Galaxy Formation: Bayesian SAMs Results

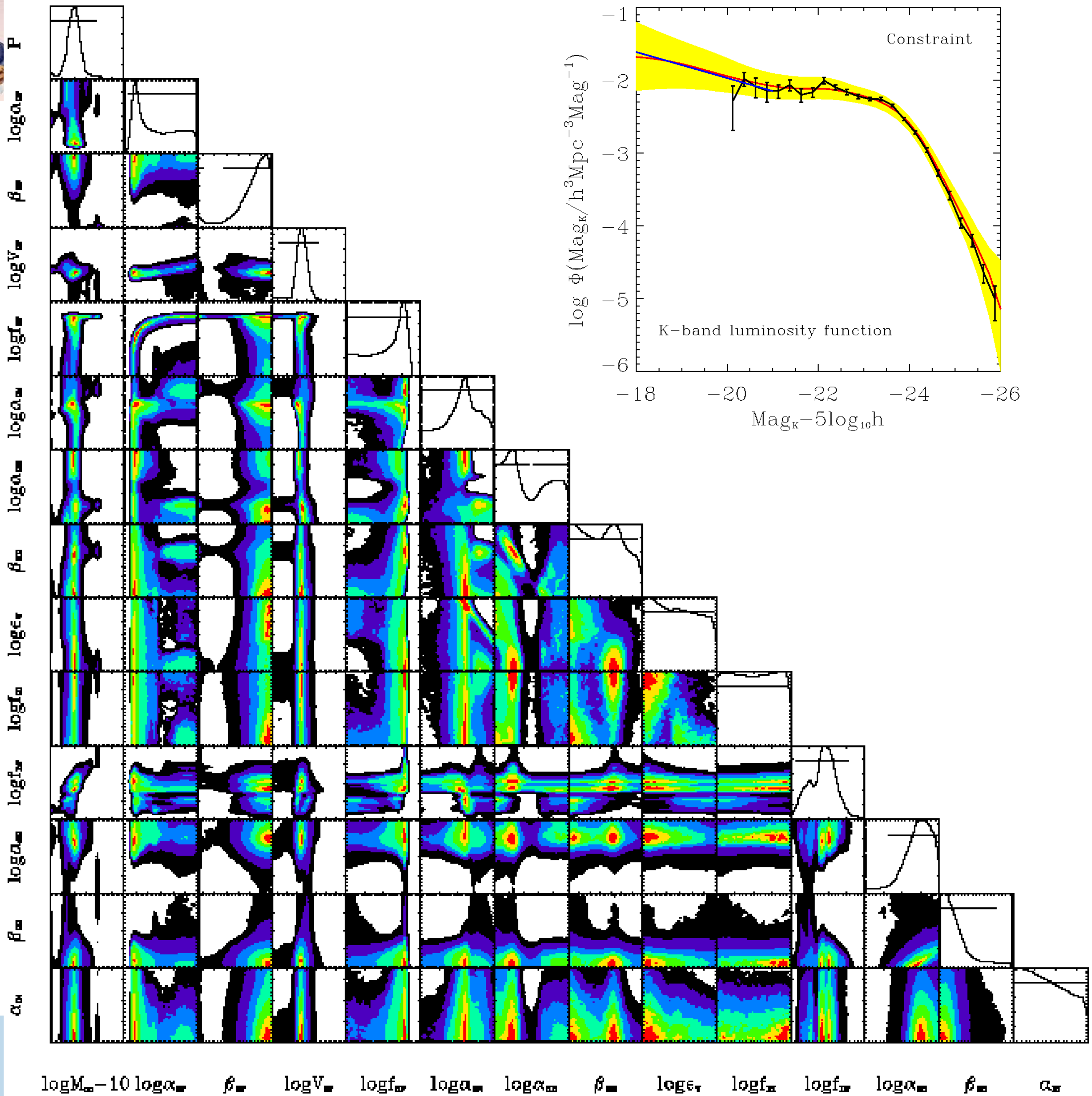
Neal Katz
UMass Astronomy

Collaborators: Zhankui Lu, Yu Lu, Houjun Mo
Martin Weinberg, Frank Van den Bosch

September, 2013

Play it SAM: K-band Constrained

- K-band errors should be diagonal.
- The faint end completeness is not well understood so we parametrize it with an additional parameter and then marginalize over it.
- Again we can fit the data and find other SAMs in some of the posterior modes.
- Some differences with the stellar mass function constrained inference.
  - Now only the mode with M_{CC} about 10^{12} and f_{DF} about one is allowed.
  - Allowed ranges of β_{SF} and β_{RH} have also changed.

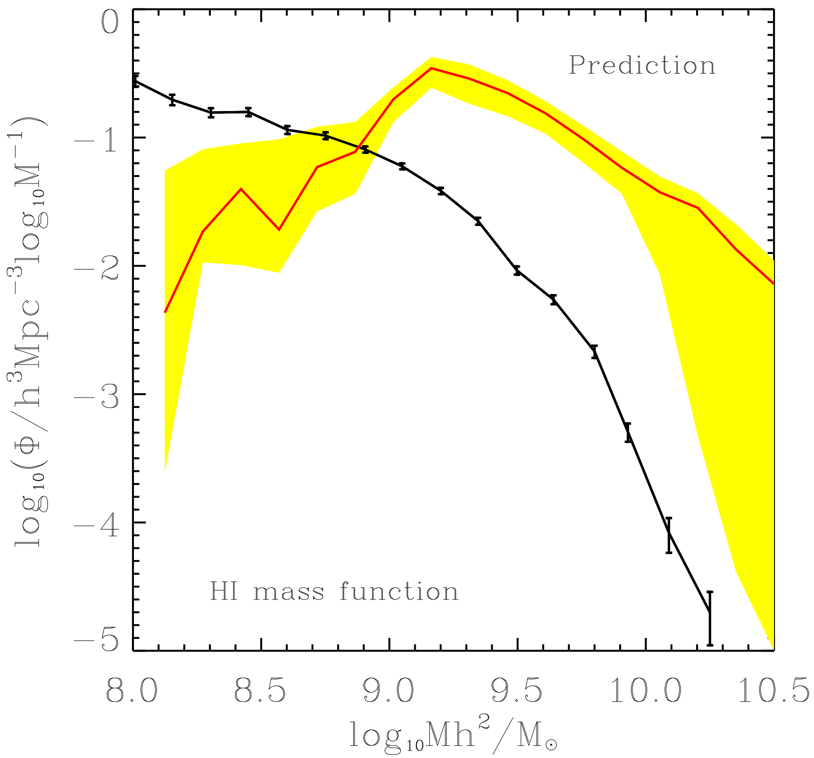


P
$\log a_{\rm sf}$
$\beta_{\rm sf}$
$\log V_{\rm sf}$
$\log f_{\rm df}$
$\log a_{\rm ej}$
$\log \alpha_{\rm cool}$
$\beta_{\rm ej}$
$\log \epsilon_{\rm r}$
$\log f_{\rm tr}$
$\log f_{\rm bf}$
$\log \alpha_{\rm ret}$
$\beta_{\rm ret}$
$\alpha_{\rm ff}$
$\log M_{\rm cool}-10$
Constraint
K-band luminosity function
$\log\ \Phi({\rm Mag_K}/h^3{\rm Mpc}^{-3}{\rm Mag}^{-1})$
${\rm Mag_K}-5\log_{10}h$

Making predictions

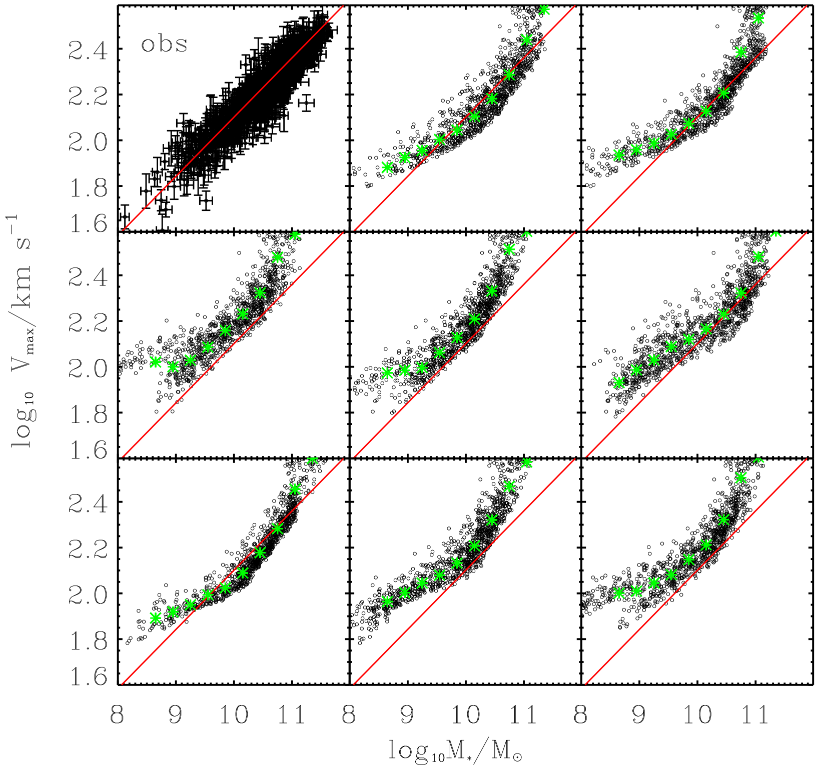
- Use the whole posterior to make predictions of other observables.
- Can use Posterior Predictive Check (PPC) using Principle Component Analysis (PCA) to statistically look for consistency with the data and the predictions.
- A lack of fit does not necessarily mean the model cannot fit both data sets.
  - Need to do an inference using both data sets as constraints.
  - Then use PPC to check for the goodness of fit to both data sets.

