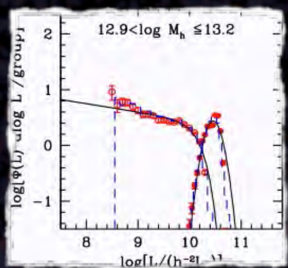
On the Assembly of Galaxies in Dark Matter Halos

...new insights from halo occupation modeling...

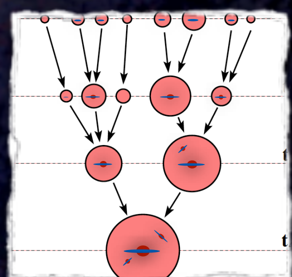
FRANK VAN DEN BOSCH
YALE UNIVERSITY

collaborators: Zhankhui Lu, Houjun Mo, Neal Katz, Martin Weinberg (UMass),
Xiaohu Yang, Youcai Zhang, Jiaxin Han (SHAO), Yu Lu (KICP)

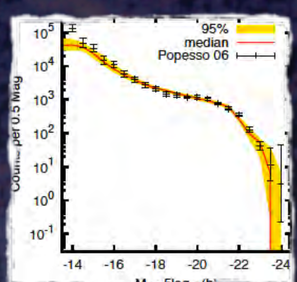
Outline



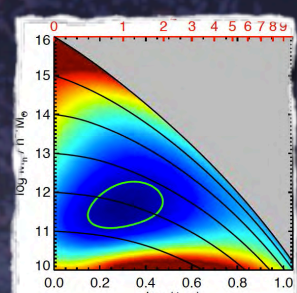
Halo Occupation Statistics



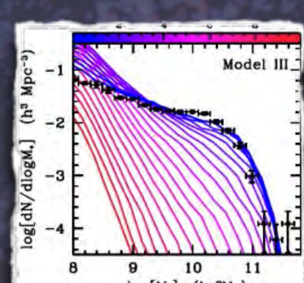
Semi-Analytical Models of Galaxy Formation



Empirical Modeling



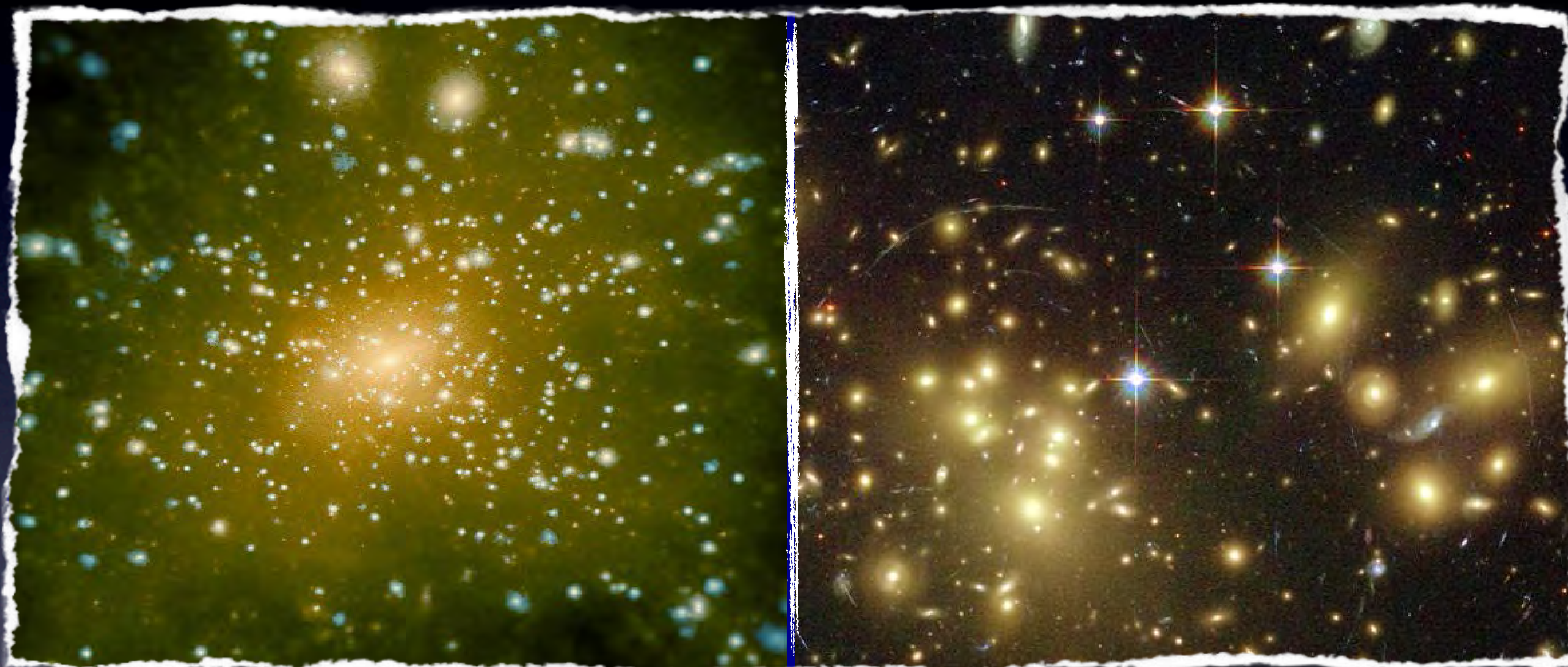
a self-consistent, dynamic model



forward modeling

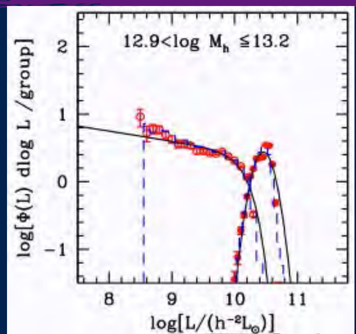
Halo Occupation Modeling: Motivation & Goal

Halo Occupation Modeling tries to establish a statistical description of the galaxy-dark matter connection, characterized by $\Phi(M_s|M_h)$



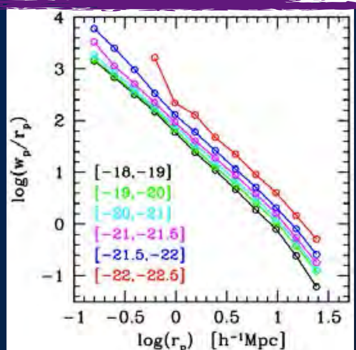
- Useful to constrain cosmological parameters
- Useful to constrain the physics of galaxy formation

Methods



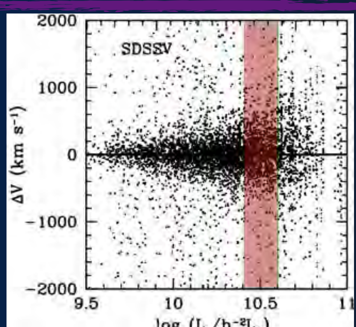
Galaxy-Group Catalogues

Eke et al. 2004; Yang et al. 2005, 2007, 2008, 2009; vdB et al. 2008; Weinmann et al. 2006a,b; Pasquali et al. 2010, 2012; Wetzel et al. 2012



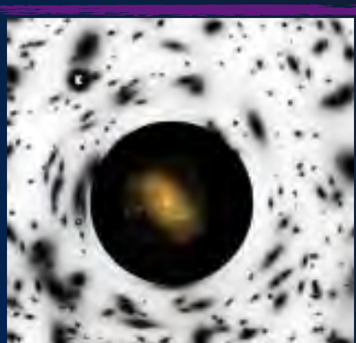
Galaxy Clustering

Jing, et al. 1998; Peacock & Smith 2000; Berlind & Weinberg 2002; Zheng 2004; Yang, Mo & vdB 2003; vdB, Yang & Mo 2003; Tinker et al. 2005; vdB et al. 2007



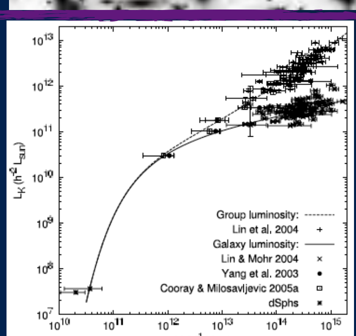
Satellite Kinematics

Zaritsky & White 1994; McKay et al 2002; Prada et al. 2003; vdB et al. 2004; Conroy et al. 2005; Norberg, Frenk & Cole 2008; More et al. 2009, 2011;



Galaxy-Galaxy Lensing

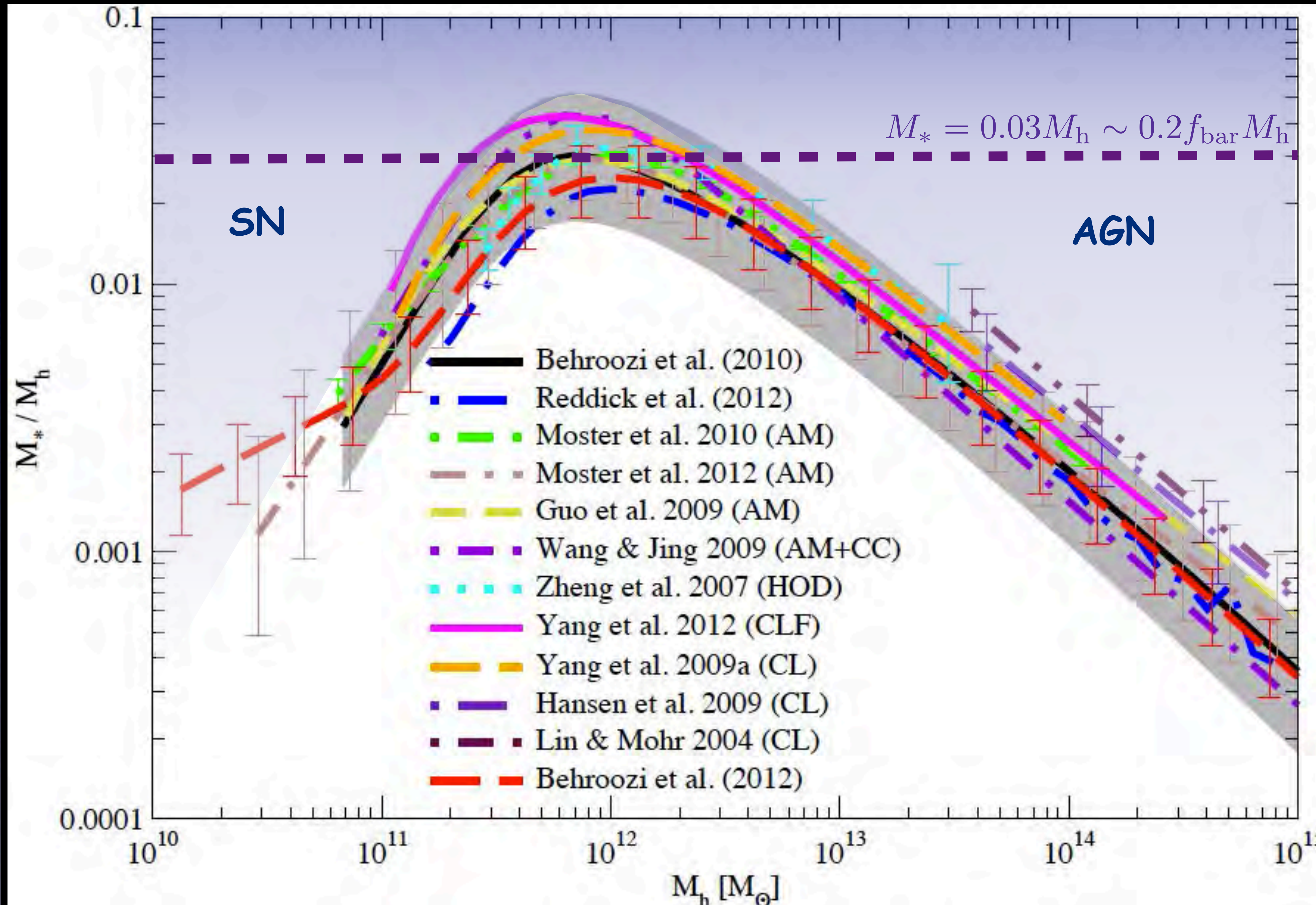
Guzik & Seljak 2002; Seljak et al. 2005; Mandelbaum et al. 2006; Yoo et al. 2006 Cacciato et al. 2009; van Uitert et al. 2011; Leauthaud et al. 2012;



Sub-Halo Abundance Matching

Vale & Ostriker 2004, 2006; Conroy et al. 2006; Shankar et al. 2006; Conroy & Wechsler 2009; Moster et al. 2010; Behroozi et al. 2010; Wetzel & White 2010

The Galaxy-Dark Matter Connection



Source: Behroozi, Wechsler & Conroy, 2013

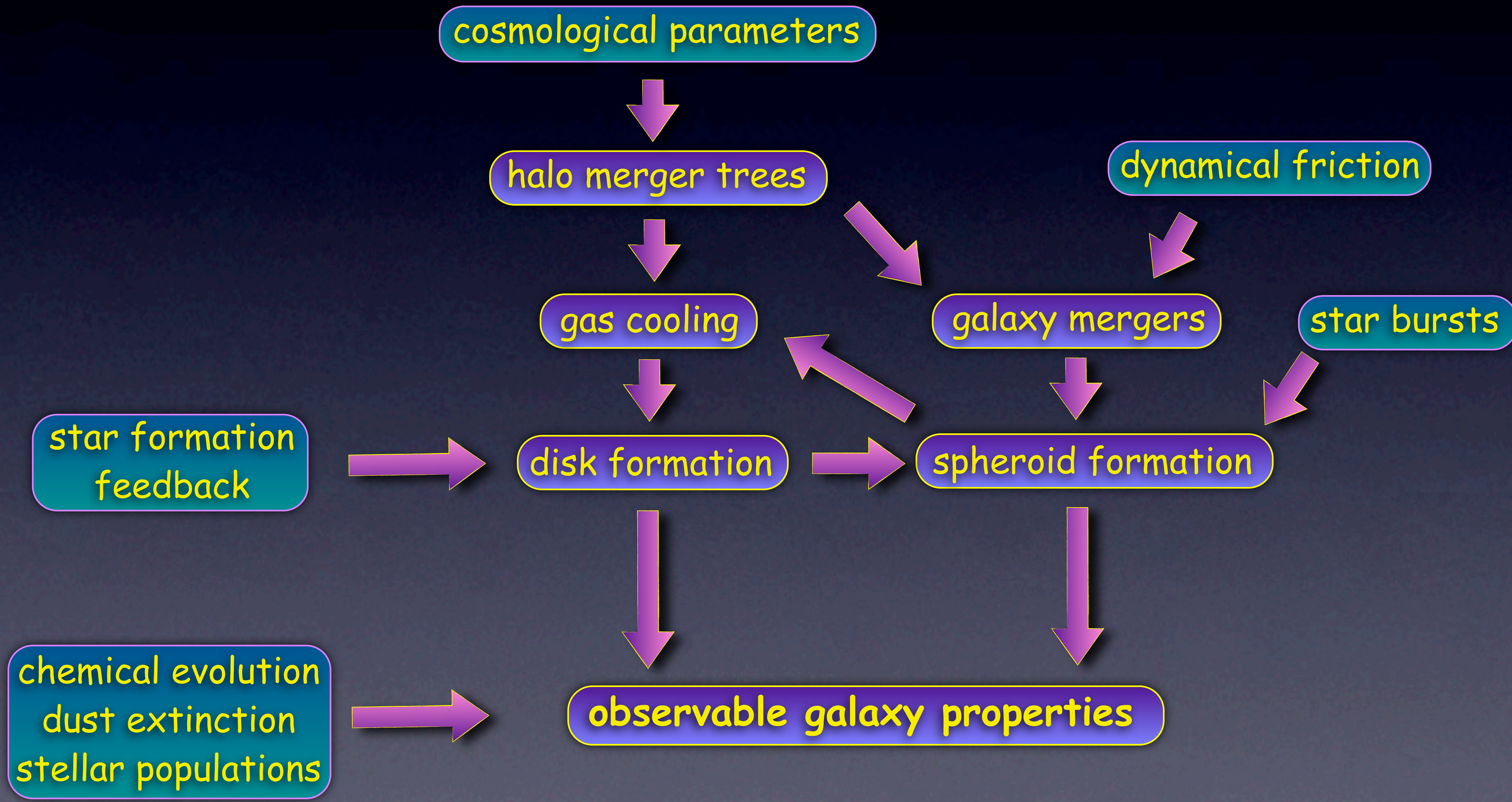
Take Home Message 1

Due to great advances in data, we now have a robust, statistical description of the galaxy-dark matter connection...

