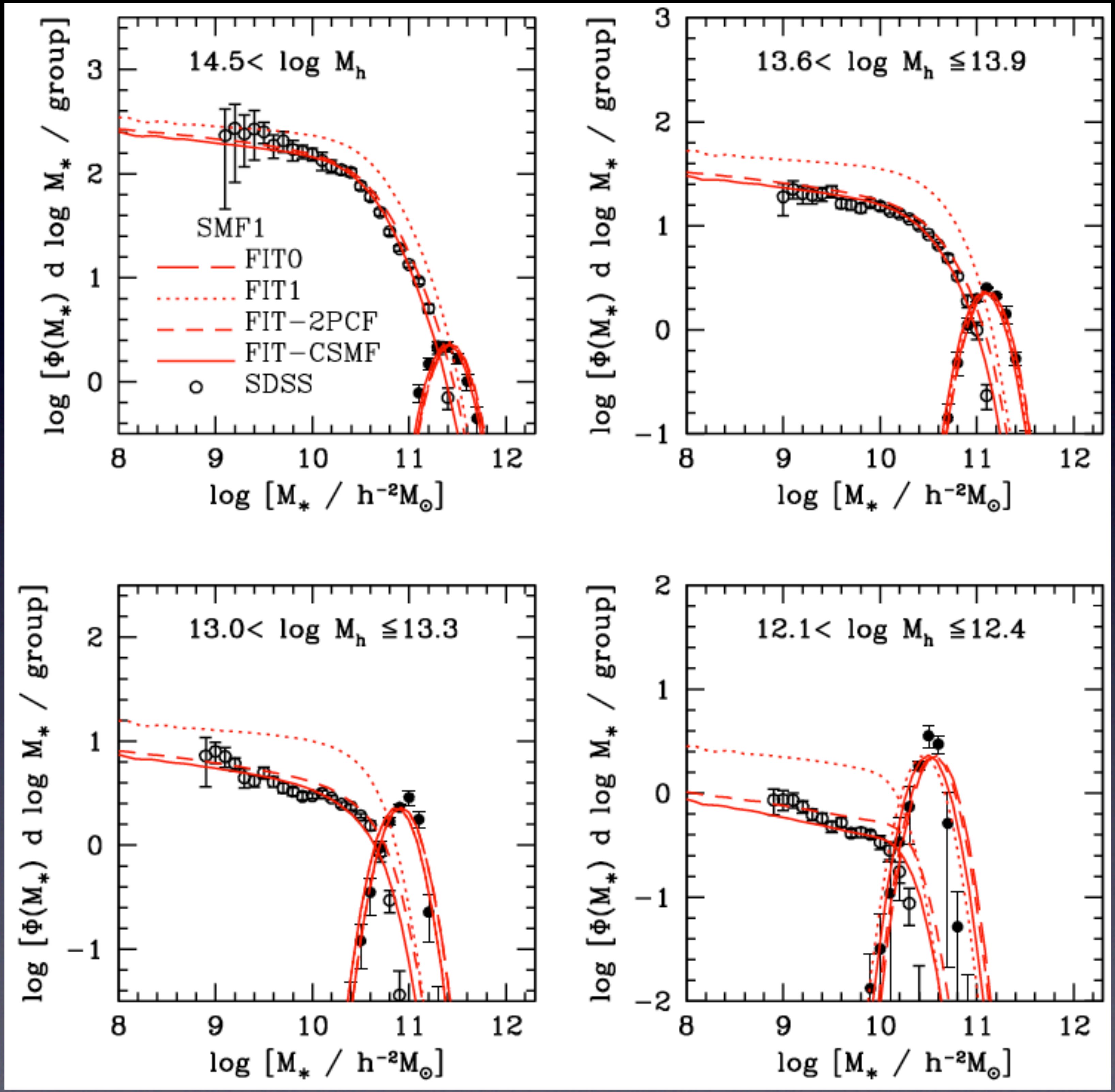


Source: Yang et al. 2012, ApJ, 752, 41

Fit 0 : $\alpha=0$ --> no sats
Fit 1 : $\alpha=\infty$ --> no evolution
2PCF : fit to $\Phi(M_\star)$ + 2PCF
CSMF : fit to $\Phi(M_\star) + \Phi(M_\star|M,z=0)$