Predicted HI Mass Function



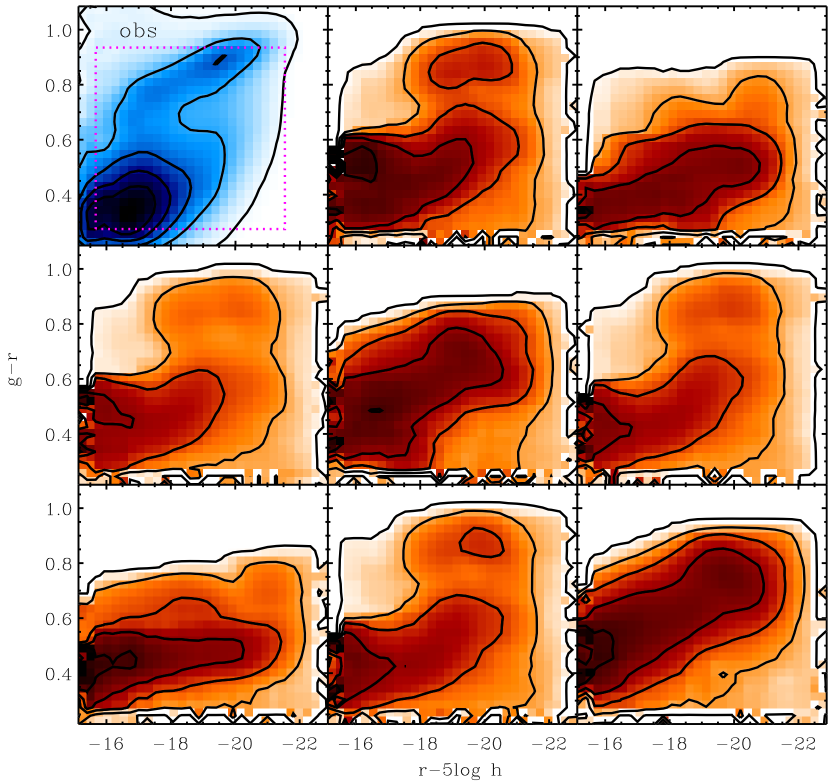
- Too high everywhere,
  - Turn down at small masses owes to mass resolution effects.
- Perhaps need to include preheating.

Predicted Tully-Fisher Relation



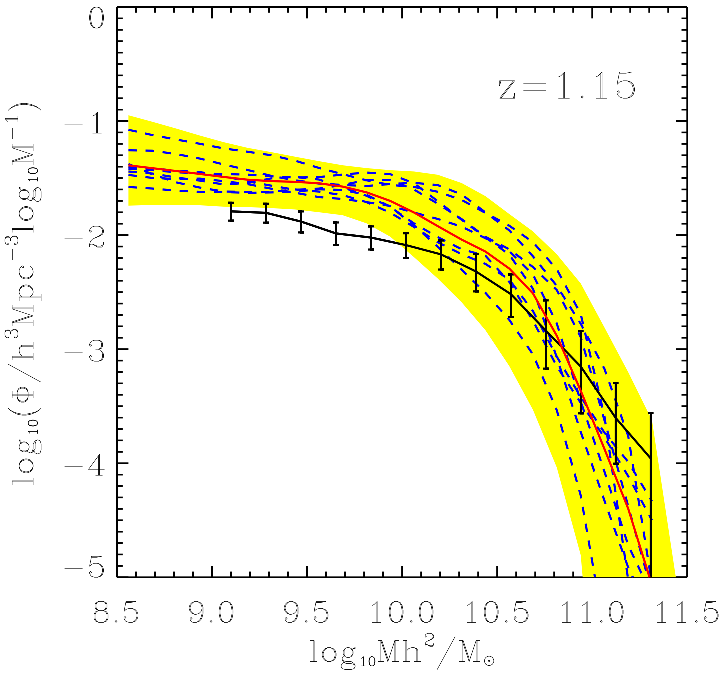
- Amplitude and shape are both wrong.
- Predictions can vary in the allowed parameter range.

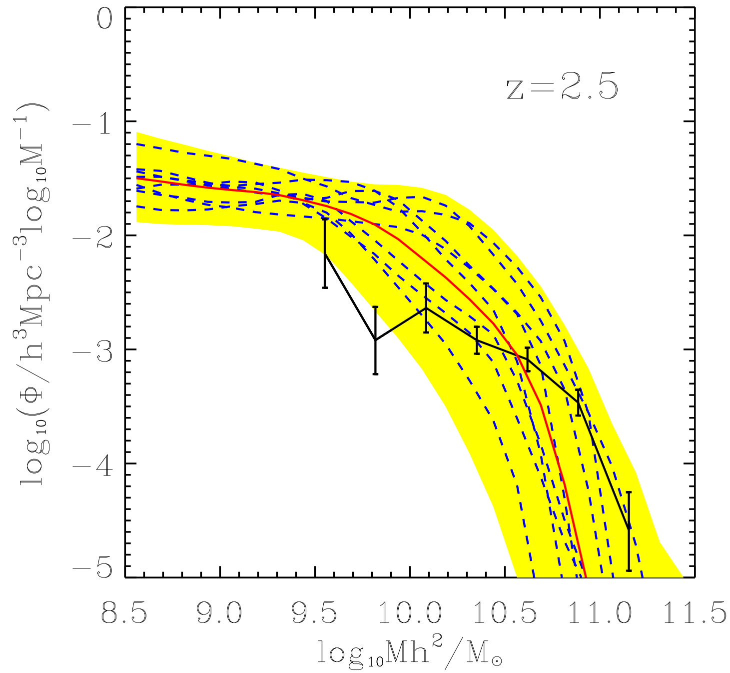
Predicted Color-Magnitude Diagram



- Sometimes it is bimodal, often it is not.
- Predictions can vary in the allowed parameter range.

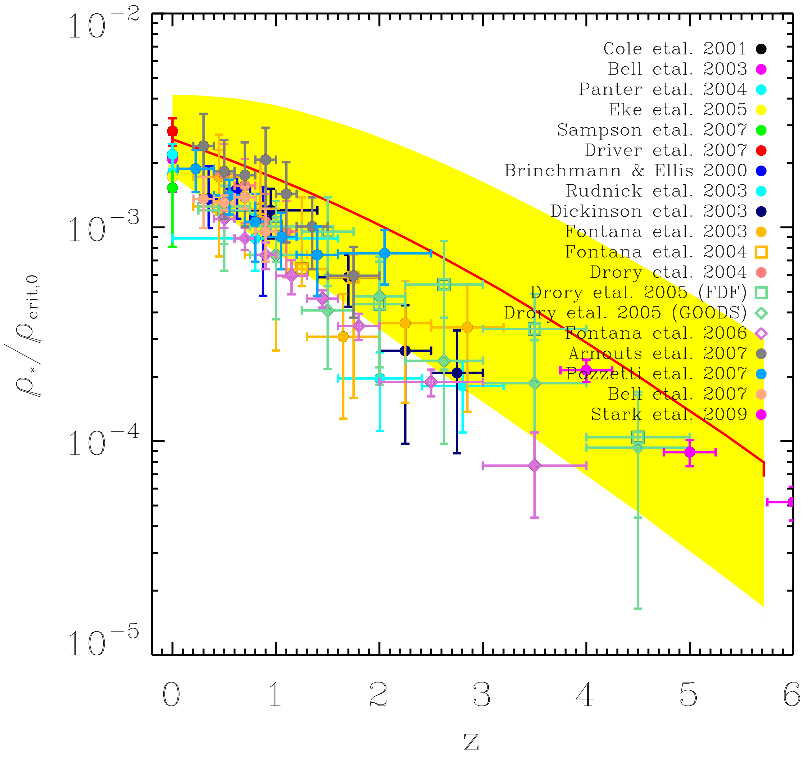
Predicted Stellar Mass Functions at High Redshift