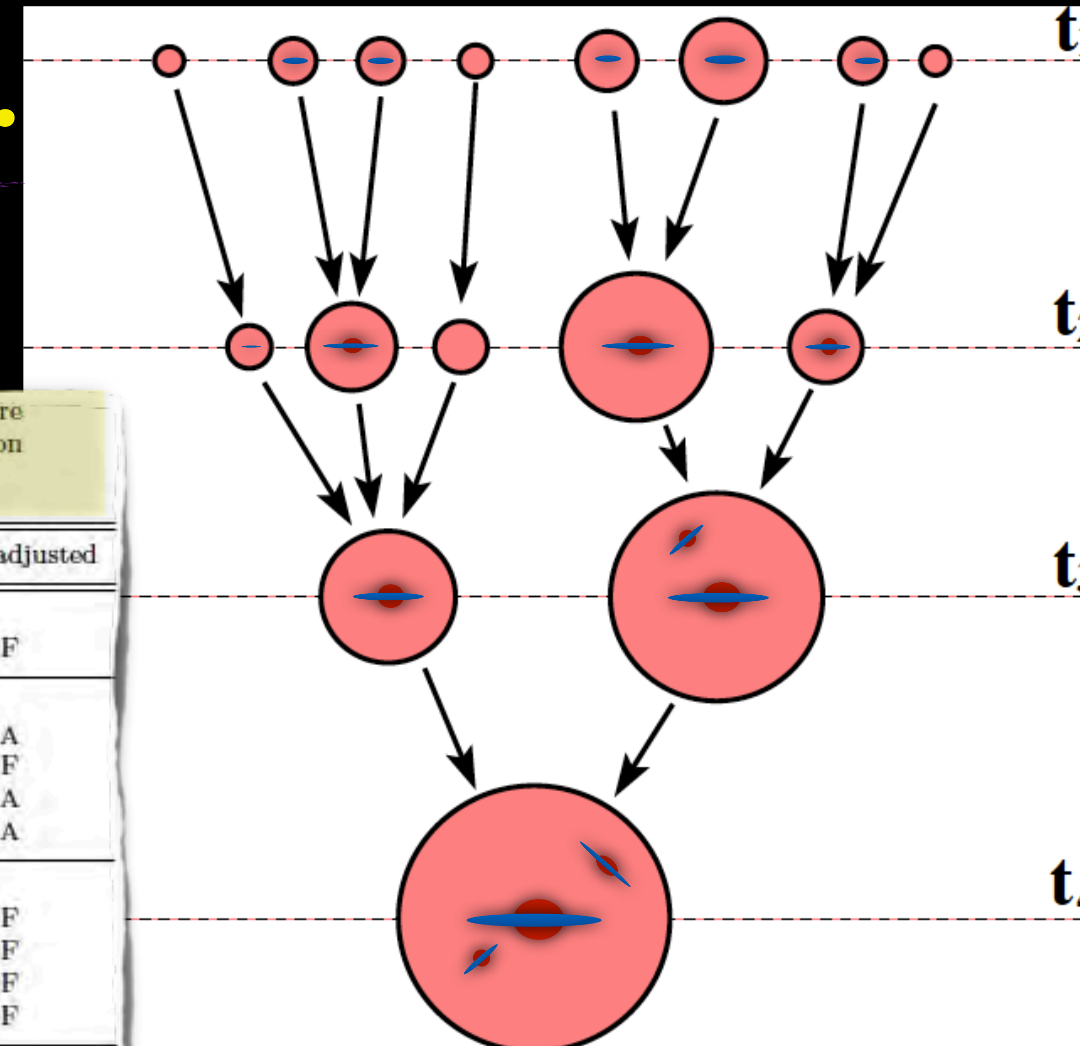
What does it tell us about galaxy formation?

Semi-Analytical Models

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.



Galaxy Formation is `complex'...



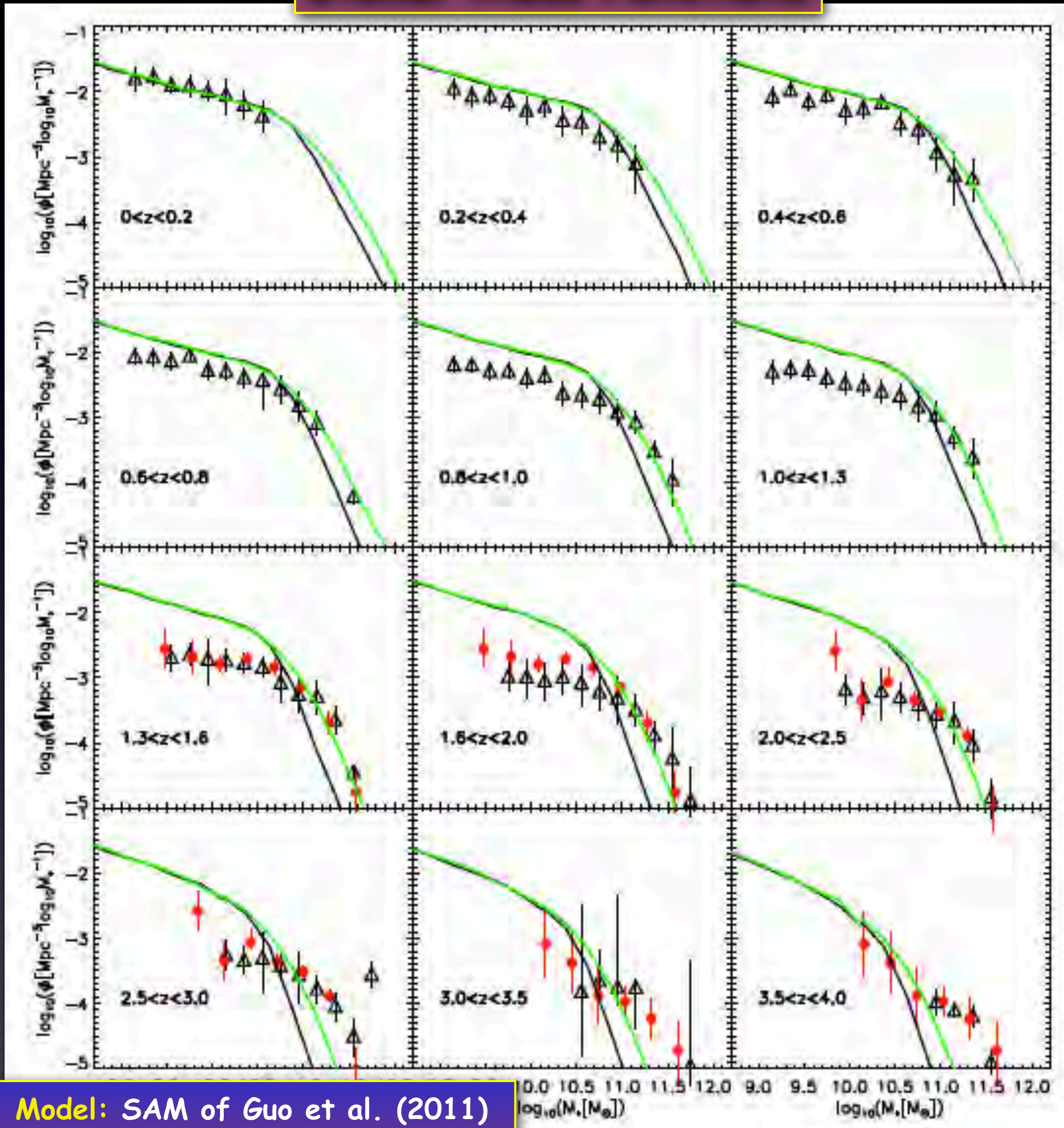
'able 2. Summary of the galaxy formation parameters in our "fiducial" model. We also specify the section in the paper where more detailed definition of each set of parameters can be found, and whether the parameter is considered to be fixed based on irect observations or numerical simulations (F), or adjusted to match observations (A).

parameter	description	fiducial value	fixed/adjusted
photoionization squelching (§2.3)			
$z_{\rm overlap}$, $z_{\rm reionize}$	redshift of overlap/reionization	11, 10	F
quiescent star formation (§2.5.1)			
$A_{\rm Kenn}$	normalization of Kennicutt Law [$M_\odot yr^{-1}$ kpc^{-2}]	8.33×10^{-5}	A
$N_{\rm K}$	power law index in Kennicutt Law	1.4	F
$\chi_{\rm gas}$	scale radius of gas disk, relative to stellar disk	1.5	A
$\Sigma_{\rm crit}$	critical surface density for star formation [$M_\odot pc^{-2}$]	6.0	A
burst star formation (§2.5.2)			
$\mu_{\rm crit}$	critical mass ratio for burst activity	0.1	F
$e_{\rm burst,0}$	burst efficiency for 1:1 merger	eqn. 9	F
$\gamma_{\rm burst}$	dependence of burst efficiency on mass ratio	eqn. 8	F
$\tau_{\rm burst}$	burst timescale	eqn. 10	F
merger remnants & morphology (§2.6)			
$f_{\rm sph}$	fraction of stars in spheroidal remnant	eqn. 11	A
$f_{\rm scatter}$	fraction of scattered satellite stars	0.4	A
supernova feedback (§2.7)			
$\epsilon^0_{\rm SN}$	normalization of reheating function	1.3	A
$\alpha_{\rm rh}$	power law slope of reheating function	2.0	A
$V_{\rm eject}$	velocity scale for ejection of reheated gas [km/s]	120	A
$\chi_{\rm reinfall}$	timescale for re-infall of ejected gas	0.1	A
chemical evolution (§2.8)			
y	chemical yield (solar units)	1.5	A
R	recycled fraction	0.43	F
black hole growth (§2.9)			
$\eta_{\rm rad}$	efficiency of conversion of rest mass to radiation	0.1	F
$M_{\rm seed}$	mass of seed BH [$M_\odot$]	100	F
$f_{\rm BH,final}$	scaling factor for mass of BH at end of merger	2.0	A
$f_{\rm BH,crit}$	scaling factor for "critical mass" of BH	0.4	F
AGN-driven winds (§2.10)			
$\epsilon_{\rm wind}$	effective coupling factor for AGN driven winds	0.5	F
radio mode feedback (§2.11)			
$\kappa_{\rm radio}$	normalization of "radio mode" BH accretion rate	3.5×10^{-3}	A
$\kappa_{\rm heat}$	coupling efficiency of radio jets with hot gas	1.0	F

Source: Somerville et al. (2008)

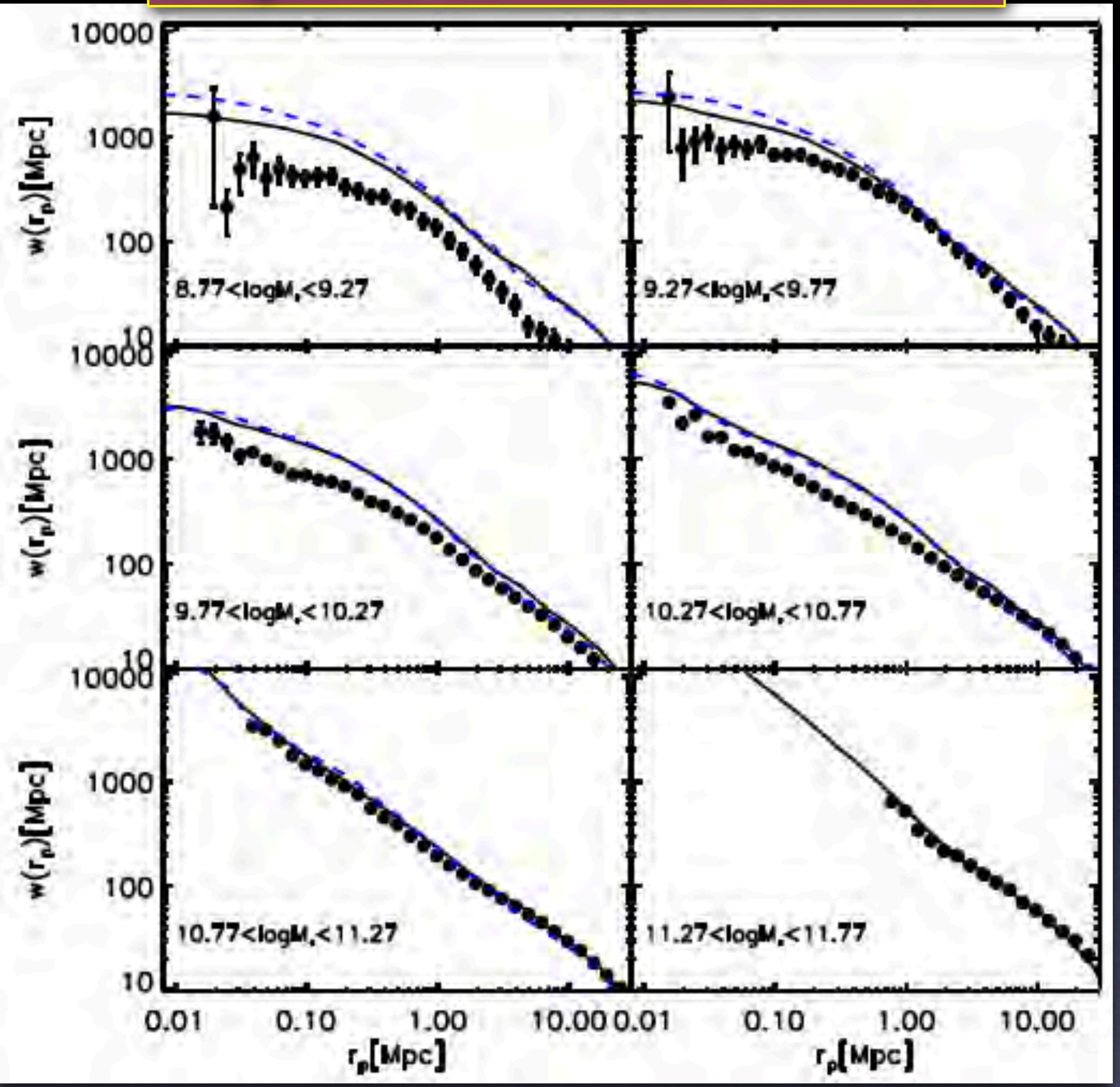
...and above all, Galaxy Formation is `unsolved'...

Stellar Mass Functions



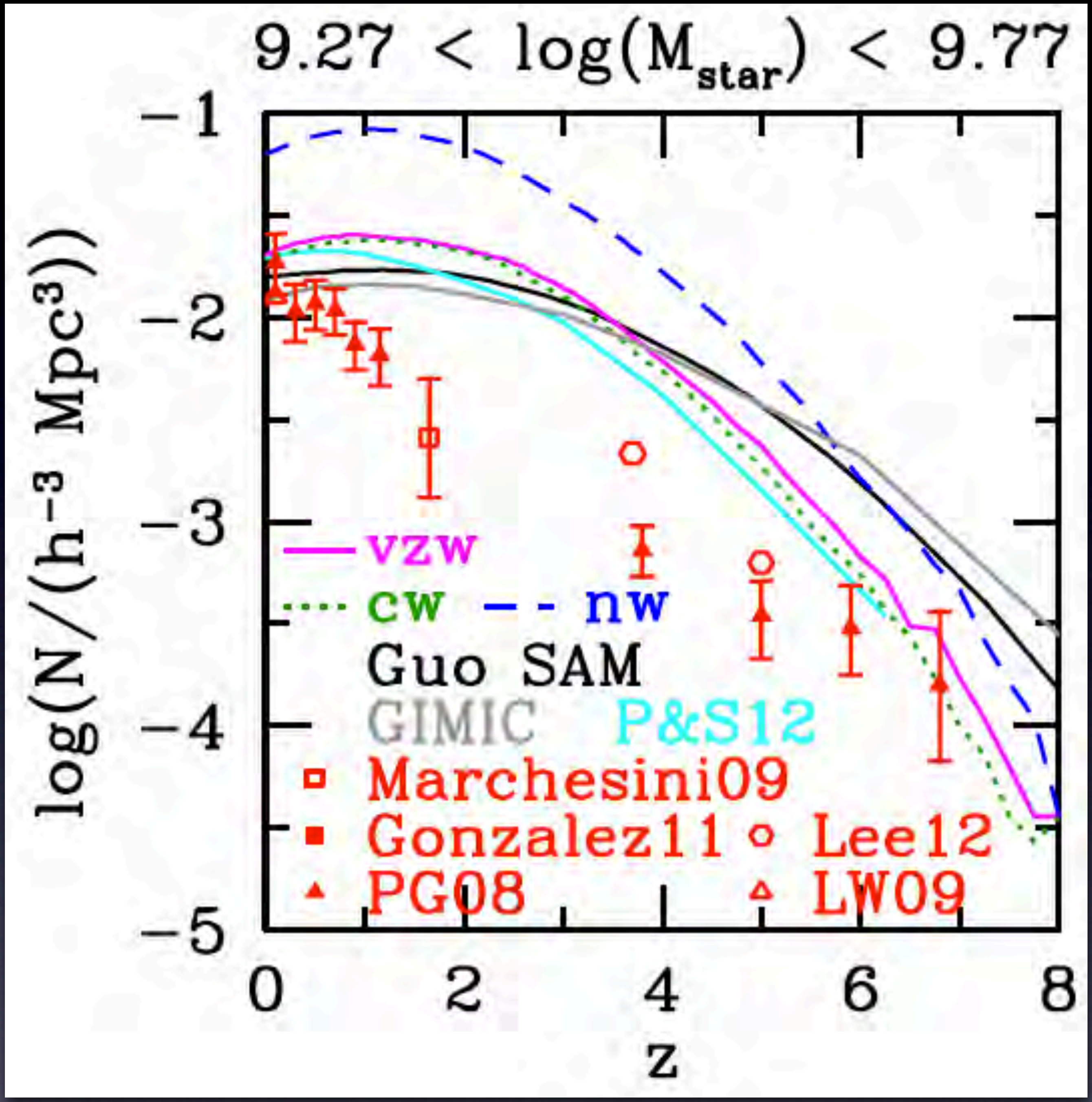
Model: SAM of Guo et al. (2011)
Data: Marchesini et al. (2009)
Perez-Gonzales et al. (2008)

Projected Correlation Functions



Source: Guo et al. (2011)

…and above all, Galaxy Formation is ‘unsolved’…



Source: Weinmann, Pasquali et al. (2011)

Neither SAMs nor SIMs reproduce assembly histories of low mass galaxies

Take Home Message 2

Despite a large number of free parameters, SAMs & SIMs fail to reproduce even the most basic observables of the galaxy population...