The Galaxy - Dark Matter Connection

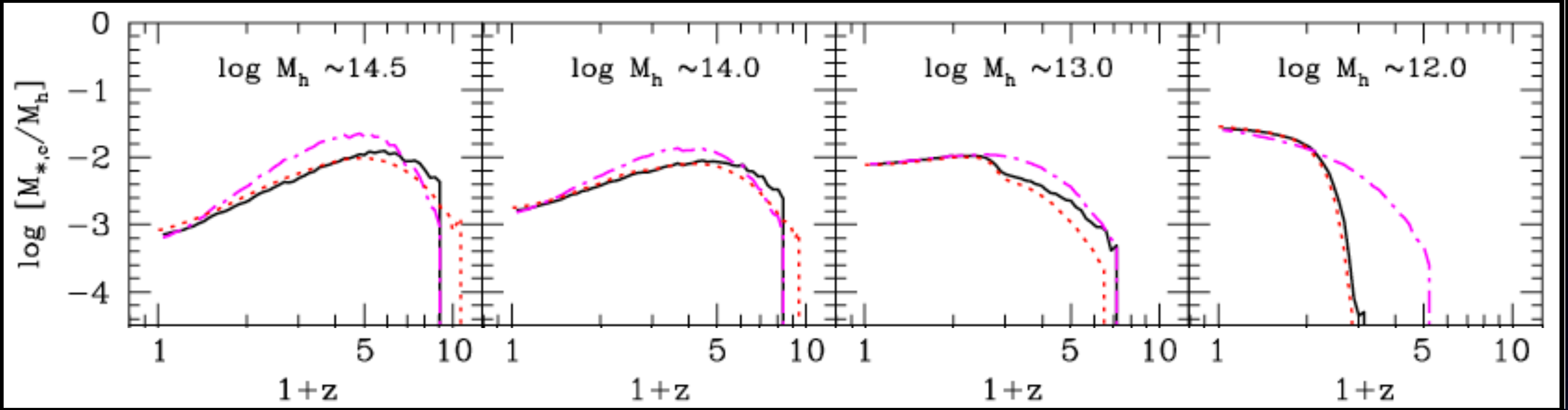
Data: SDSS Galaxy Group Catalogues
(Yang et al. 2009)



Source: Yang et al. 2012, ApJ, 752, 41

Fit 0 : $\alpha=0$ --> no sats
Fit 1 : $\alpha=\infty$ --> no evolution
2PCF : fit to $\Phi(M_\star)$ + 2PCF
CSMF : fit to $\Phi(M_\star) + \Phi(M_\star|M,z=0)$

The Galaxy - Dark Matter Connection



Source: Yang et al. 2012, ApJ, 752, 41

Stellar mass assembly history of galaxies is completely decoupled from mass assembly histories of their dark matter haloes

Virtually all galaxy formation models try to establish such a `decoupling' with feedback from supernovae and/or AGN:

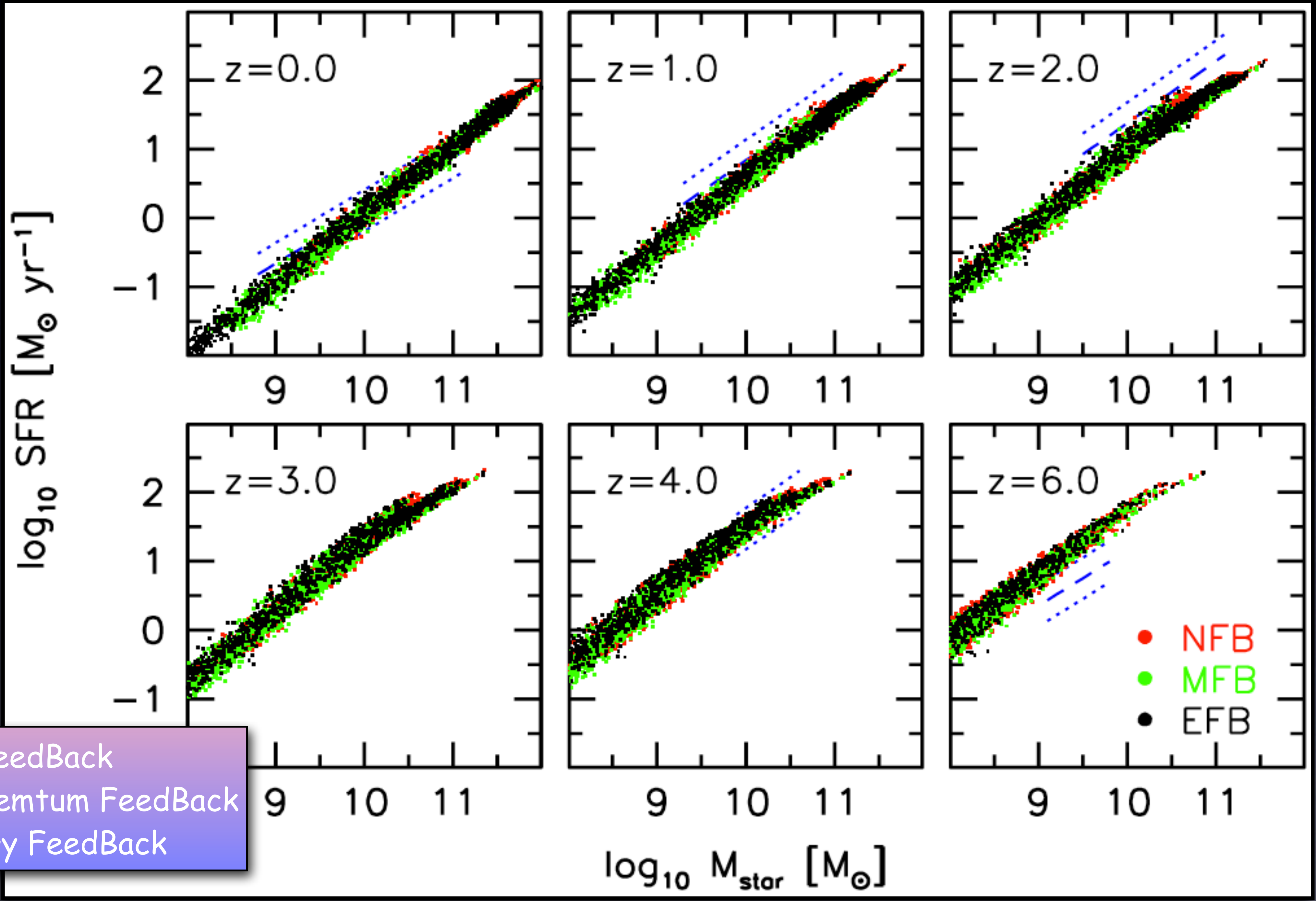
- AGN feedback is invoked to quench centrals in massive haloes
- SN feedback is invoked to suppress SF in low mass haloes.

The MAIN problem with current SAMs (and hydro-sims) is that the latter doesn't decouple stellar mass growth from halo mass growth, as I now demonstrate..

Origin of the Galaxy-SFR Sequence

Dutton et al. (2010) used detailed, spatially resolved SAM of (disk) galaxy formation, and showed that model naturally reproduces SFR-M* relation, independent of (SN) feedback. This is consequence of self-regulation that drives system to steady-state!