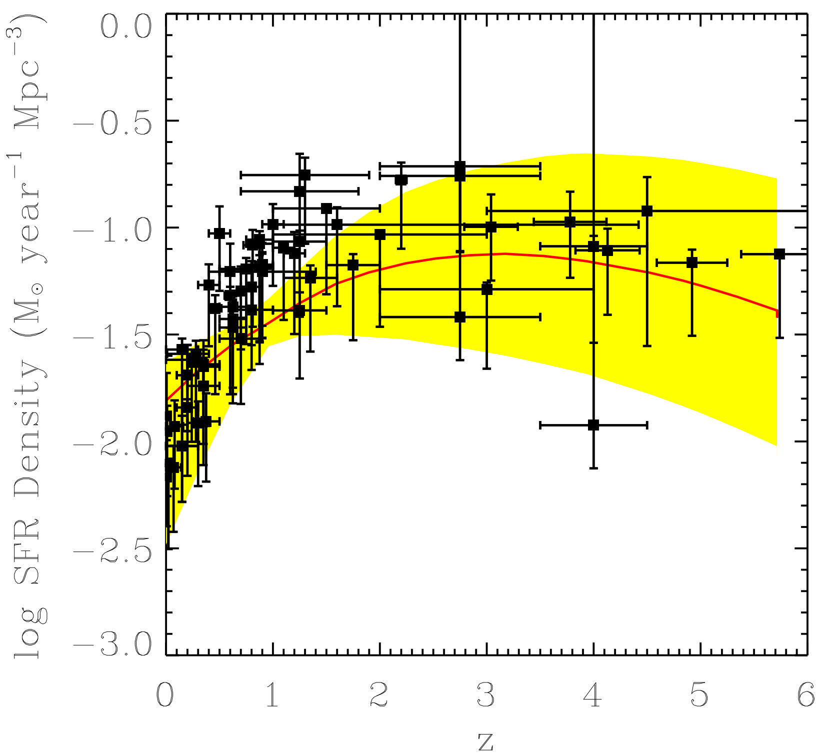


- Shape is wrong.
- Does better if one uses a cooling model that explicitly includes cold mode accretion.

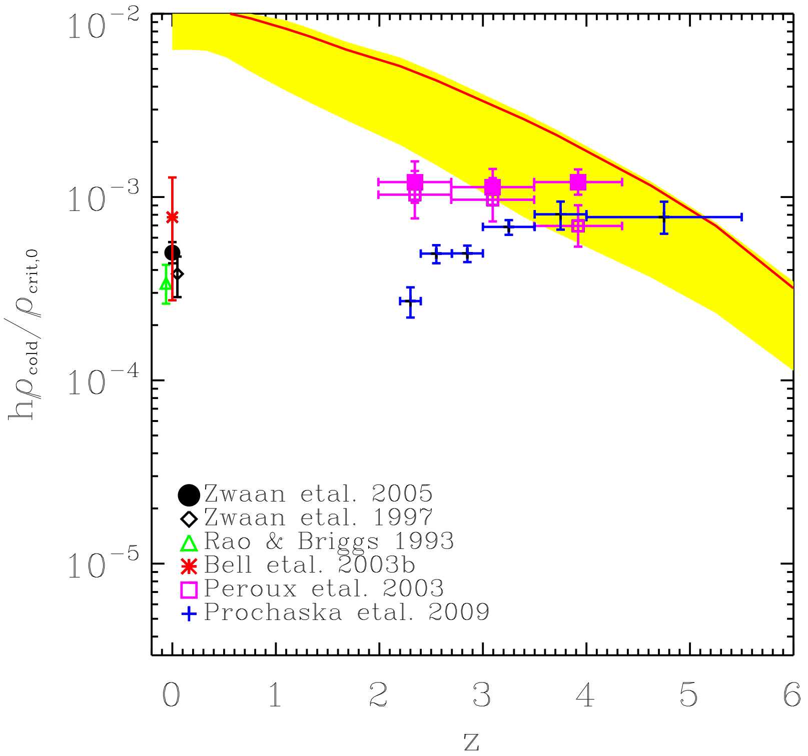
Predicted Stellar Mass Density Evolution



Predicted Global Star Formation History



Predicted Cold Gas Mass Density Evolution



A New Approach

- How does one learn about galaxy formation within the LCDM paradigm?

A New Approach

- How does one learn about galaxy formation within the LCDM paradigm?
- Usual approach using Hydro simulations or SAMs.
  - Make ab initio models of galaxy formation including all the physical processes that one thinks are important.
  - Make predictions from these models and compare them with observations.
  - Change the model to improve any deficiencies.

The Approach

- Our approach.
  - Start with the observations.
  - See in general terms what the observations require of the galaxy formation models.
  - Put this on a firm statistical footing using Bayesian Inference.

The Approach

- Our approach.
  - Start with the observations.
  - See in general terms what the observations require of the galaxy formation models.
  - Put this on a firm statistical footing using Bayesian Inference.
- Similar in flavor to models by Behroozi et al 2012, Moster et al 2013, Yang et al 2013, & Bethermin et al 2013 but improved.
  - We follow galaxies from one time to another instead of just matching abundances at different times.
  - We increase the complexity of the model as required by the data assessed using Bayes ratios.

A Simple SAM

- Use dark halo merger trees.
- Subsume most baryonic physics by assuming for central galaxies

$$\dot{M}_{\star} = \dot{M}_{\star}(M_{\rm vir}(z), z).$$

- Assume star formation in satellite galaxies is quenched exponentially with a timescale that depends on the stellar mass (two free parameters).
- Galaxy mergers and stripping are treated as in other SAMs (one free parameter).

A Simple Model for the Central Galaxy SFR

$$\dot{M}_\star = \mathcal{E}\frac{f_B M_{\rm vir}}{\tau_0}(1+z)^{3/2}(X+1)^\alpha\left(\frac{X+\mathcal{R}}{X+1}\right)^\beta\left(\frac{X}{X+\mathcal{R}}\right)^\gamma.$$

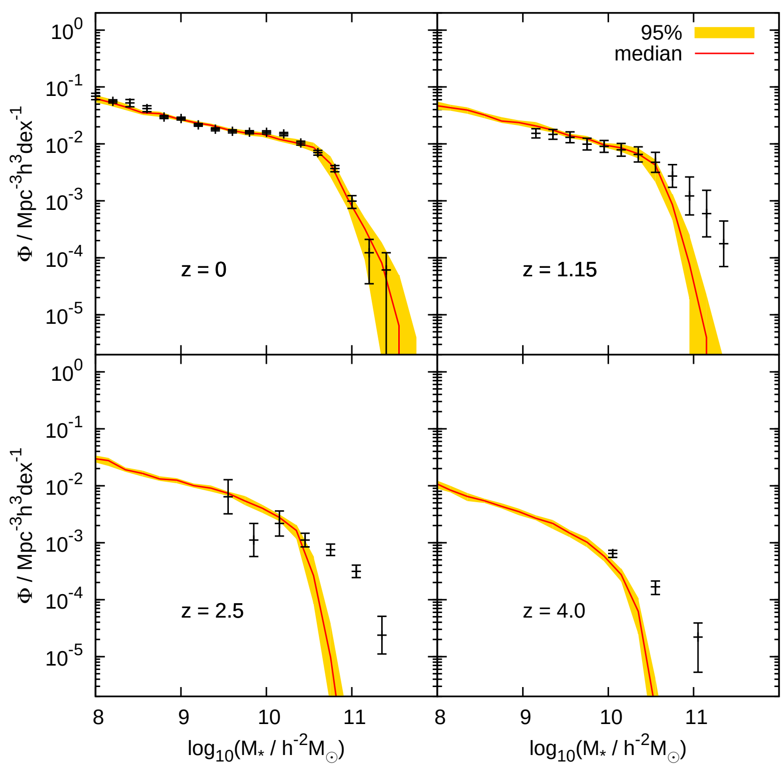
- $\mathcal{E}$ is an overall efficiency.
- f_B is the cosmic baryon mass fraction (fixed).
- τ_0 is a present dynamic timescale of the halos (fixed).
  - $\tau_0 \equiv 1/(10H_0)$.
- Two characteristic masses: $M_{\rm c}$ ($X \equiv M_{\rm vir}/M_{\rm c}$) and $\mathcal{R}M_{\rm c}$ with $\mathcal{R} < 1$.
- The two masses divide 3 power laws with slopes α, β, and γ:

$$\dot{M}_\star \propto \frac{M_{\rm vir}}{\tau_0}\begin{cases} M_{\rm vir}^\alpha & \text{if } M_{\rm vir} \gg M_{\rm c} \\ M_{\rm vir}^\beta & \text{if } M_{\rm c} > M_{\rm vir} > \mathcal{R}M_{\rm c} \\ M_{\rm vir}^\gamma & \text{if } M_{\rm vir} \ll \mathcal{R}M_{\rm c}\,. \end{cases}$$