↓

ask Google for help

Google: how do galaxies evolve?

Samsung Galaxy S

Samsung Galaxy SII

Samsung Galaxy SIII

Take Home Message 2

Despite a large number of free parameters, SAMs & SIMs fail to reproduce even the most basic observables of the galaxy population...

Back to the drawing board

Empirical Modelling

can we construct self-consistent models for stellar mass assembly of galaxies in dark matter halos that are consistent with

1) the data

2) the LCDM paradigm ???

if 'yes'; what does it tell us about galaxy formation?

Towards a Self-Consistent, Empirical Model

Step 1: constrain conditional stellar mass function across cosmic time

data $\Phi(M_*, z)$	$\Phi(M_*\vert M_{\rm h}, z) = \Phi_{\rm c}(M_*\vert M_{\rm h}, z) + \Phi_{\rm s}(M_*\vert M_{\rm h}, z)$
theory $\Phi(M_{\rm h}, z)$	

Self-consistency constraint: $\Phi_{\rm s}(M_*|M_{\rm h}, z)$ must depend on $\Phi_{\rm c}(M_*|M_{\rm h}, > z)$

Step 2: combine with mass assembly histories of dark matter halos to construct stellar mass assembly histories (for centrals)

model $\Phi_{\rm c}(M_*\vert M_{\rm h}, z)$	$M_{*,{\rm c}}(z\vert M_{\rm h,0})$
theory $M_{\rm h}(z\vert M_{\rm h,0})$	

Step 3: Time derivative yields SFR after correcting for stellar evolution (mass loss) and mass accretion (cannibalism)

model $\Phi_{\rm s}(M_*\vert M_{\rm h}, z)$	$\dot{M}_{*,{\rm c}}(z\vert M_{*,0})$
model $M_{*,{\rm c}}(z\vert M_{\rm h,0})$	

A Dynamic, Self-Consistent Model

Yang et al. 2011, ApJ, 741, 13
Yang et al. 2012, ApJ, 752, 41
Yang et al. 2013, ApJ, 770, 115

central galaxies

$$\Phi_c(M_*|M,z) = \frac{1}{2\pi\sigma_c} \mathrm{EXP}\left[-\frac{(\log M_*/\overline{M}_*)^2}{2\sigma_c^2}\right]$$

$$\left.\begin{array}{l} \overline{M}_* = \overline{M}_*(M,z) \\ \sigma_c = \sigma_c(z) \end{array}\right\} \text{9 free parameters}$$

satellite galaxies are centrals at infall:

$$\Phi_s(M_*|M,z) = \int_0^\infty \mathrm{d}M_{*,a} \int_0^M \mathrm{d}m_a \int_z^\infty \mathrm{d}z_a \int_0^M \mathrm{d}M_a \int_0^1 \mathrm{d}\eta \; \Phi_c(M_{*,a}|m_a,z_a)\, n_{\rm sub}(m_a,z_a|M,z)$$

$$P(M_*,z|M_{*,a},z_a;m_a;M_a,\eta)\, P(M_a,z_a|M,z)\, P(\eta)$$

a simplified model for the evolution of satellites:

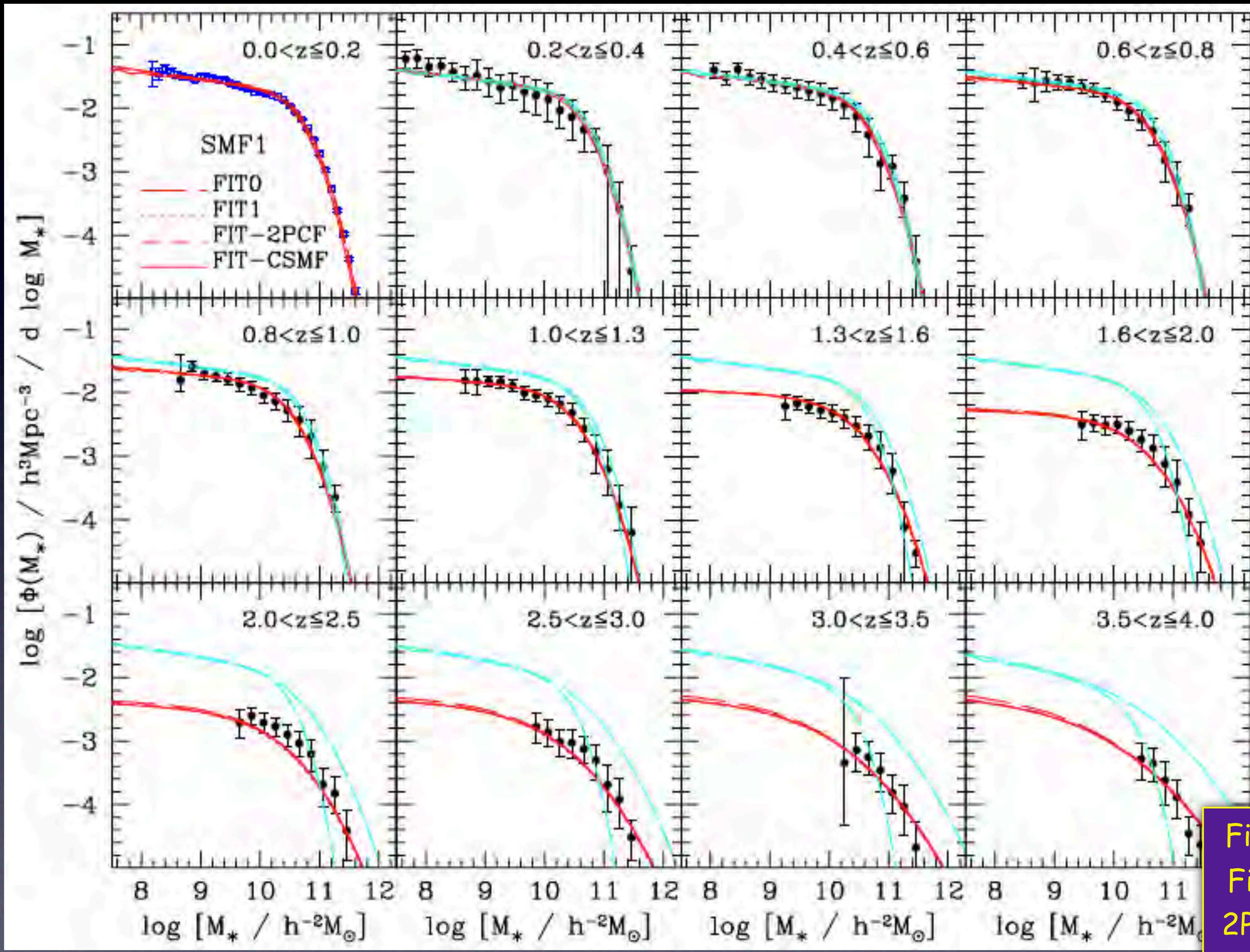
$$P(M_*,z|M_{*,a},z_a;m_a;M_a,\eta) = \begin{cases} \delta^{\rm D}(M_* - M'_*) & \text{if } \Delta t < \alpha\, t_{\rm df}(m,M,z,\eta) \\ 0 & \text{otherwise} \end{cases}$$

$$M'_* = (1-c)\, M_{*,a} + c\, \overline{M}_{*,c}(m_a,z)$$

α `satellite disruption' parameter
c `satellite mass growth' parameter

Fit to Stellar Mass Functions across Cosmic Time

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

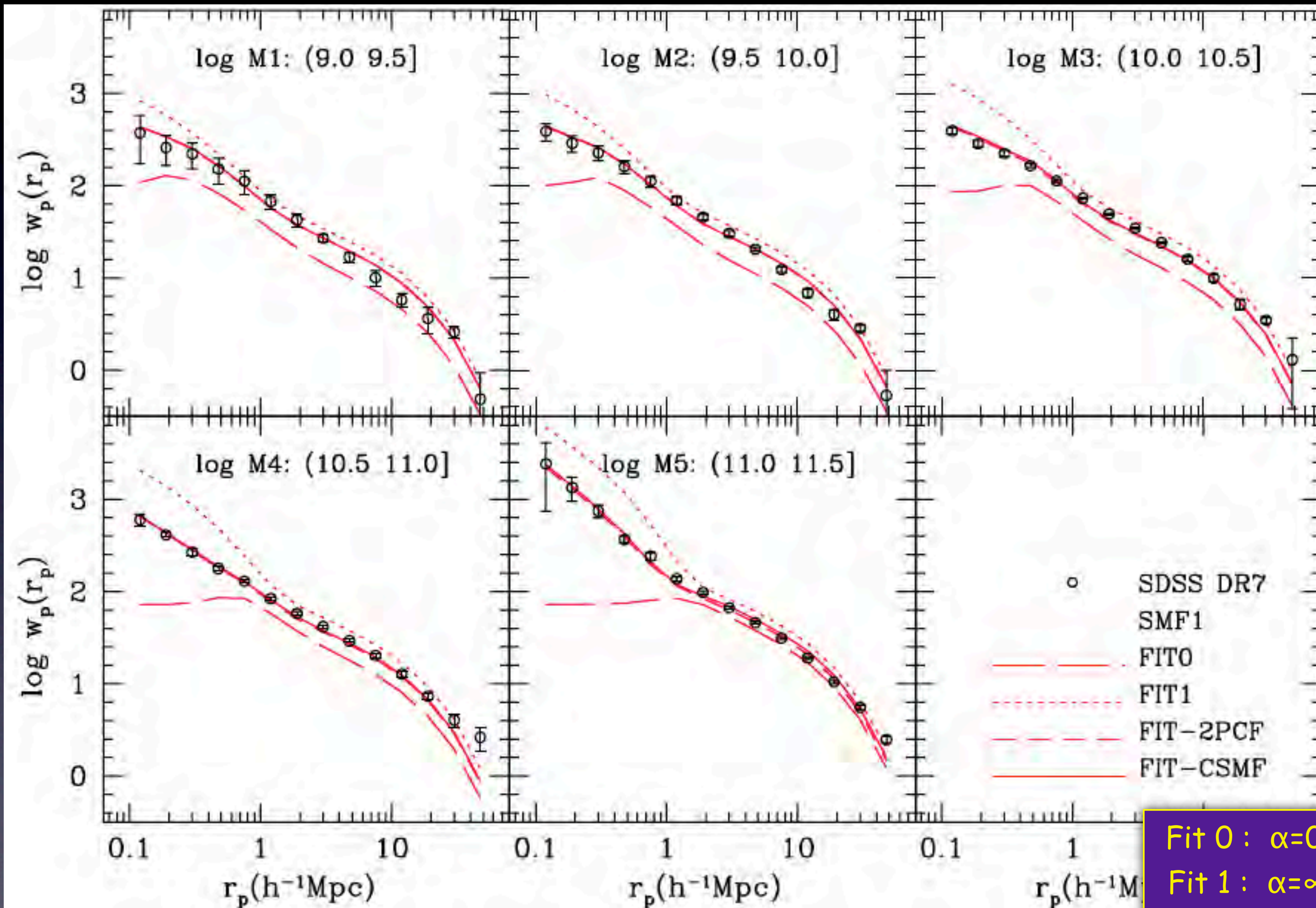


Fit 0 : α=0 --> no sats
Fit 1 : α=∞ --> no evolution
2PCF : fit to $\Phi(M_\star)$ + 2PCF
CSMF : fit to $\Phi(M_\star)$ + $\Phi(M_\star|M,z=0)$

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

Fit to Two-Point Correlation Functions at z=0.1

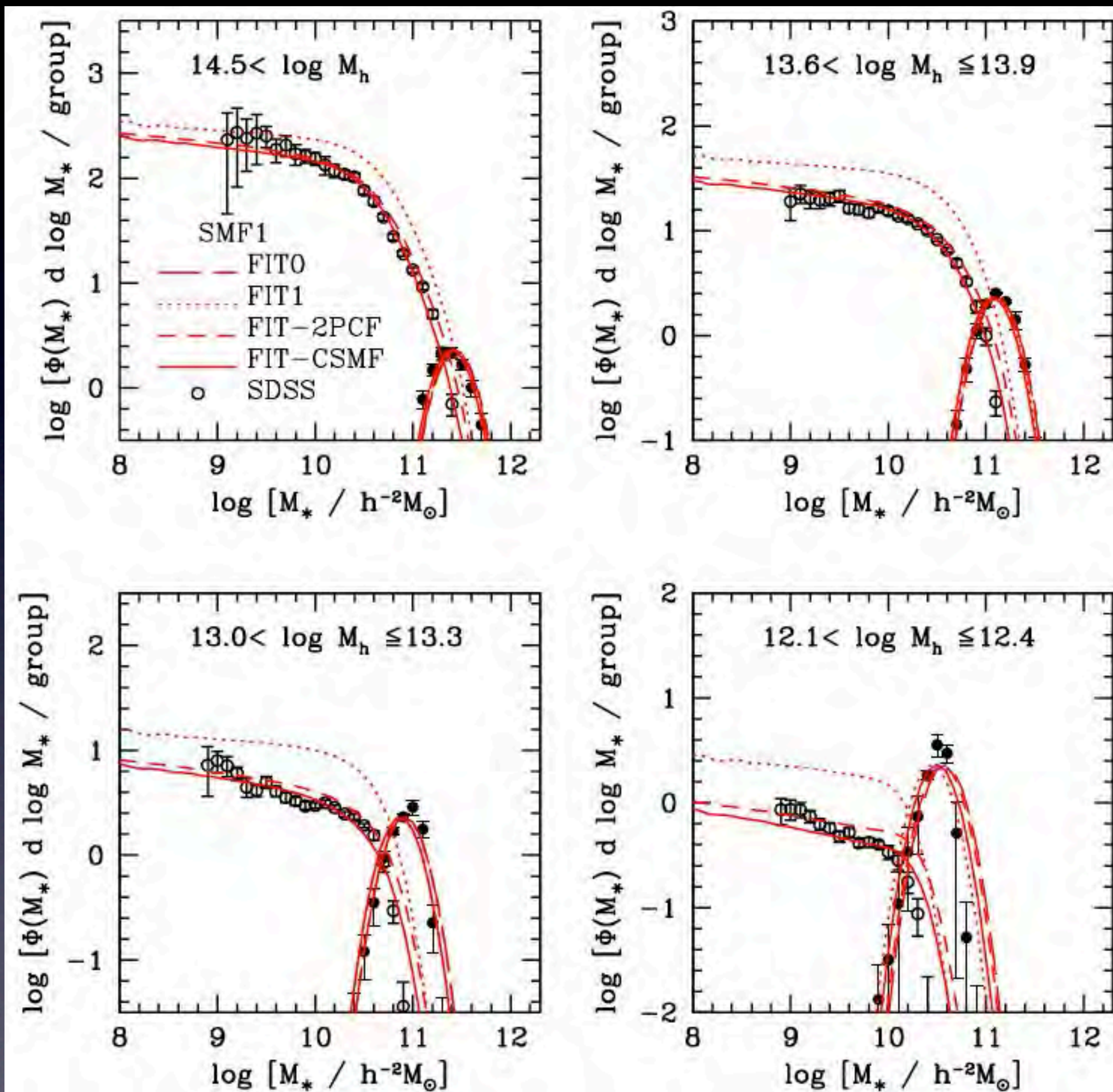
Data: SDSS DR7
(Yang et al. 2012)