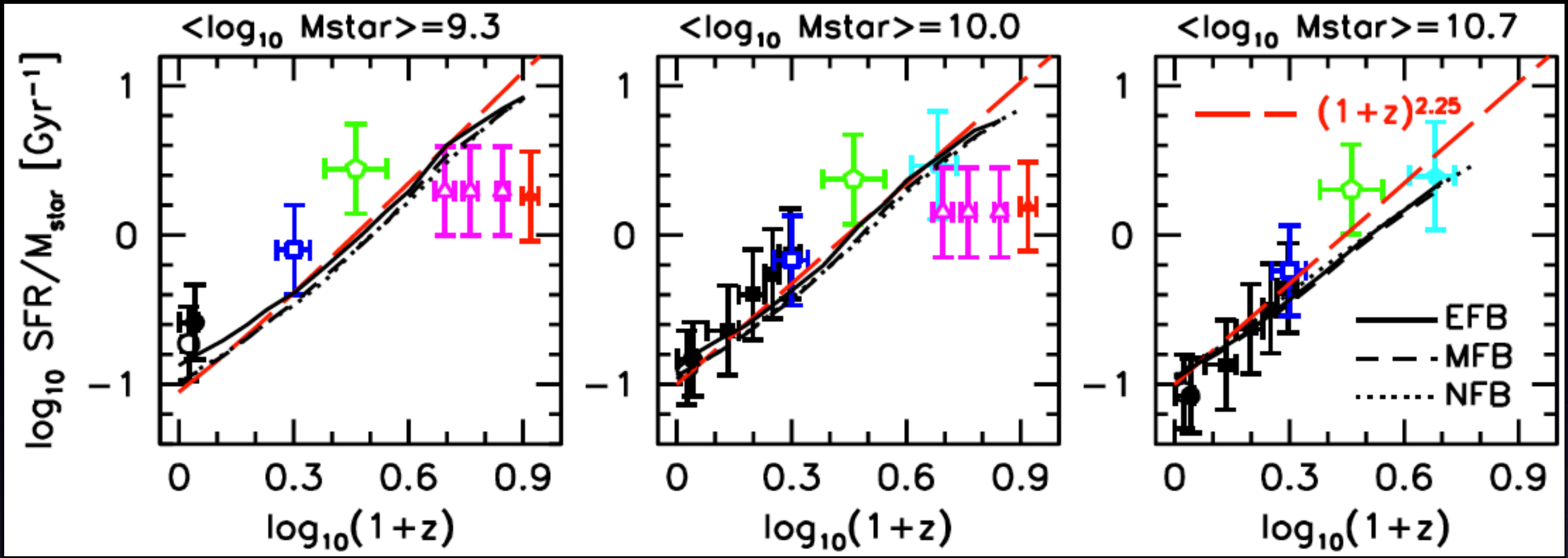
[cf. Bouché et al. (2010), Davé et al. (2010) as well as the lectures by Davé and Lilly]