Modi Operandi

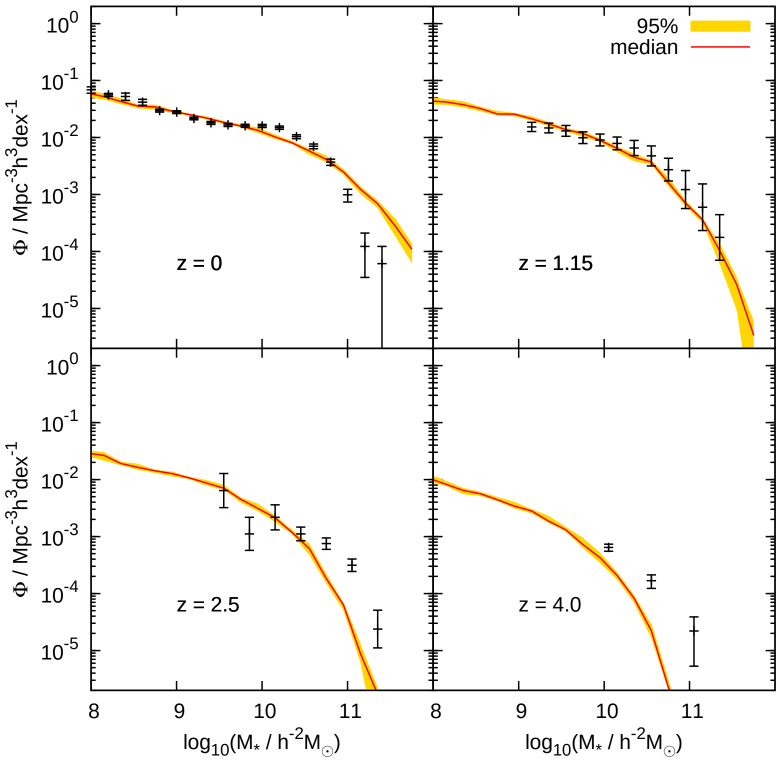
- Use Bayesian Inference to determine the Posterior Distribution of the parameters given the constraining data.
- Use one of three observational data sets as constraints:
  - 1. The $z = 0$ galaxy stellar mass function (Baldry et al 2012).
  - 2. Above plus the $z = 1.15$, $z = 2.5$, and $z = 4.0$ galaxy stellar mass functions.
  - 3. Above plus the low redshift cluster galaxy luminosity function (Popesso et al 2006).
- Increase the complexity of the model by allowing some parameters to be redshift dependent.
- Use Bayes Ratios to decide whether or not the observations require the increased complexity.

Model I



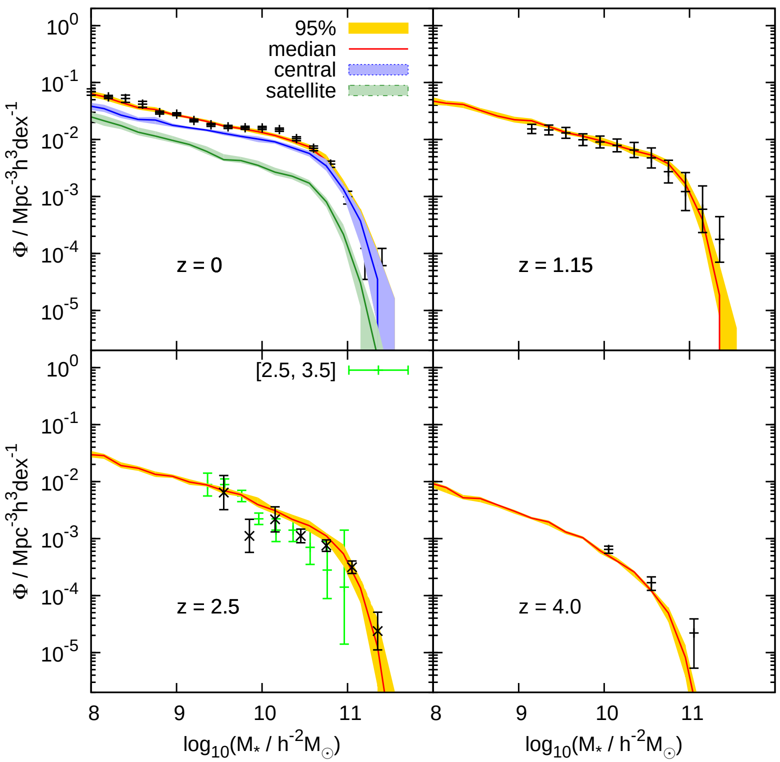
- All parameters are redshift independent.
- Use $z = 0$ galaxy stellar mass function as data constraint (Data 1).
- Matches at $z = 0$ but gets the massive end wrong at higher redshifts.
- Fit using only a nine parameter model.

Model I: Redux



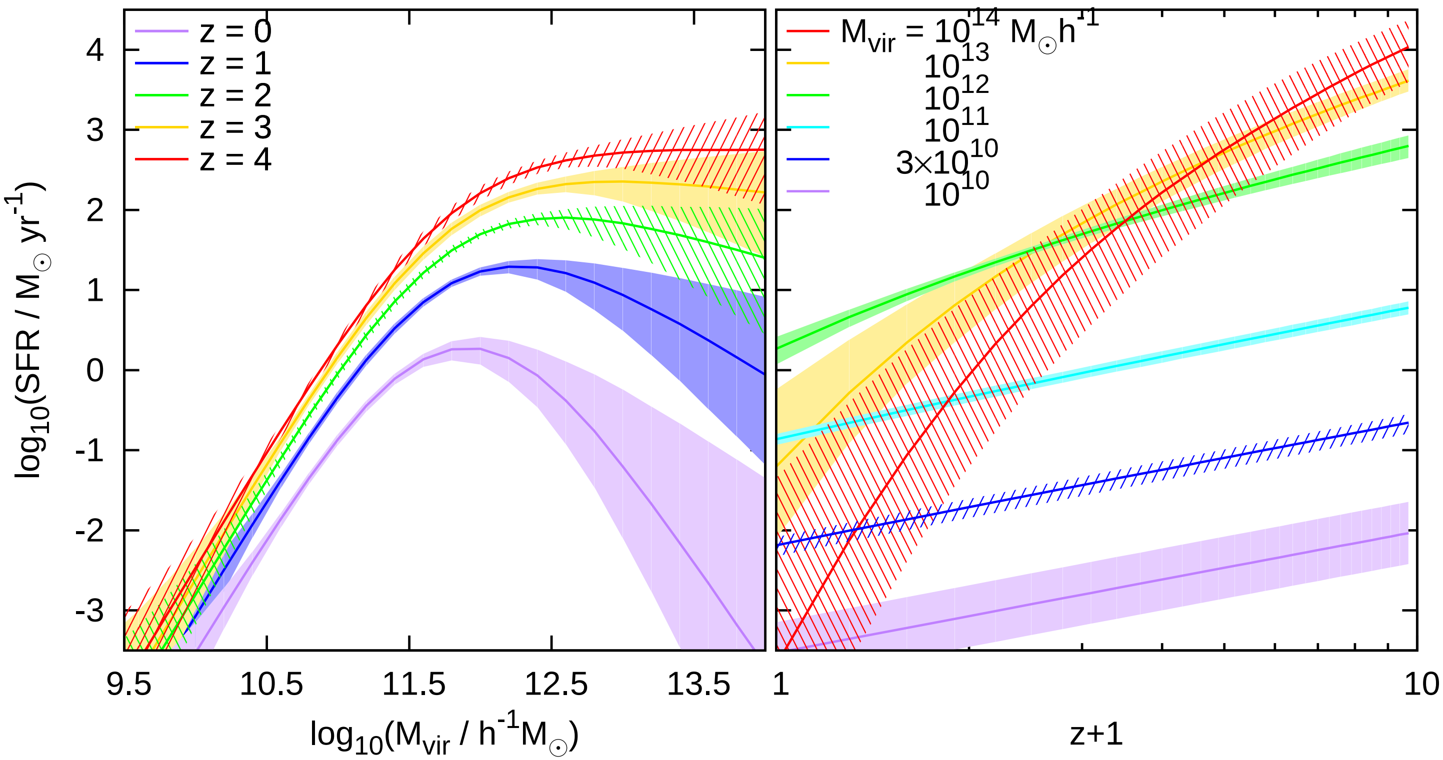
- All parameters are still redshift independent.
- Use GSMFs from four redshifts as data constraints (Data 2).
- Worse at $z = 0$ and now misses the massive end at all redshifts.
- Need the SFR at the massive end to be redshift dependent.