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

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

Fit to Conditional Stellar Mass Functions at z=0.1



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

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

best-fit value for c ~ 0.95 +/- 0.05 indicating that sats continue to grow in stellar mass after accretion, in excellent agreement with recent results by Wetzel et al. (2012)

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

Take Home Message 3

Empirical models can easily fit all available data with only a modest set of free parameters

for the critics

"but there is no physics in your model"

- this does not make the model unphysical
- empirical models are not inhibited by restricted parameterizations of physical processes that are poorly understood
- empirical models are not the end-goal; they are first step in two-step `reverse engineering' approach
- empirical models "translate" opague data into a language more directly interpretable in framework of galaxy formation

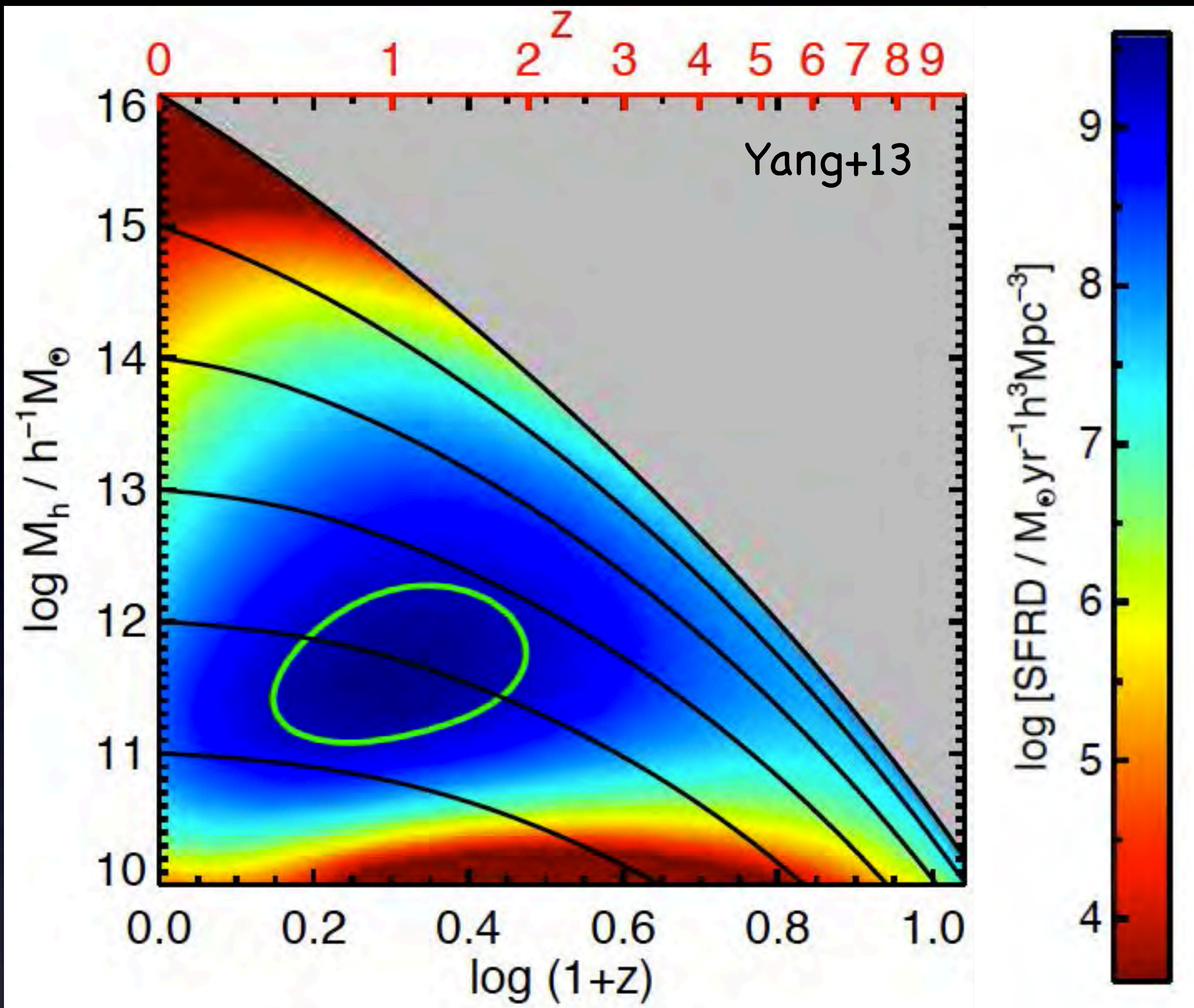
Empirical modeling is useful for <u>informing</u> galaxy formation theory

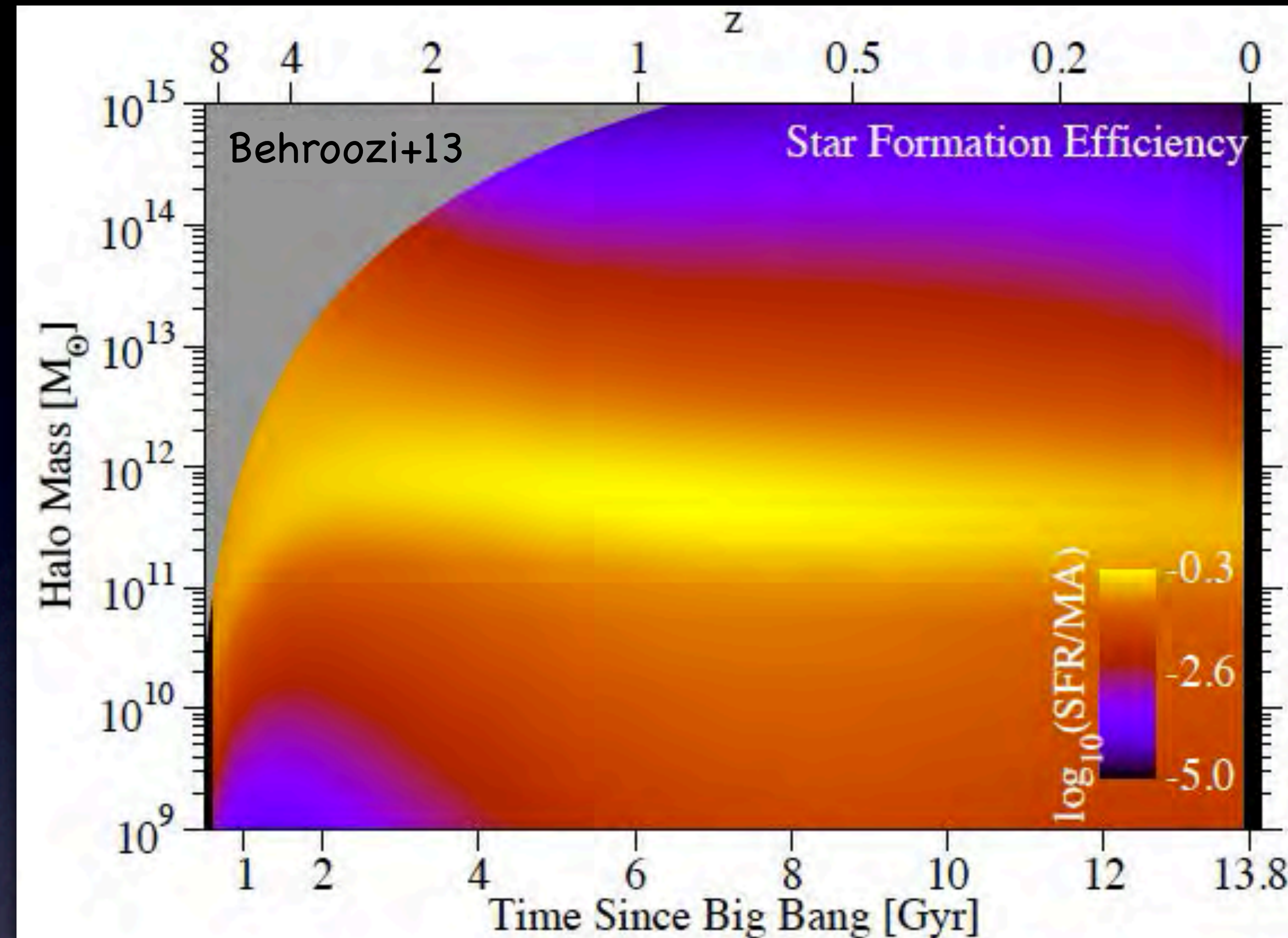
Take Home Message 3

Empirical models can easily fit all available data with only a modest set of free parameters

What insights can we gain regarding the physics of galaxy formation?

Star Formation Efficiencies across Cosmic Time



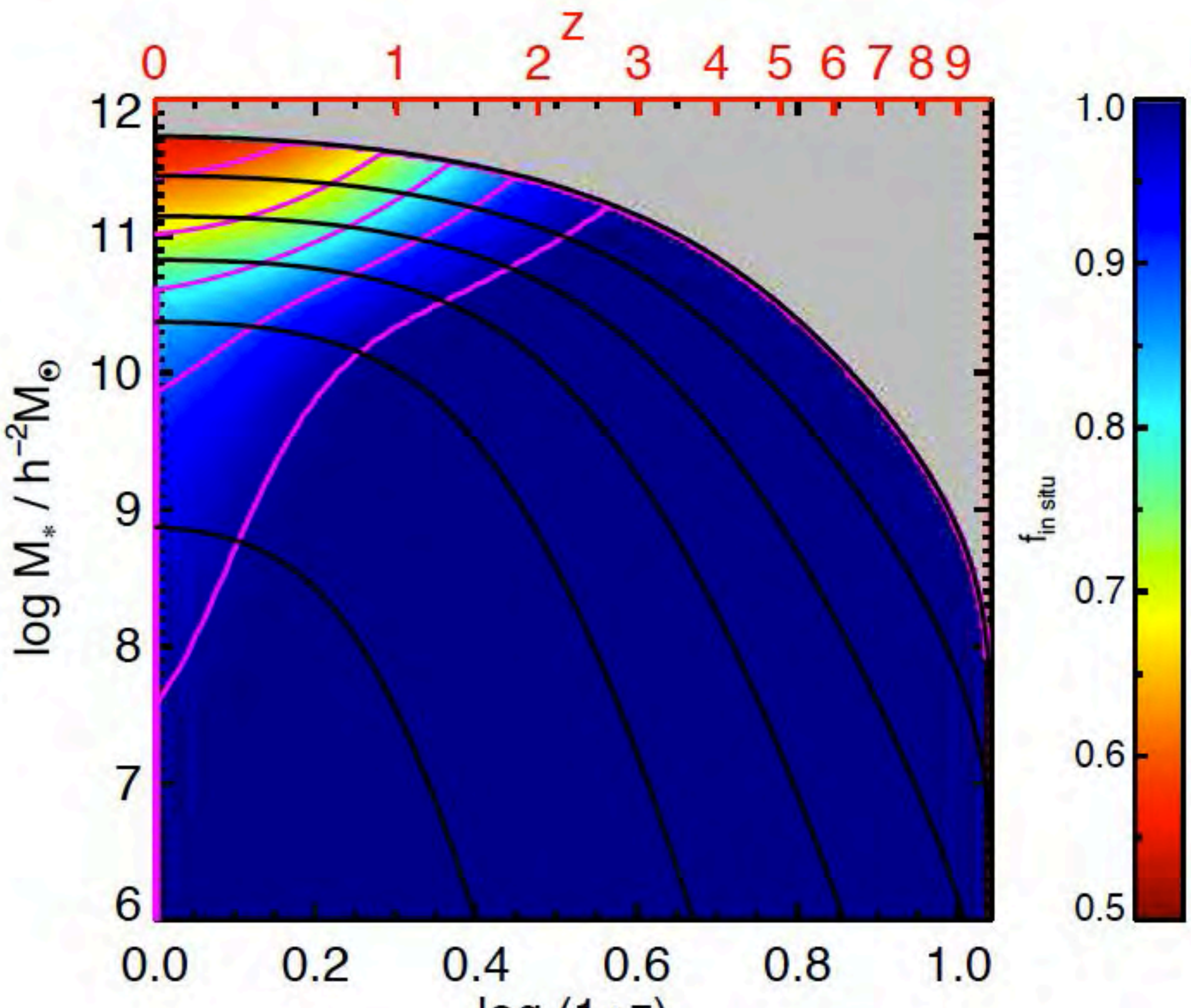


Empirical models show that the majority of stars form in dark matter halos with 10^{11} Msun < M_{halo} < 10^{12} Msun around z ~ 1 - 2.

see also: Bouche+10; Behroozi+13, Yang+13; Moster+13; Mutch+13

In-Situ Fractions

in-situ fraction: fraction of stars that formed in-situ, as opposed to were accreted via mergers.



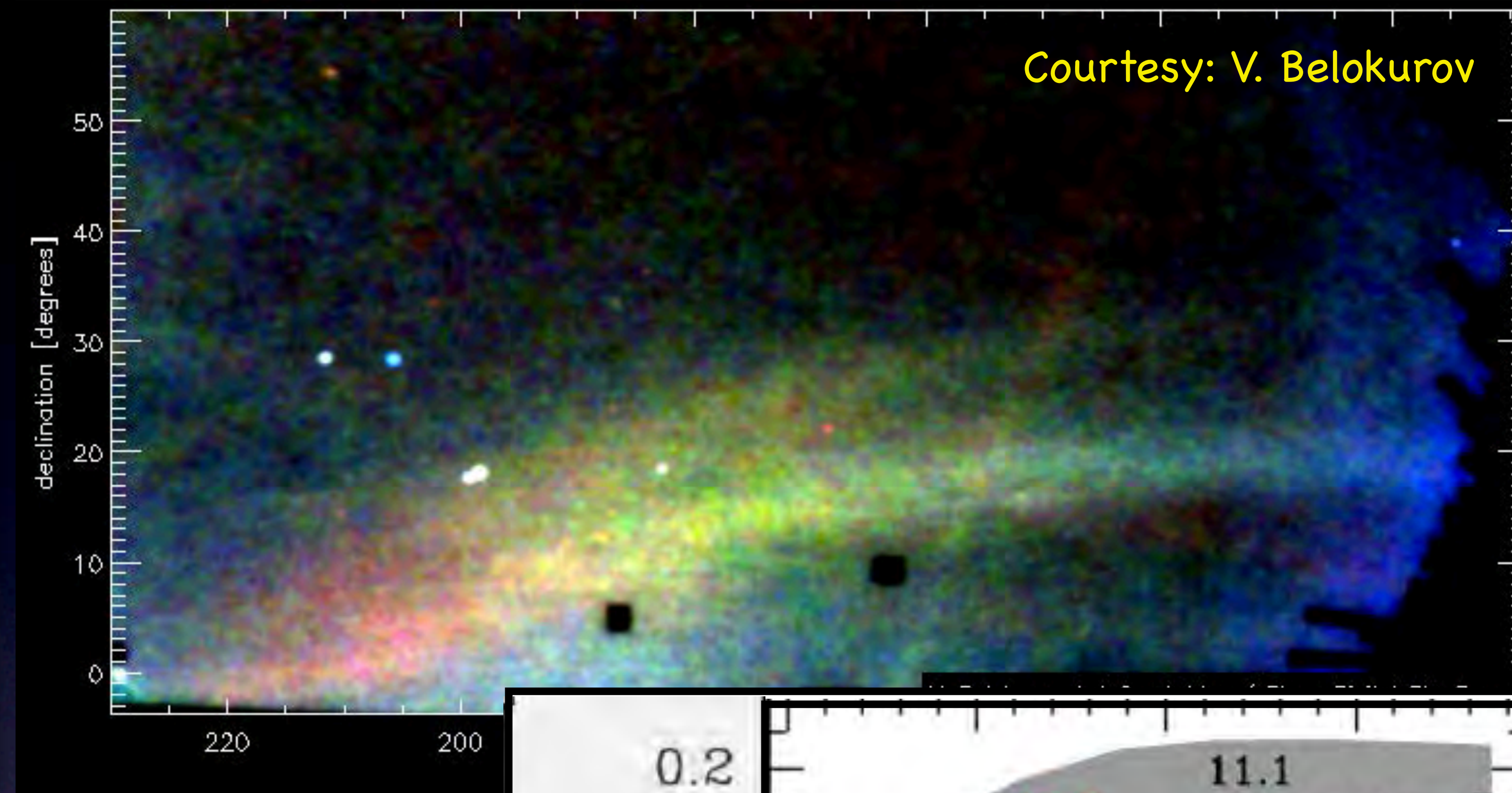
Yang et al. (2013)