NFB = No FeedBack
MFB = Momemtum FeedBack
EFB = Energy FeedBack

Source: Dutton, vdB & Dekel, 2010 MNRAS, 405, 1690

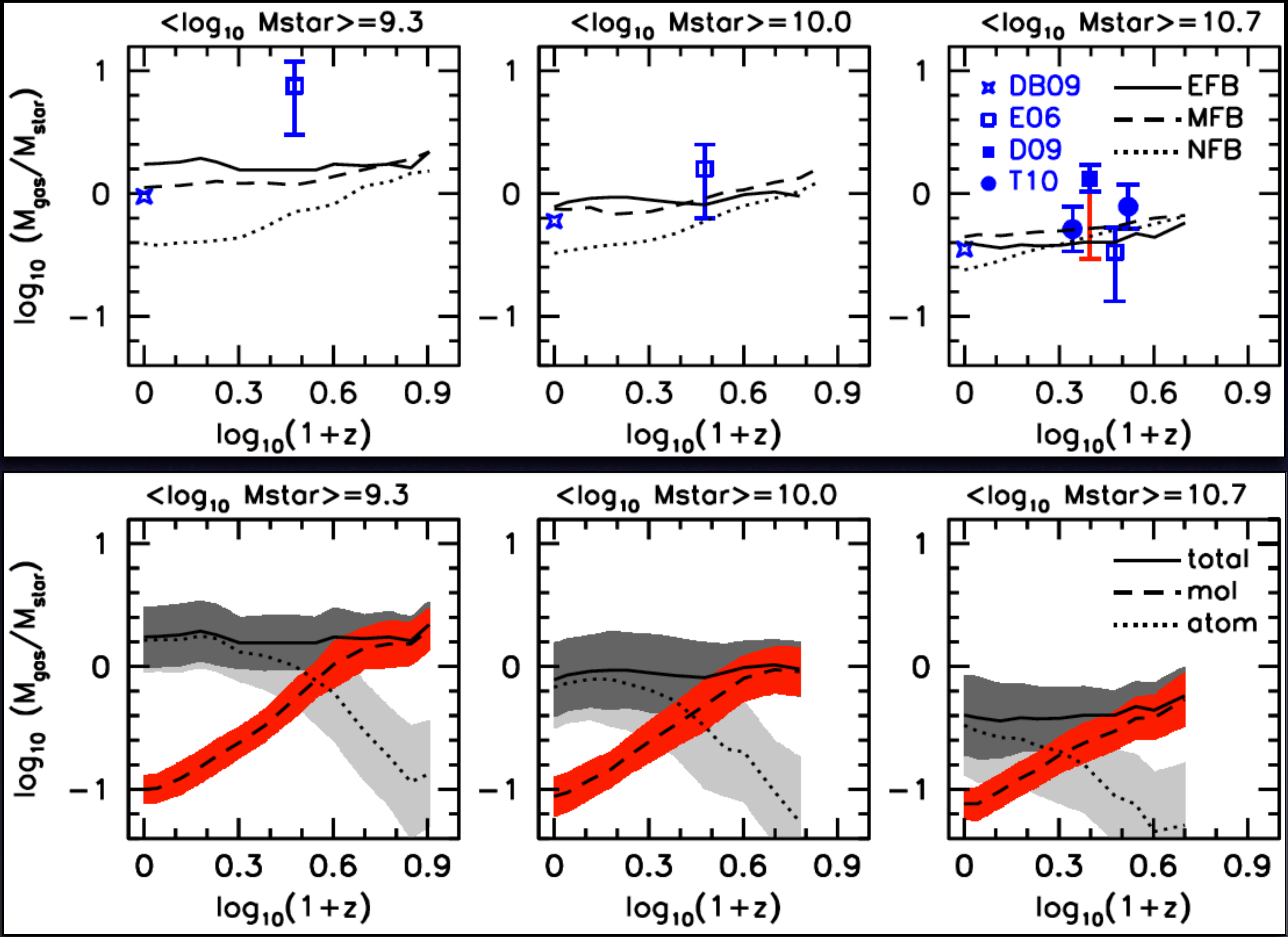
Origin of the Galaxy-SFR Sequence



Source: Dutton, vdB & Dekel, 2010 MNRAS, 405, 1690

The stellar mass growth closely follows the cosmological accretion rate of dark matter haloes, independent of SN feedback; Hence, SN feedback cannot decouple halo growth from galaxy growth...