Model II



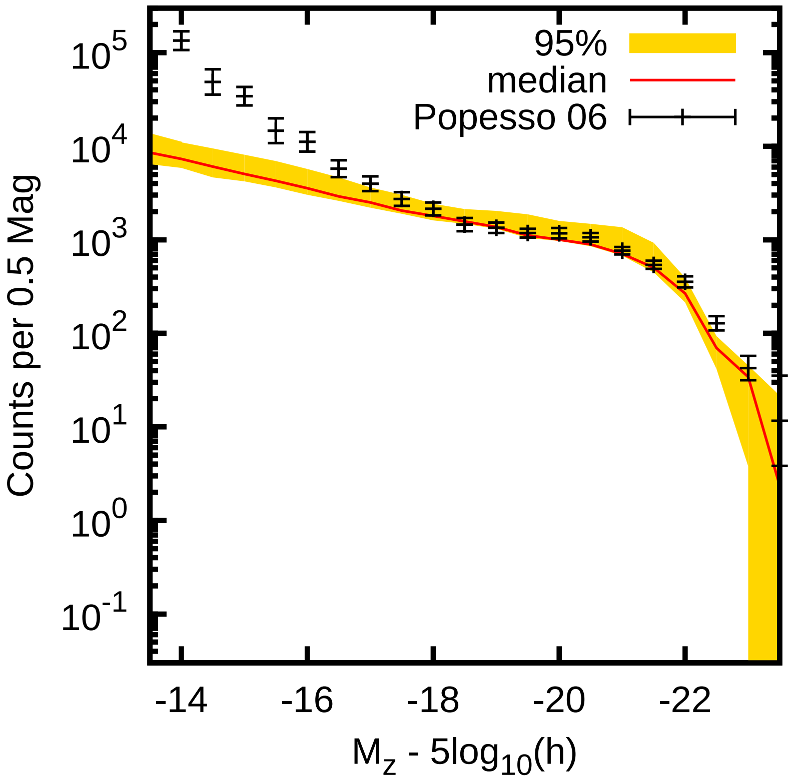
- Allow α to be redshift dependent: $\alpha = \alpha_0(1+z)^{\alpha'}$.
- Use GSMFs from four redshifts as data constraints (Data 2).
- Matches at all redshifts.
- Preferred over Model I by over 10^{38} to one!

Model II: SFRs



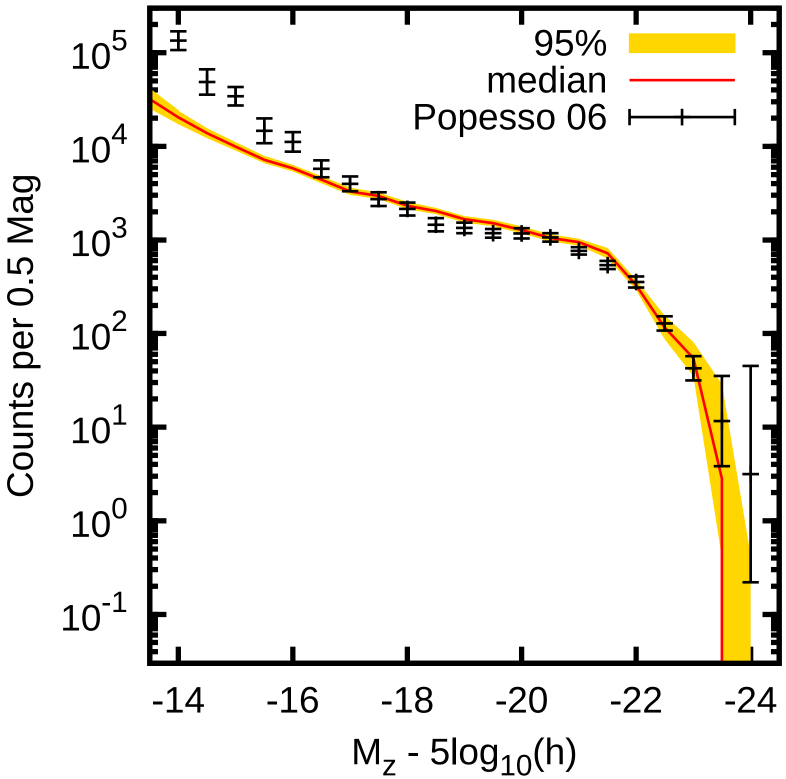
- The star formation rate peaks at around $10^{12}\,h^{-1}\mathrm{M}_{\odot}$.
- In low mass halos SFR increases with halo mass as $\sim M_{\mathrm{vir}}^{2.5}$.
- For high mass halos SFR increases as $(1+z)^{2.3}$.
- For low mass halos SFR increases as $(1+z)^{1.5}$.

Model II: Clusters



- Model II constrained by Data 2 fails to match the cluster galaxy luminosity function data.

Model II: Clusters



- Even when the cluster data is added as a constraint (Data 3), Model II fails to match the cluster galaxy luminosity function data.

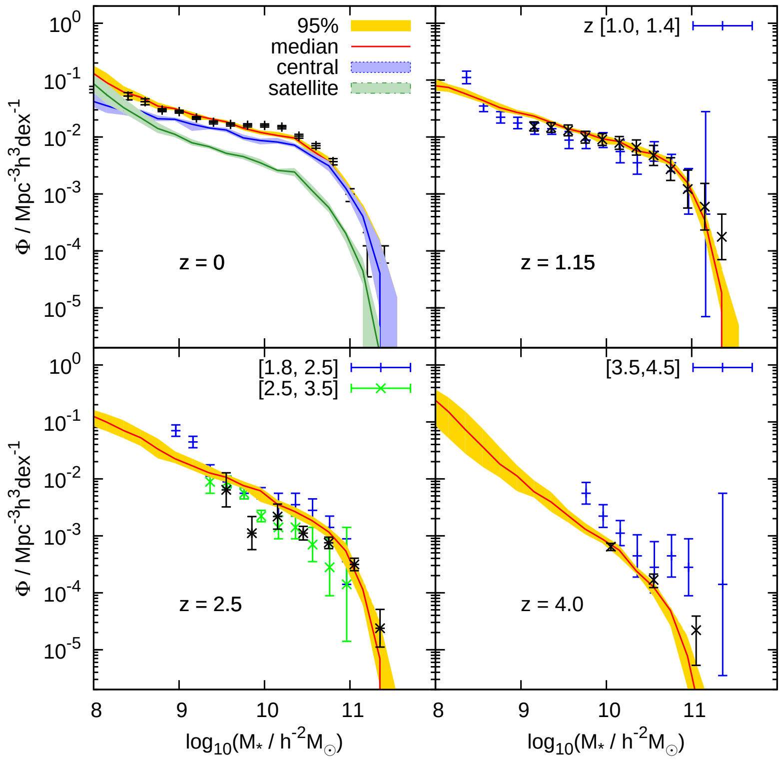
Model III: Requirements

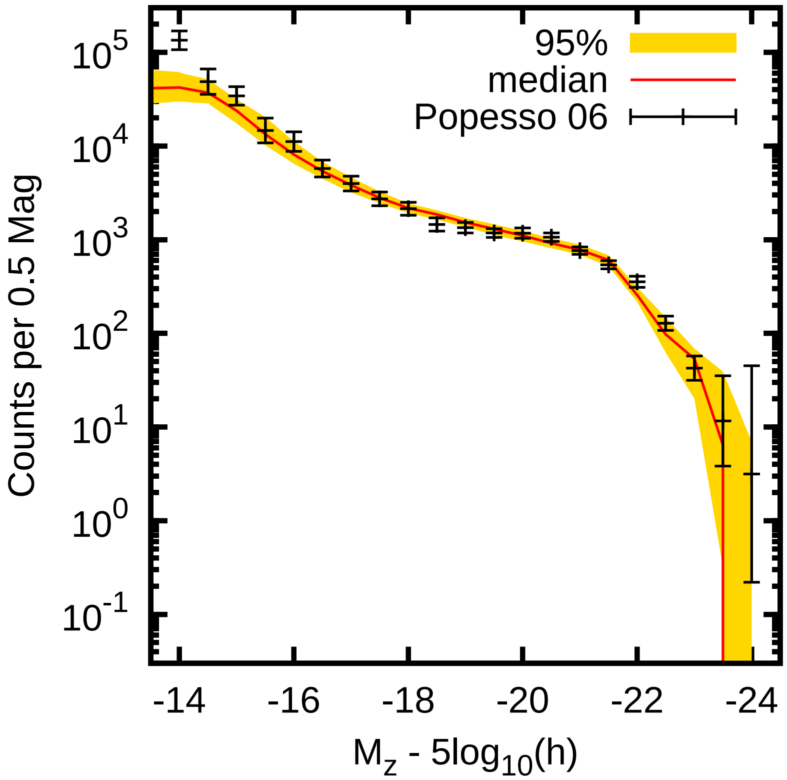
- Dwarf galaxies in clusters formed at high redshift and then accreted.
- Need to affect the formation of dwarfs at high redshift.
- Allow γ to be redshift dependent:

$$\gamma = \begin{cases} \gamma_{\mathrm{a}} & \text{if } z < z_c \\ (\gamma_{\mathrm{a}} - \gamma_{\mathrm{b}}) \left(\frac{z+1}{z_c+1}\right)^{\gamma'} + \gamma_{\mathrm{b}} & \text{otherwise} \,. \end{cases}$$

  - Introduced a critical redshift, z_c, where the slope changes.