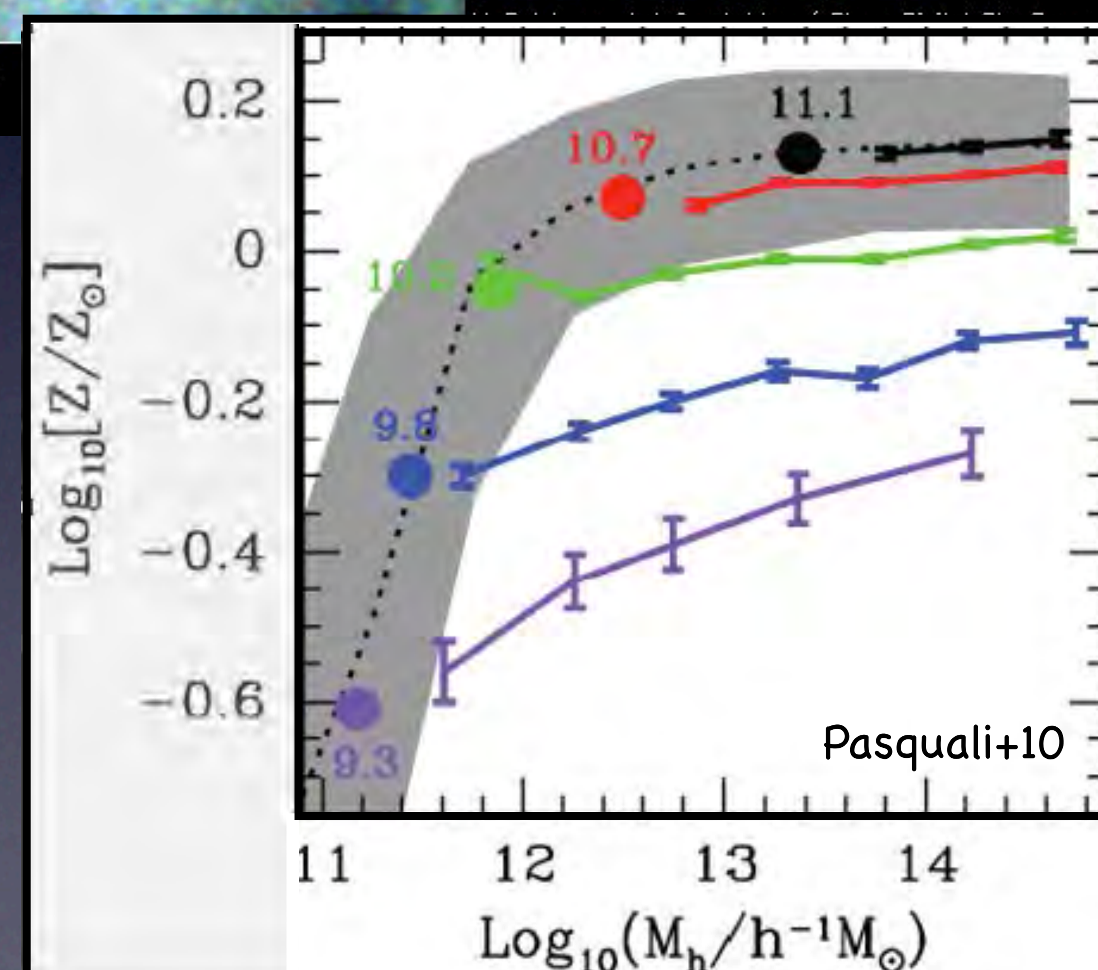
Mass assembly via mergers is only important for the most massive galaxies ($M_* > 10^{11}$ M_{sun}) and at low redshift ($z < 1$).

This idea that merging is only relevant in most massive galaxies is consistent with shape of M_*-M_{halo} relation.

It also implies that tidal disruption of satellites is very important!!!

Stripping & Disruption rules





If you look hard enough, you see evidence for stripping and disruption everywhere...

Take Home Message 4

Virtually all star formation occurs in halos in narrow range of halo mass ($10^{11} < M_h < 10^{12}$)

Merging is irrelevant, except for most massive galaxies

Satellite disruption is utterly important

Forward Approach; Galaxy Formation Simplified

Model

- Central galaxies form stars according to $SFR[M_h,z]$
- Satellite galaxies merge with centrals a time $t_{df}[M_s/M_h,z]$ after accretion.
- At time of merger, a fraction f_{ICL} of satellite stars go to stellar halo.

Method

Construct a set of halo merger trees

Propragate model through merger trees

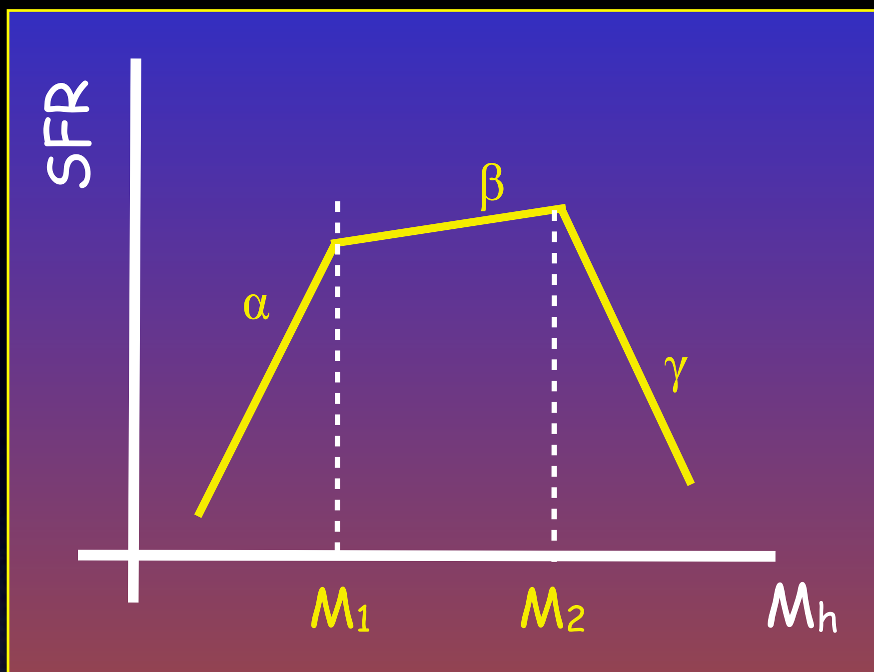
Compare model with data

Change model parameters

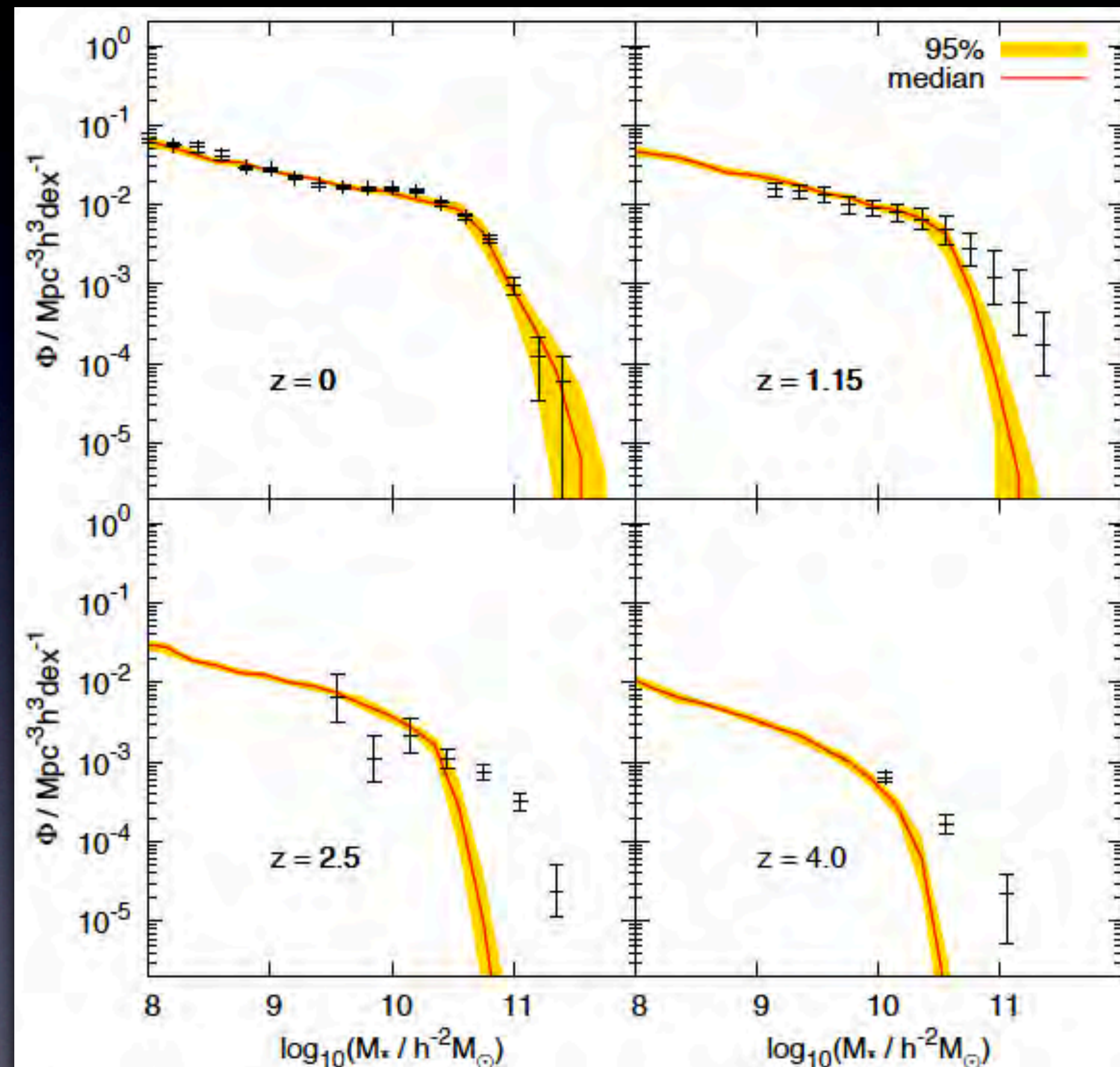
Posterior Distribution of Model Parameters

Model I

Lu et al. 2013 (arXiv:1306.0605)



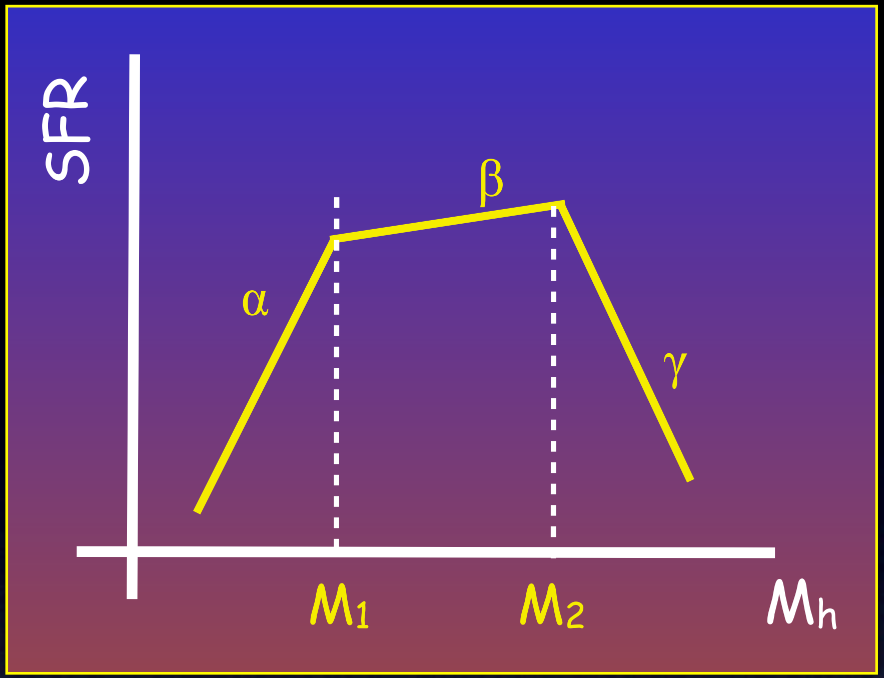
As starting point, we pick a simple model with only 7 free parameter:
$\{\alpha, \beta, \gamma, M_1, M_2, f_{ICL}, \epsilon_{SF}\}$



This model is able to fit stellar mass function at z=0, but fails at higher redshifts....

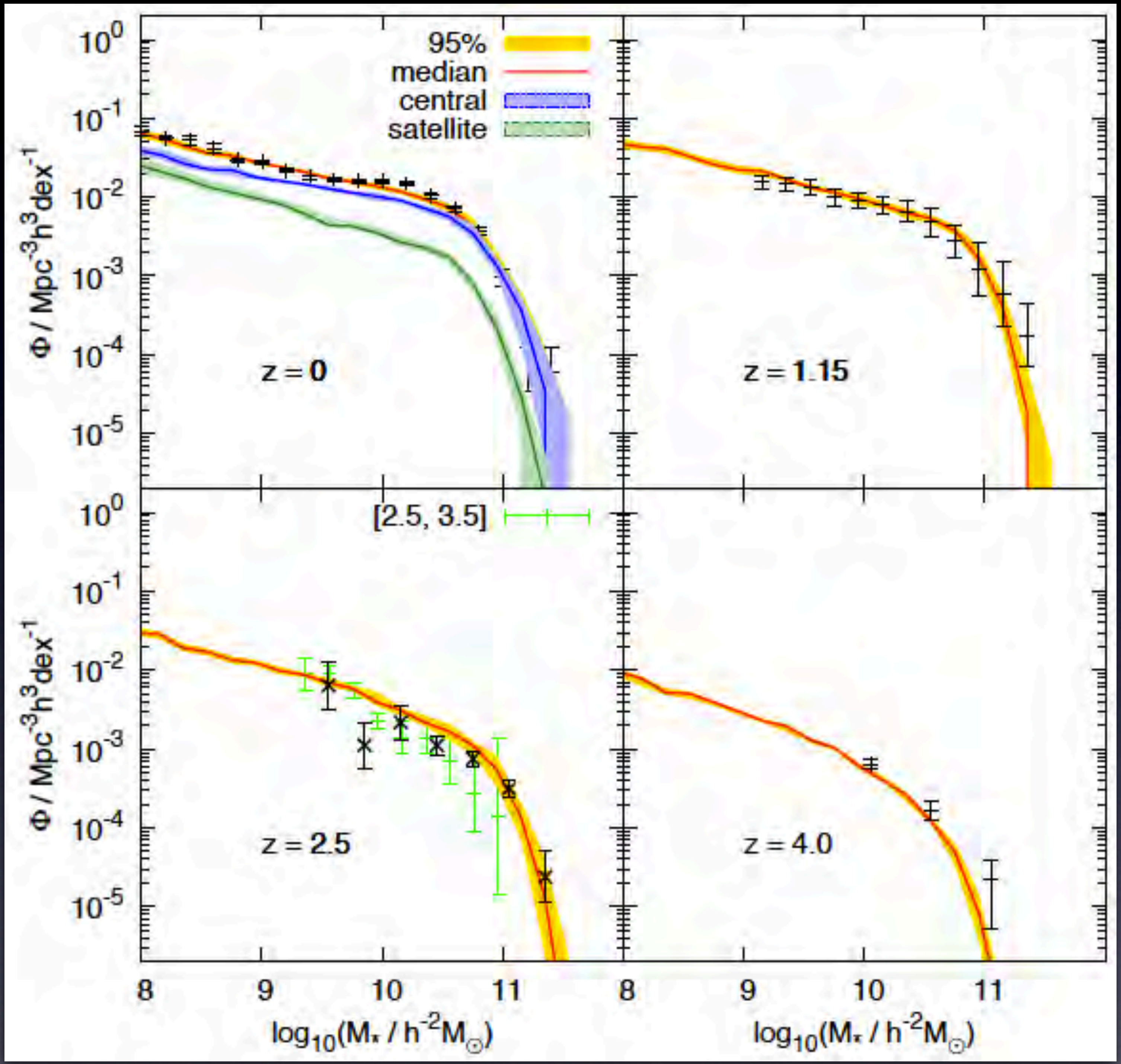
Model II

Lu et al. 2013 (arXiv:1306.0605)