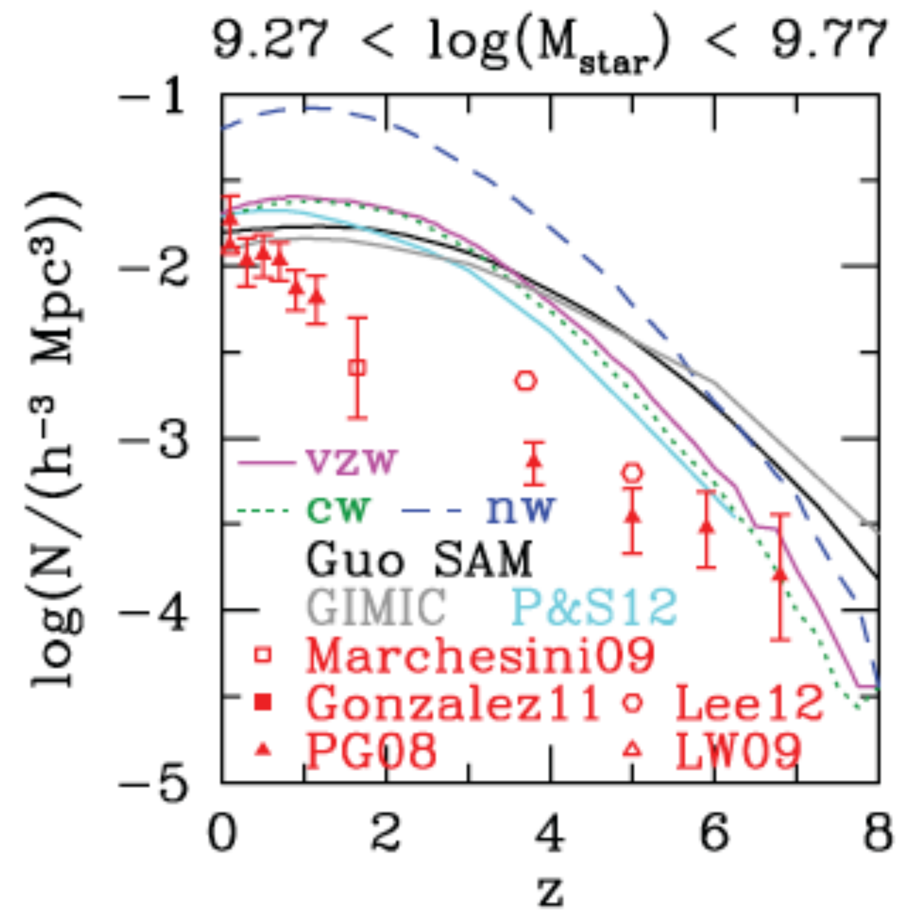
Origin of the Galaxy-SFR Sequence



Source: Dutton, vdB & Dekel, 2010 MNRAS, 405, 1690

Model predicts little to no evolution in gas mass fractions at fixed stellar mass, but strong evolution in ratio of molecular-to-atomic (consequence of density evolution).

Outstanding Problems



Source: Weinmann et al. 2012, MNRAS, 426, 2797

In my opinion, the most daunting problem for galaxy formation is understanding why stellar mass assembly is so detached from halo mass assembly.

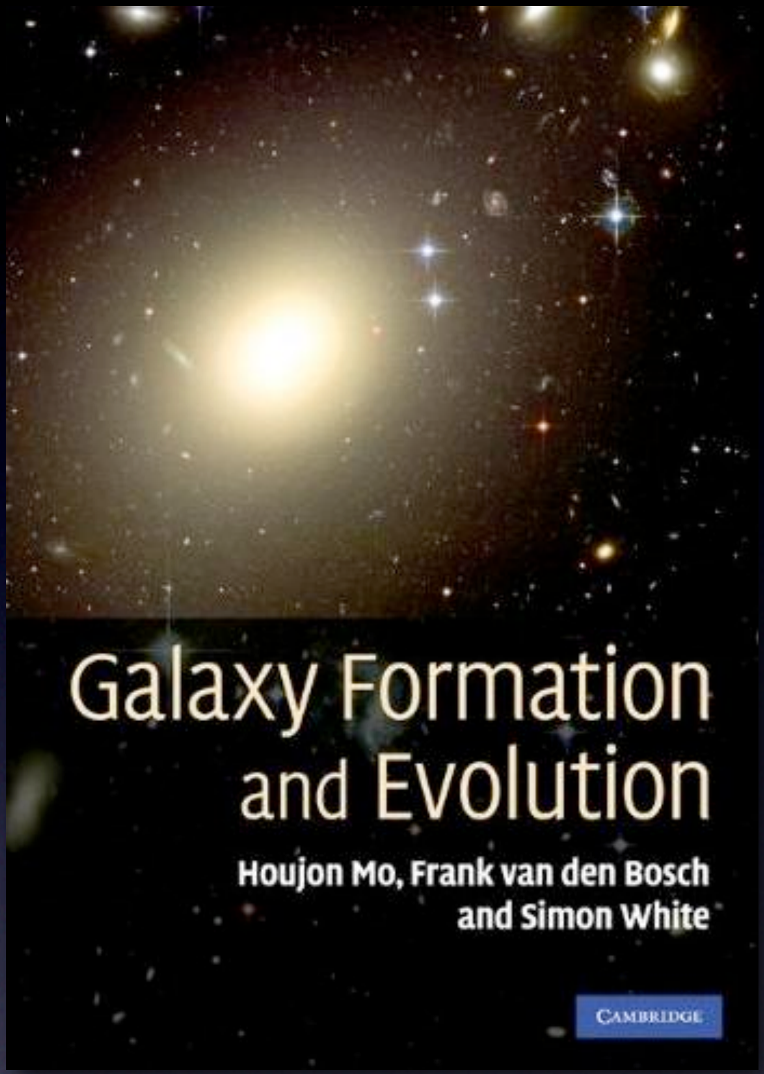
Most SAMs try to accomplish this using a combination of reionization and feedback from SN & AGN

The problem with SN feedback is that you need star formation to prevent star formation...

The problem with AGN feedback is that currently it is nothing but wishful thinking...

Galaxy Formation is far from a solved problem

For Further Reading...



A detailed treatment of the physical processes relevant for galaxy formation, and included in SAMs, can be found in the textbook Galaxy Formation and Evolution.

An excellent and up-to-date review of galaxy formation is

- Benson, 2010, Physical Reports, 495, 33

A detailed review of Semi-Analytical Models is

- Baugh, 2006, Reports on Progress in Physics, 69, 3101