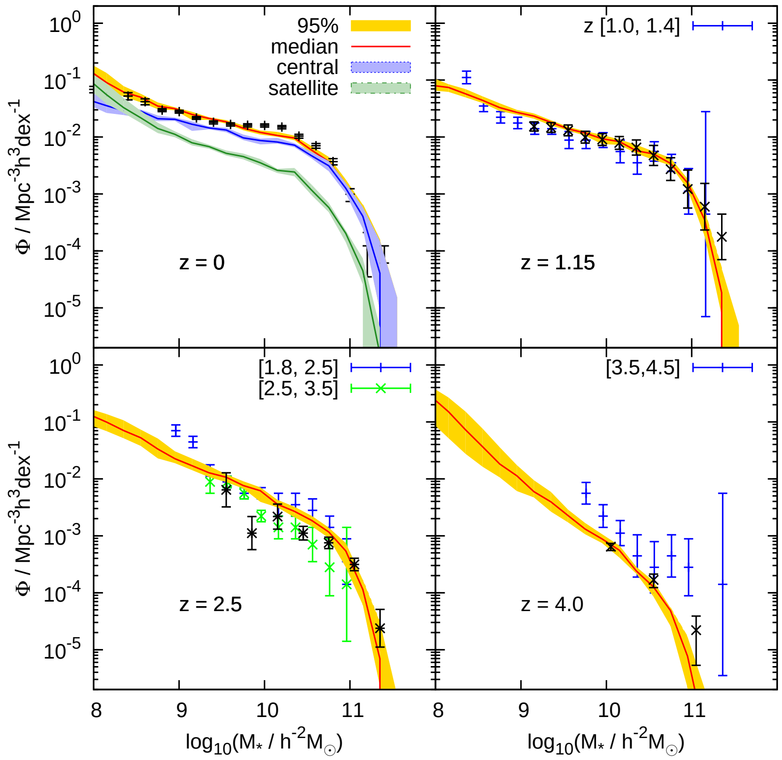
Model III

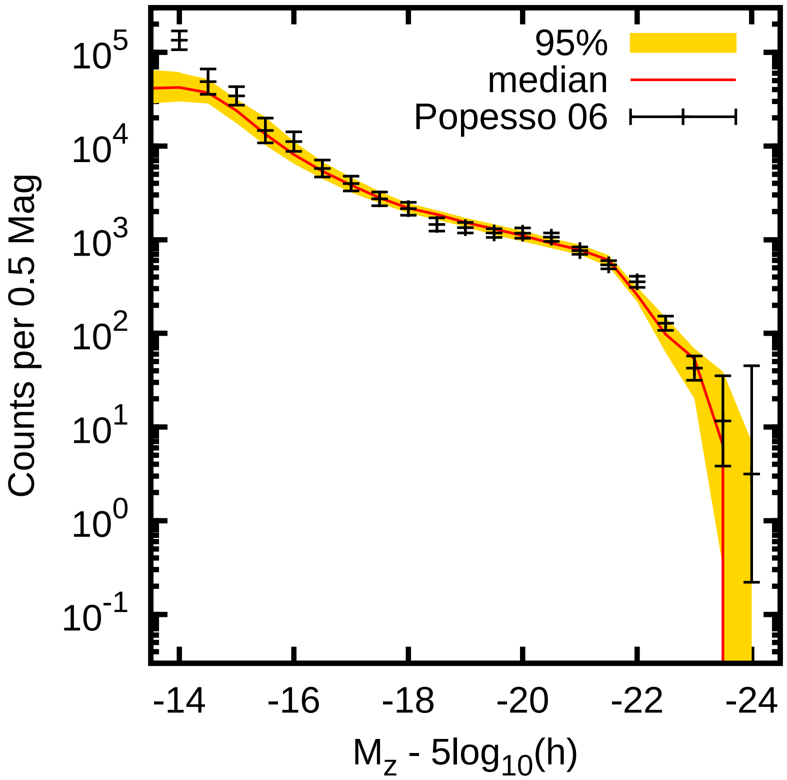




- Use all the observations as data constraints (Data 3).
- Matches all the constraining data.
- Preferred over Model II by over 10^{11} to one!

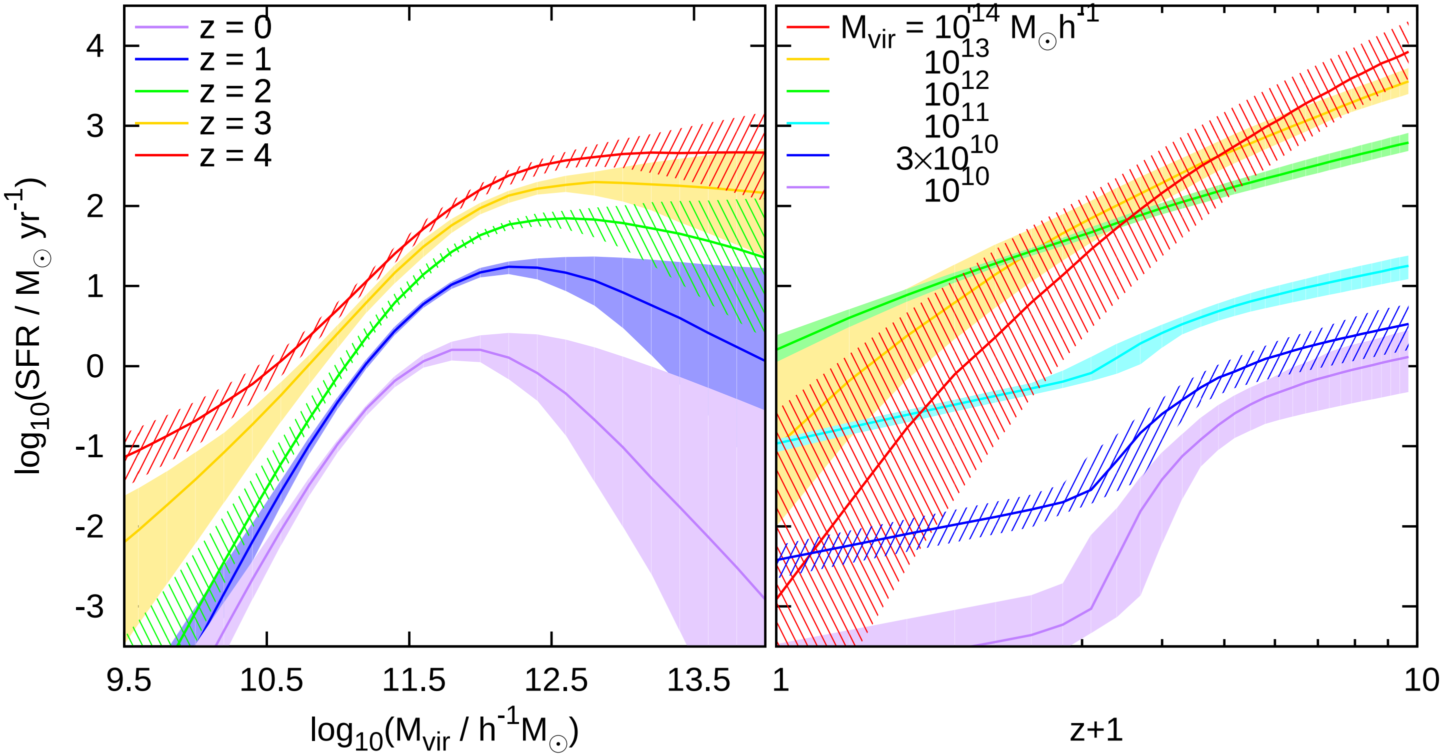
Model III





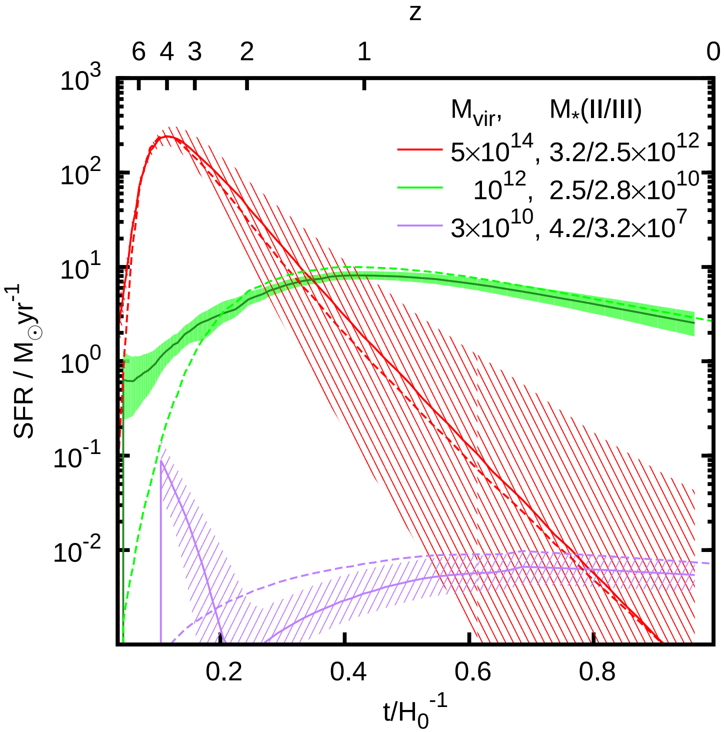
- Use all the observations as data constraints (Data 3).
- Matches all the constraining data.
- Preferred over Model II by over 10^{11} to one!
- Adding more complexity does not statistically improve the fit.