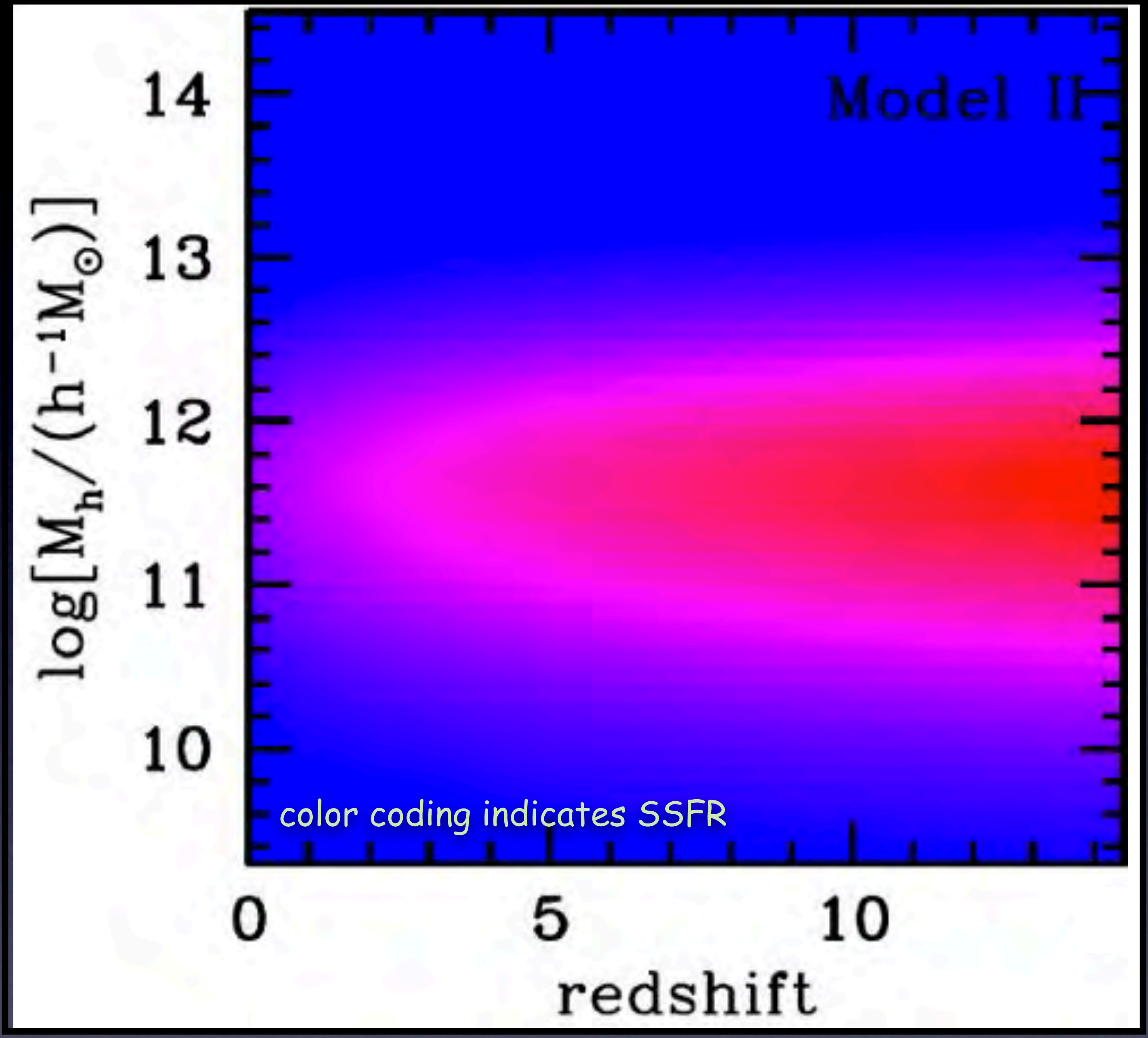
We can solve this problem by adding one additional parameter:

$$\gamma \rightarrow \gamma_0\,(1+z)^c$$



Model accurately fits stellar mass functions out to z=4, and predicts that central galaxies dominate the stellar mass function at z=0 down to at least 10^8 Msun...

Galaxy Formation is Simple



Empirical modeling suggests <u>simplicity</u>.

Star formation occurs mainly in halos with masses in narrow mass range;

$$10^{11} h^{-1} M_\odot < M_\mathrm{h} < M^{12} h^{-1} M_\odot$$

Excellent agreement with a number of similar studies:

Bouche+10, Behroozi+13, Yang+13, Moster+13, Mutch+13

SAMs apparently cannot reproduce this, despite many more free parameters...

Take Home Message 3 (once more)

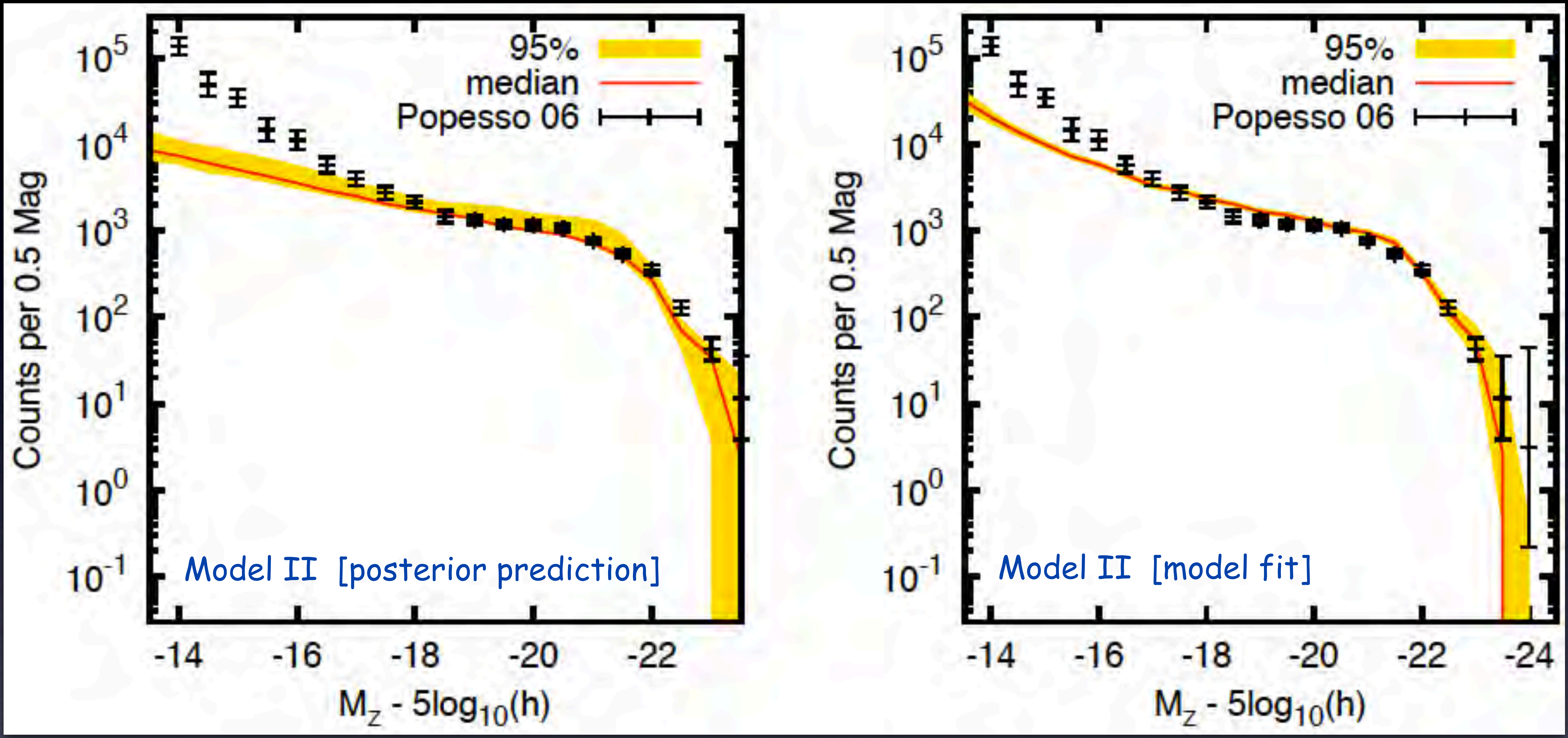
Empirical models can easily fit all available data with only a modest set of free parameters

↓

Are SAMs & SIMs missing relevant physics?

The Cluster Luminosity Function

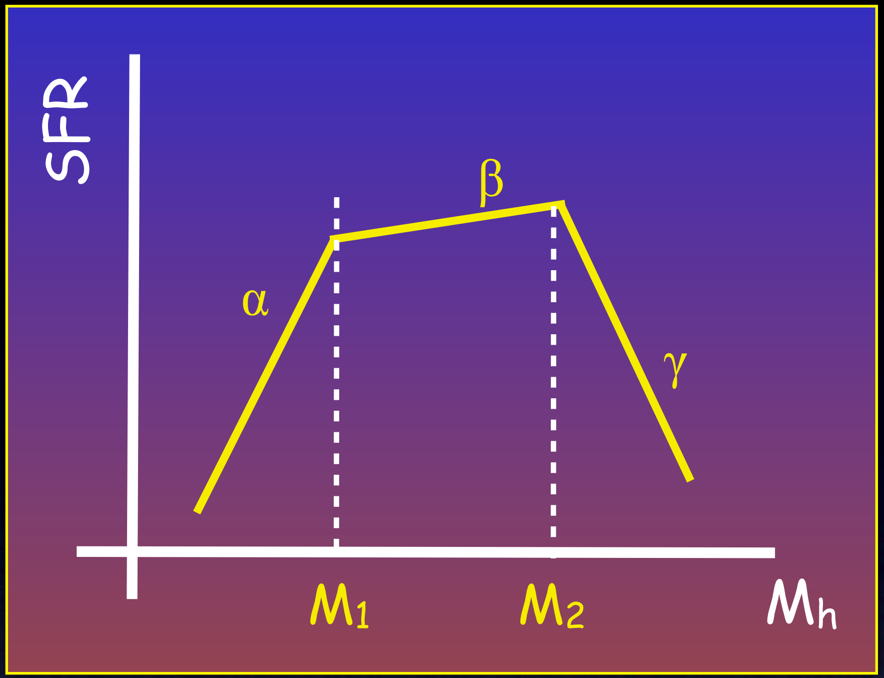
Lu et al. 2013 (arXiv:1306.0605)

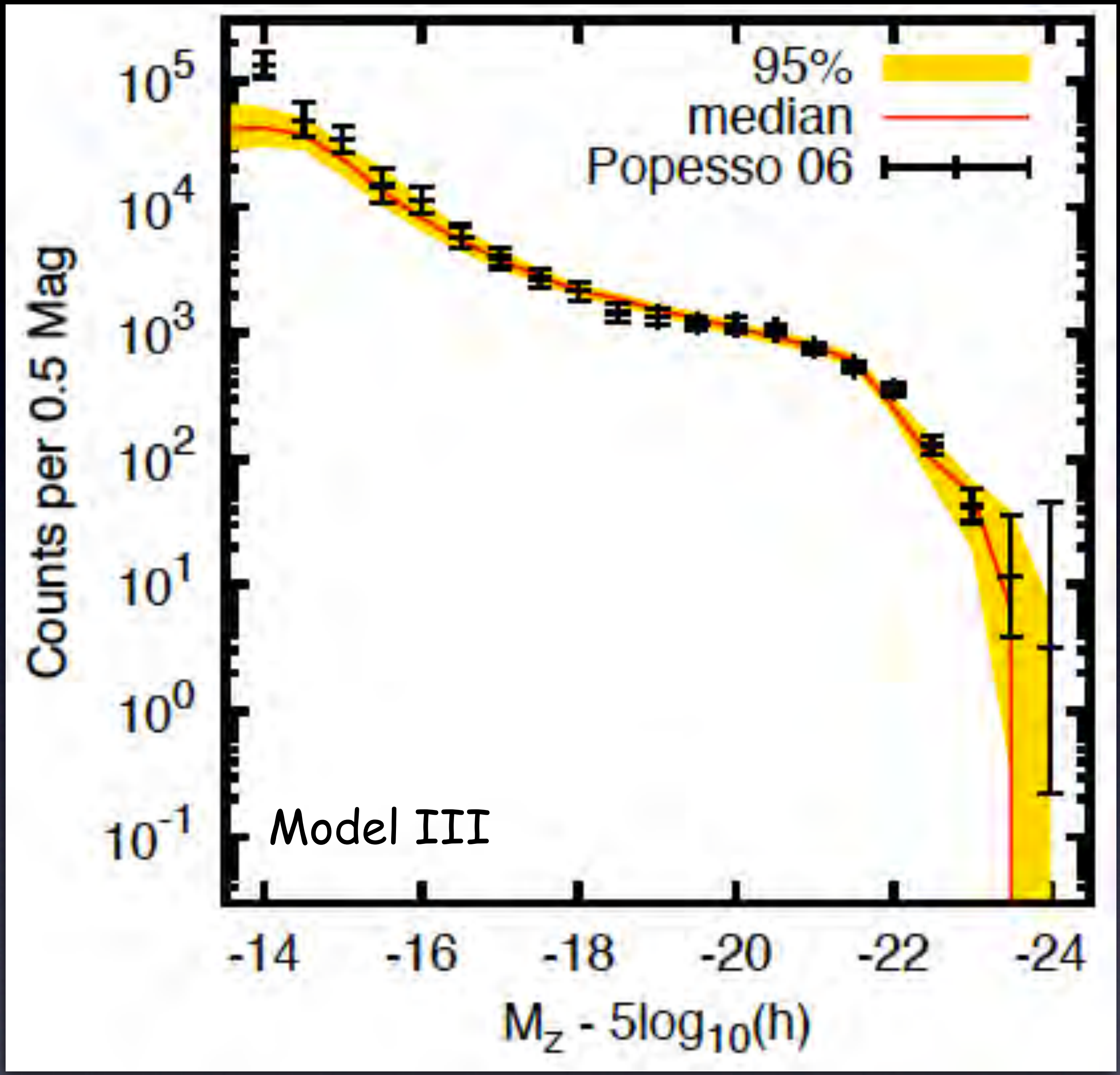


...but, Model II fails to reproduce the steep faint-end slope of the cluster LF...

Model III

Lu et al. 2013 (arXiv:1306.0605)





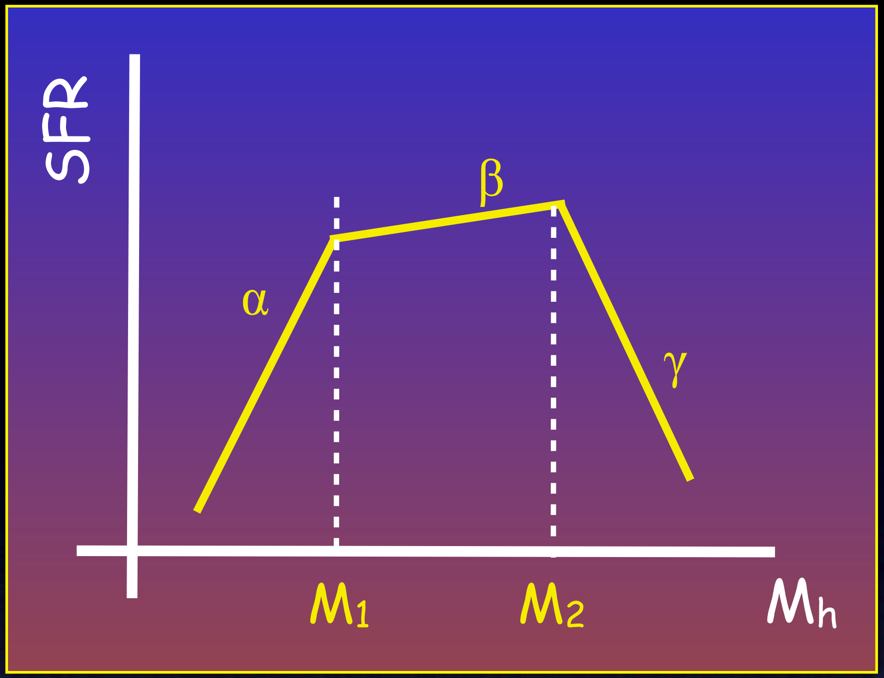
Fitting the cluster LF requires yet more model freedom:

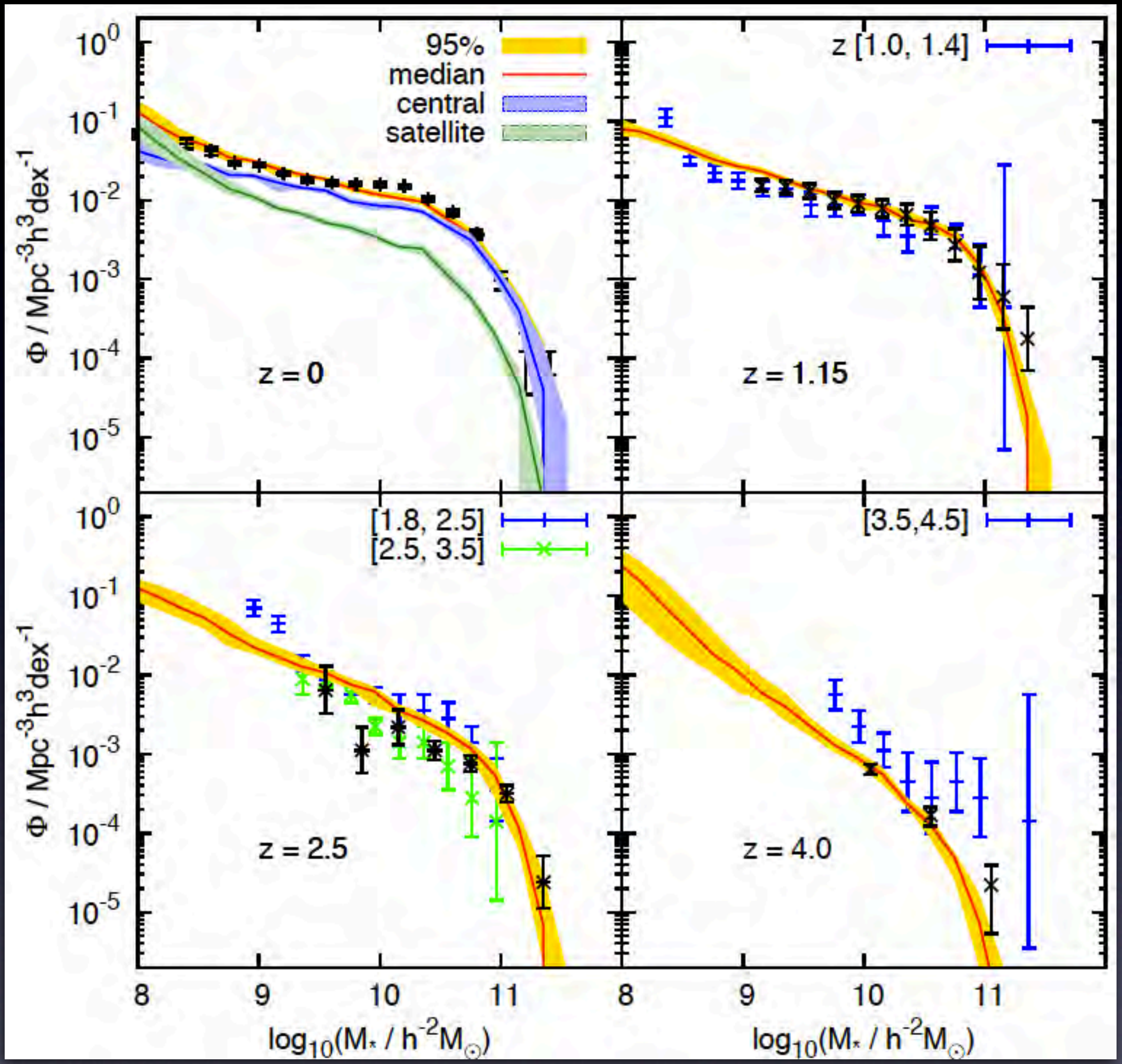
$$\alpha \propto \begin{cases} \alpha_0 & z < z_c \\ (1+z)^a & z > z_c \end{cases}$$

This model is still able to fit the stellar mass functions out to z=4, but predicts a larger fraction of satellites at z=0 at the low mass end...

Model III

Lu et al. 2013 (arXiv:1306.0605)





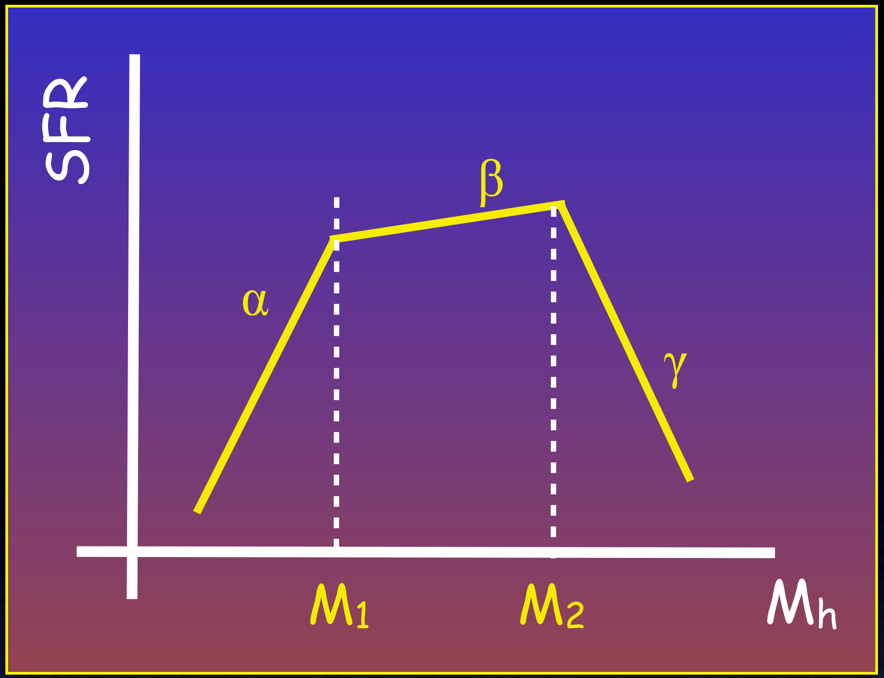
Fitting the cluster LF requires yet more model freedom:

$$\alpha \propto \begin{cases} \alpha_0 & z < z_c \\ (1+z)^a & z > z_c \end{cases}$$

This model is still able to fit the stellar mass functions out to z=4, but predicts a larger fraction of satellites at z=0 at the low mass end...

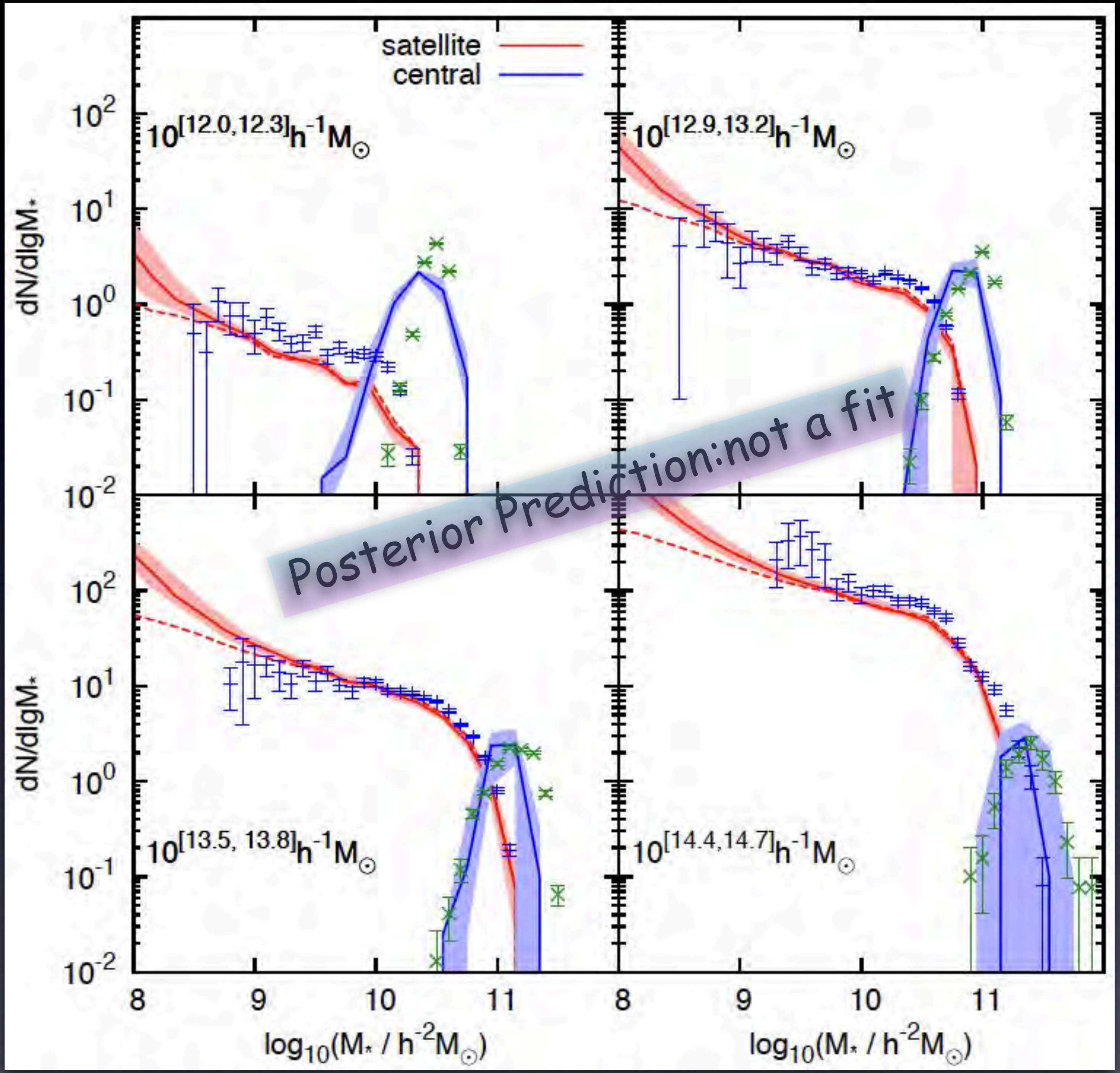
Model III

Lu et al. 2013 (arXiv:1306.0605)



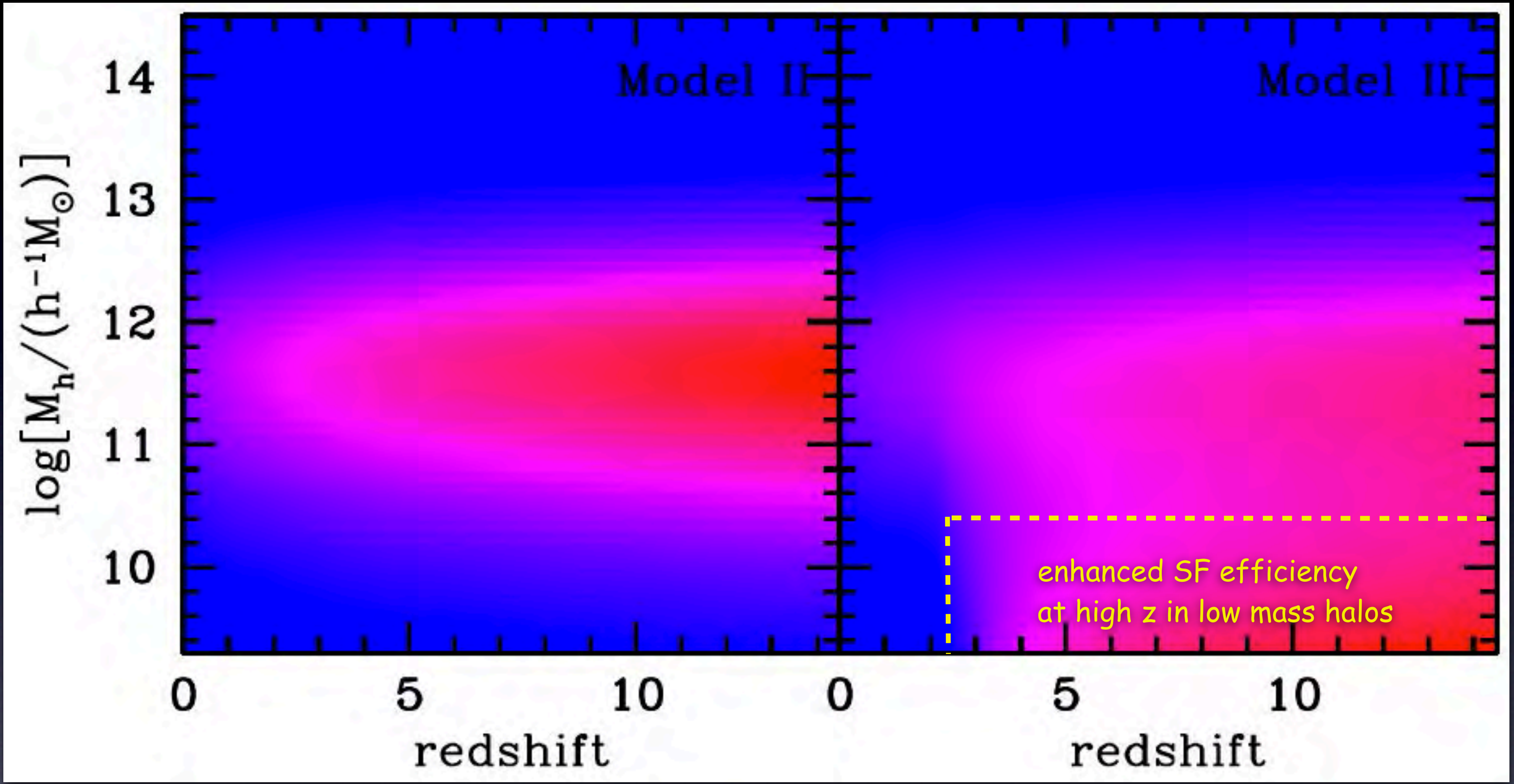
Fitting the cluster LF requires yet more model freedom:

$$\alpha \propto \begin{cases} \alpha_0 & z < z_c \\ (1+z)^a & z > z_c \end{cases}$$



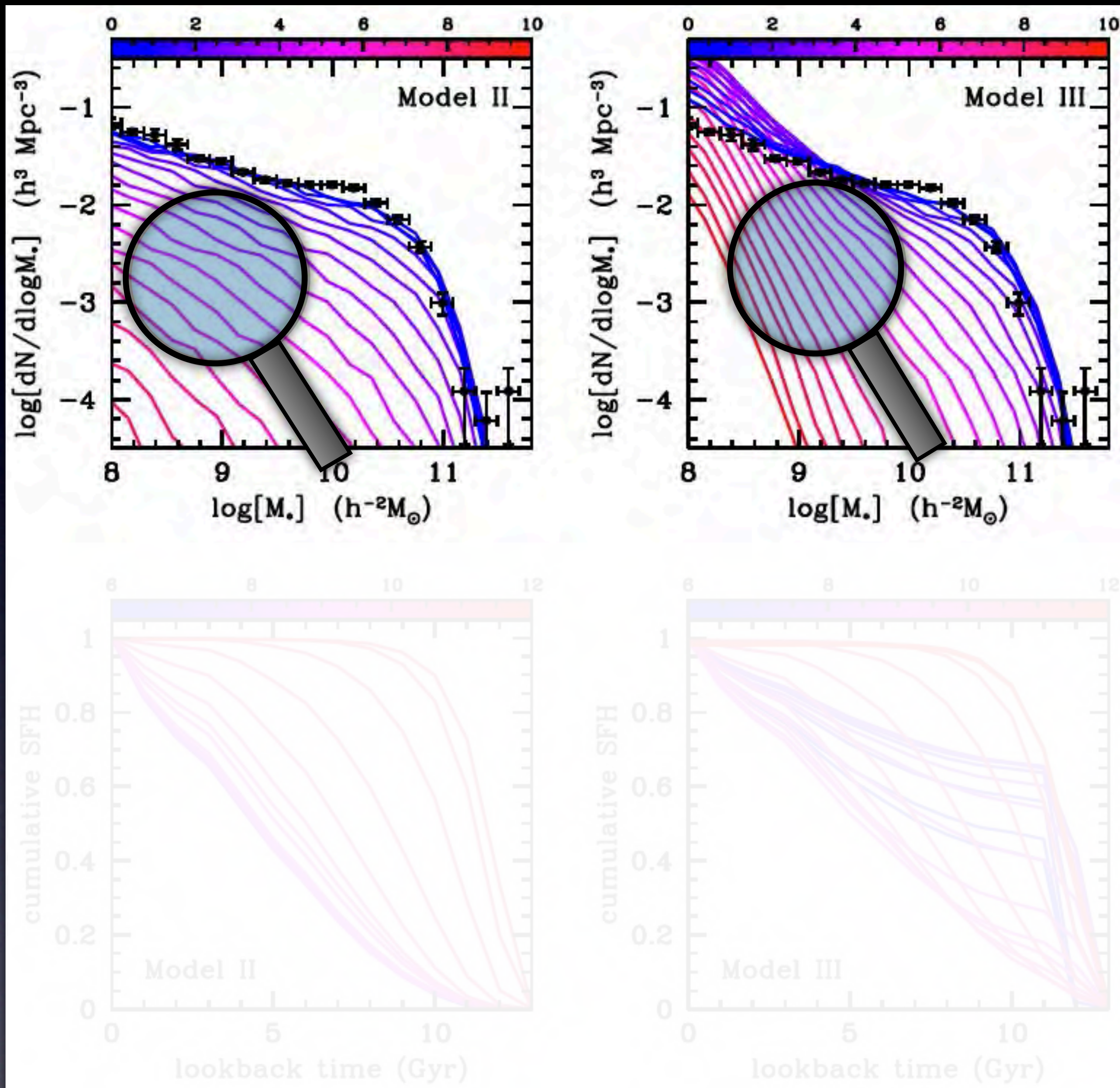
And model also does reasonable job in matching Conditional Stellar Mass functions obtained by Yang, Mo & vdB (2008) using SDSS Galaxy Group Catalogs...