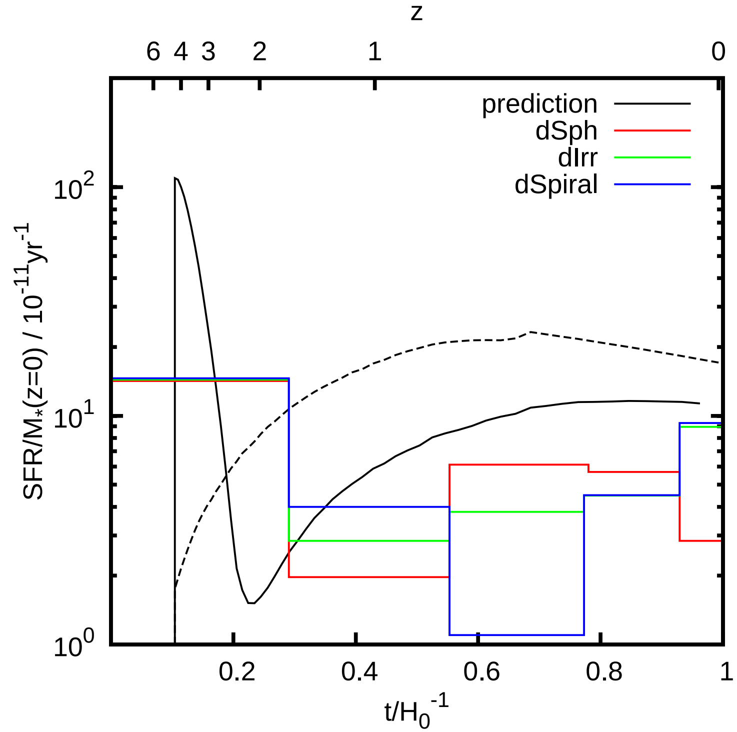
Model III: SFRs



- The SFRs for halos $\geq 10^{12}\, h^{-1}\mathrm{M}_{\odot}$ are the same as in Model II.
- For lower mass halos star formation is less efficient below $z_c \sim 2$.
  - Could owe to some IGM preheating process that occurs at $z \sim 2$, e.g. supernova heating, blasar heating, or pancake heating.

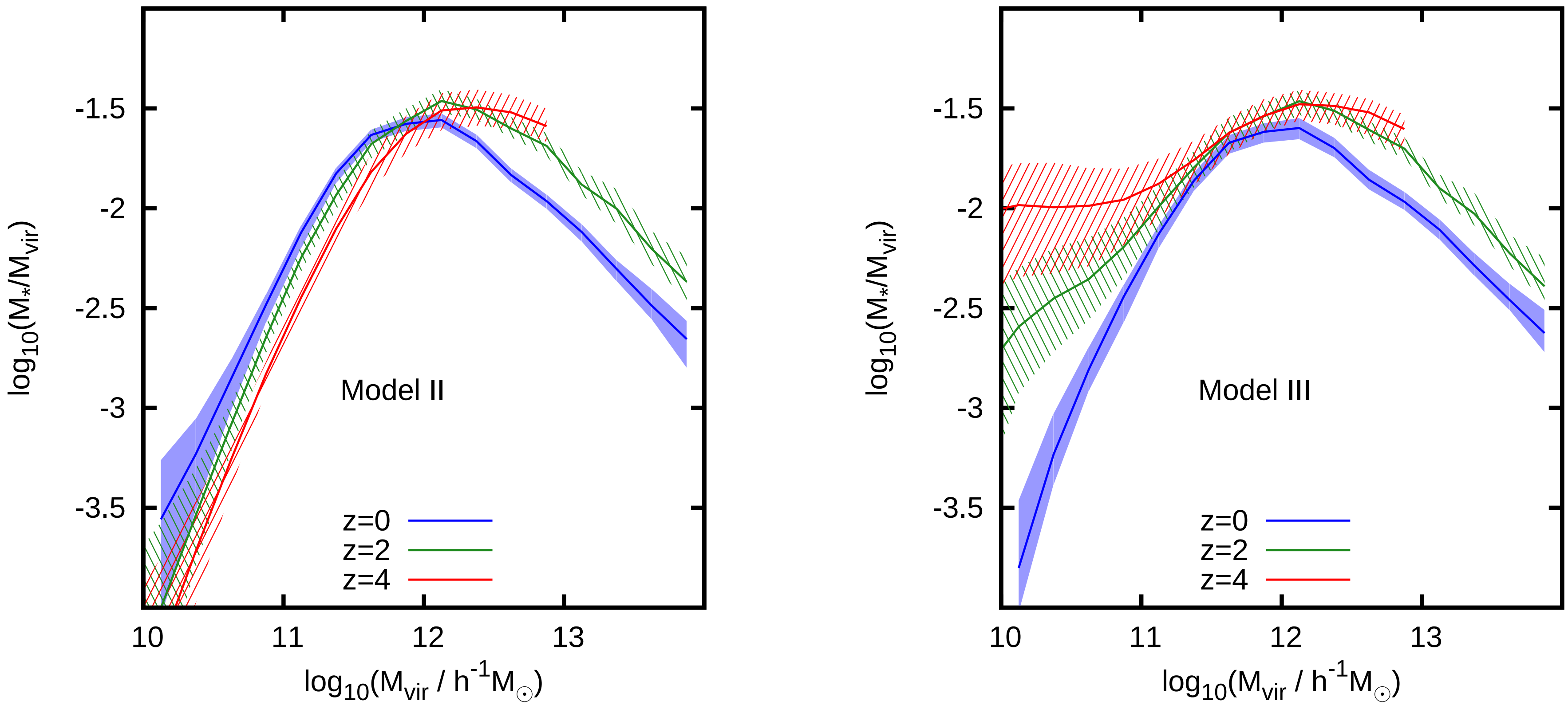
Star Formation Histories





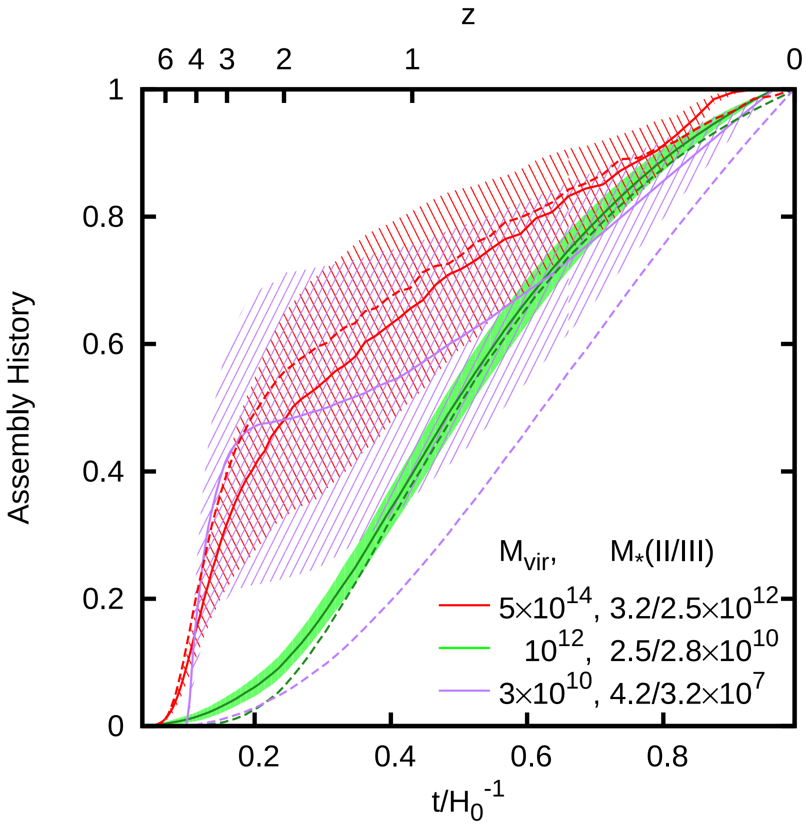
- The most massive galaxies form most of their stars at $z > 3$.
- Milky Way's have an almost constant SFR after $z \sim 1$.
- Unlike Model II, Model III has a bimodal star formation history for dwarfs with $> 60\%$ of their stars formed by $z \sim 2$.
  - The predicted old stellar population is observed in star counts.