A new Characteristic Scale in Galaxy Formation?



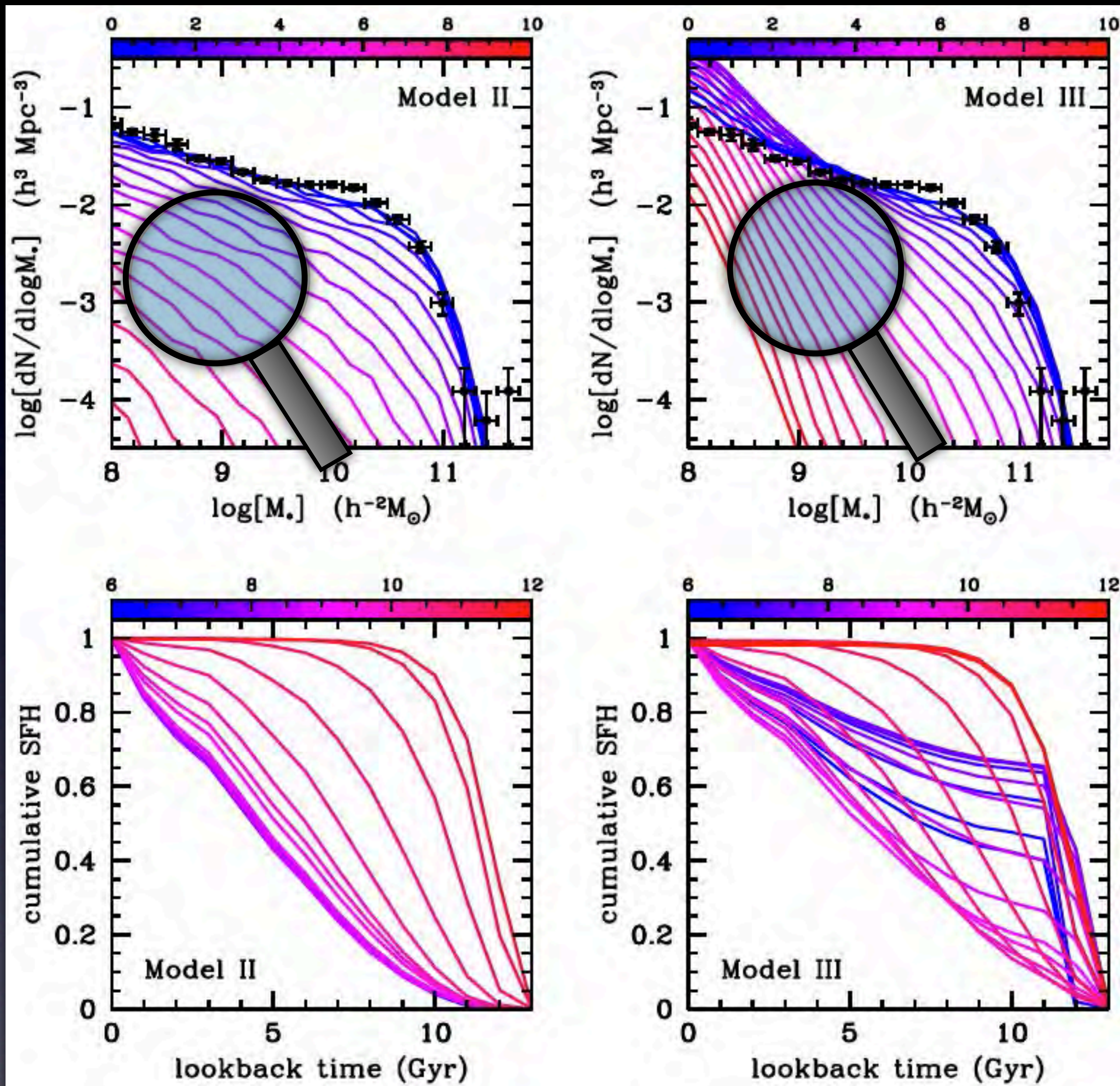
Upturn at faint end in cluster LF requires a boost of SF efficiencies in low mass halos, but only at high redshift ($z > z_c \sim 2$)

How to Distinguish between Models II and III ?

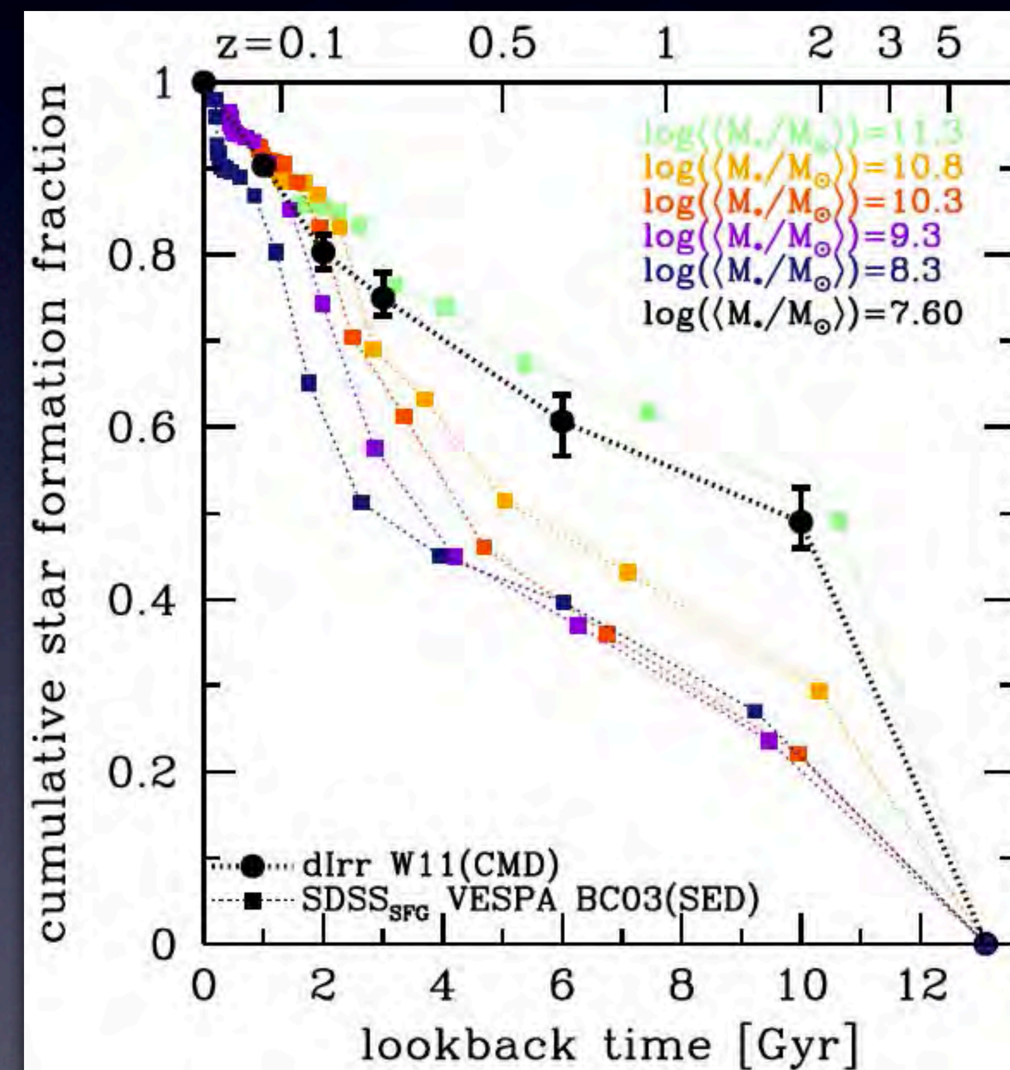


Model III predicts high-z mass functions that are much steeper at the low mass end.

How to Distinguish between Models II and III ?



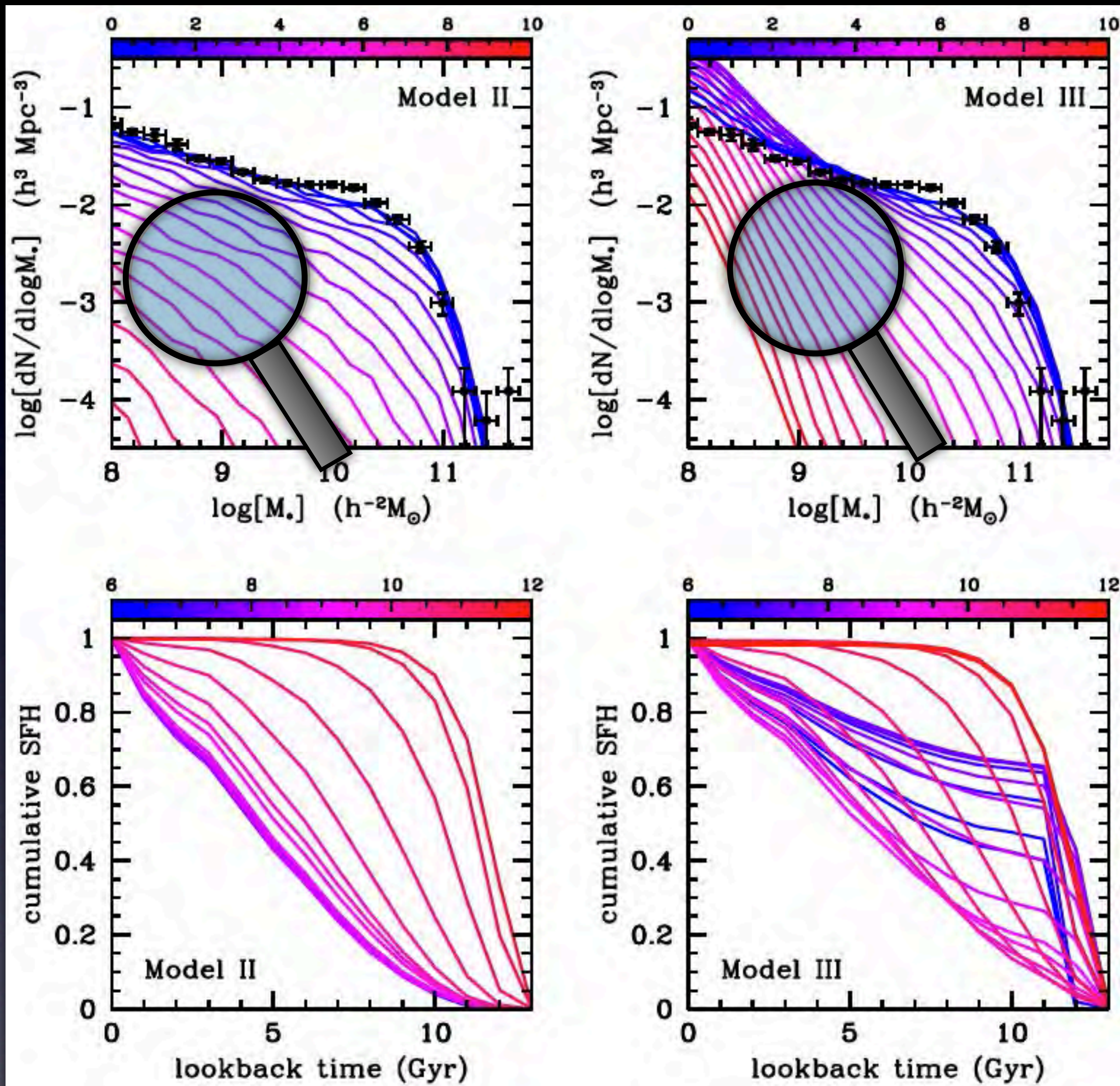
Model III predicts high-z mass functions that are much steeper at the low mass end.

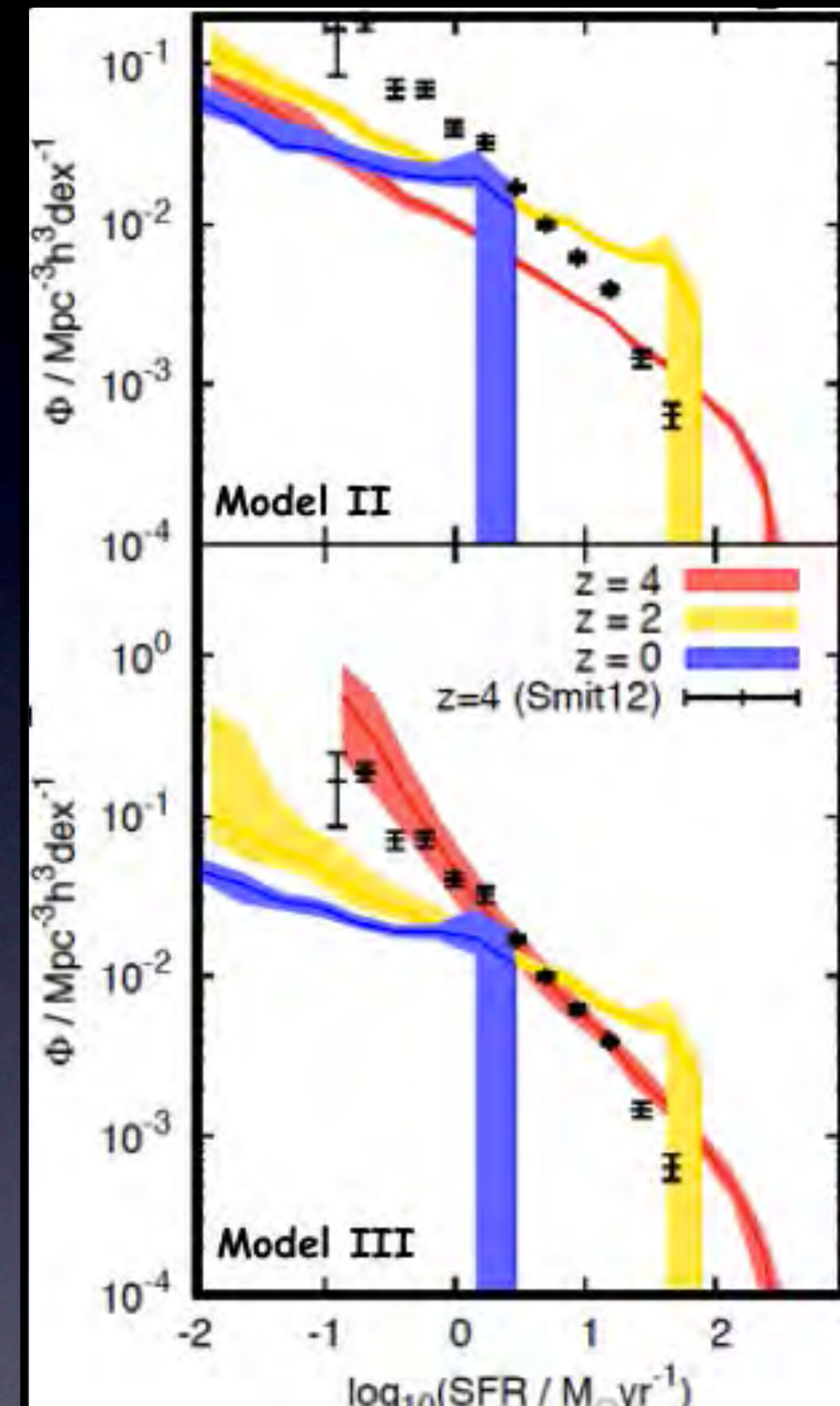


from: Leitner et al. (2012)

Model III also predicts a `break' in the monotonicity of star formation histories. This has observational support from resolved stellar populations!

How to Distinguish between Models II and III ?





and finally, model III also predicts SFR functions at high-z in much better agreement with the data than model II...

Take Home Message 5

Data suggests dramatic change in star formation efficiency in low mass halos around Z~2

new characteristic scale/epoch in galaxy formation

Conclusions

- Due to great advances in data, we now have an accurate, statistical description of the galaxy-dark matter connection.
- Empirical modeling, based on halo occupation models, is able to accurately fit all existing data regarding the abundances of galaxies across cosmic time.
- These models suggest an extremely simple $\dot{M}_*[M_{\rm h}, z]$
- Surprisingly; SAMs, with all their freedom, seem unable to produce such a $\dot{M}_*[M_{\rm h}, z]$; are they missing relevant physics?
- Data on dwarf galaxies suggests a new, characteristic epoch in galaxy formation: star formation becomes strongly suppressed in low mass halos ($M_{\rm h} < 10^{11} h^{-1} M_\odot$) around z~2.
- What is cause of this transition? Preheating by TeV blazars?