Stellar Mass-Halo Mass Relation



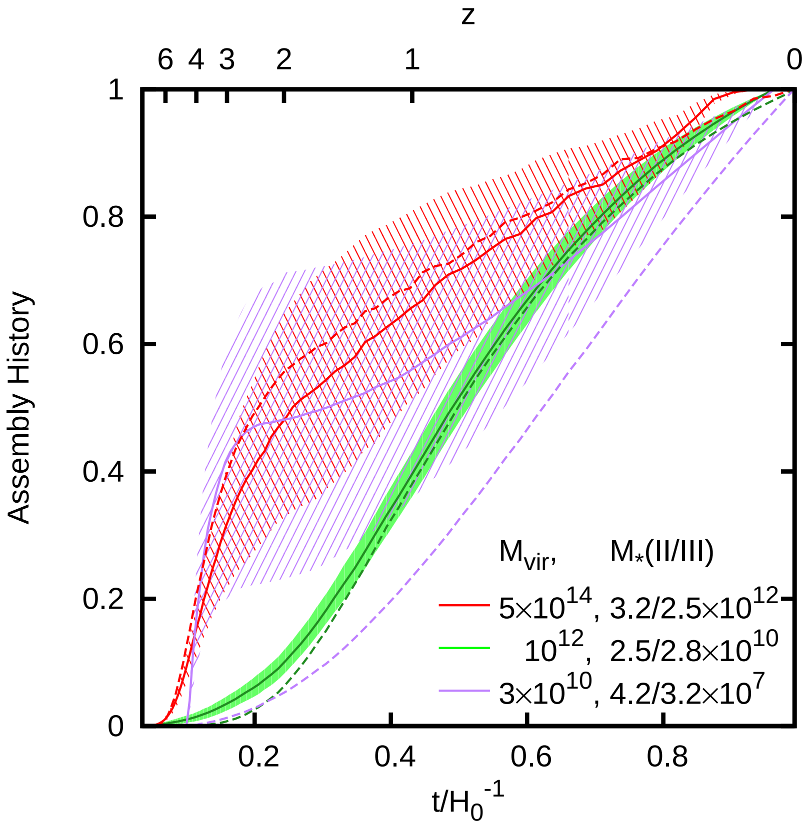
- The results for Model II are similar to past work.
- Above $\sim 10^{12}\,h^{-1}\mathrm{M}_{\odot}$ the stellar mass to halo mass ratio drops.
- In Model III, as one goes to higher redshifts the stellar mass to halo mass ratio becomes almost constant $< 10^{12}\,h^{-1}\mathrm{M}_{\odot}$.

Galaxy Assembly



- Model III: massive and dwarf galaxies have similar assembly histories.
- Model II: Dwarfs assemble like Milky Way galaxies.

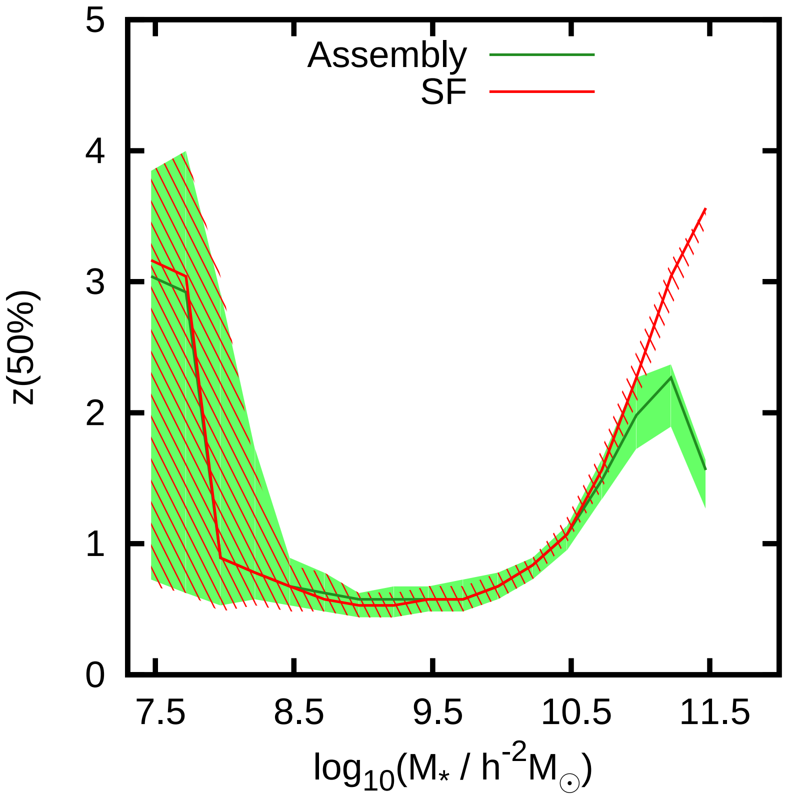
Galaxy Assembly



- Model III: massive and dwarf galaxies have similar assembly histories.
- Model II: Dwarfs assemble like Milky Way galaxies.

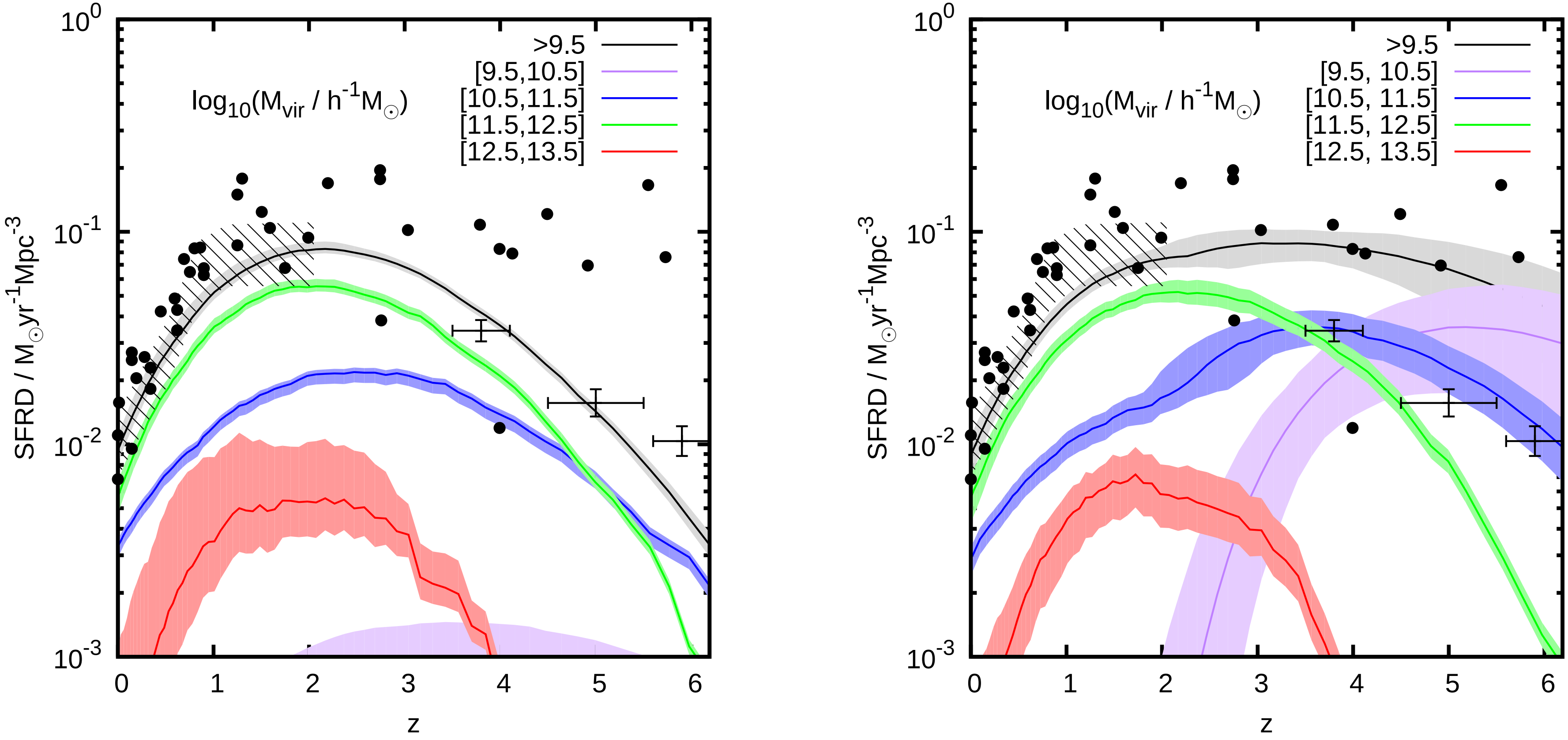
- Progenitor Milky Way galaxies have only assembled a few percent of their stellar mass by $z = 2$ but 25% of their dark halo mass.
  - To observe Milky Way progenitors at $z \geq 2$ requires observing galaxies with stellar masses $< 10^9\, h^{-1}\mathrm{M}_\odot$.

Galaxy Assembly vs. Star Formation



- Massive and dwarf galaxies form half their stars before $z = 2$.
- Intermediate mass galaxies, e.g. the Milky Way, do not form half their stars until $z < 1$.
- Galaxy assembly follows star formation except for $M_\star > 10^{11}\, h^{-1}\mathrm{M}_\odot$.

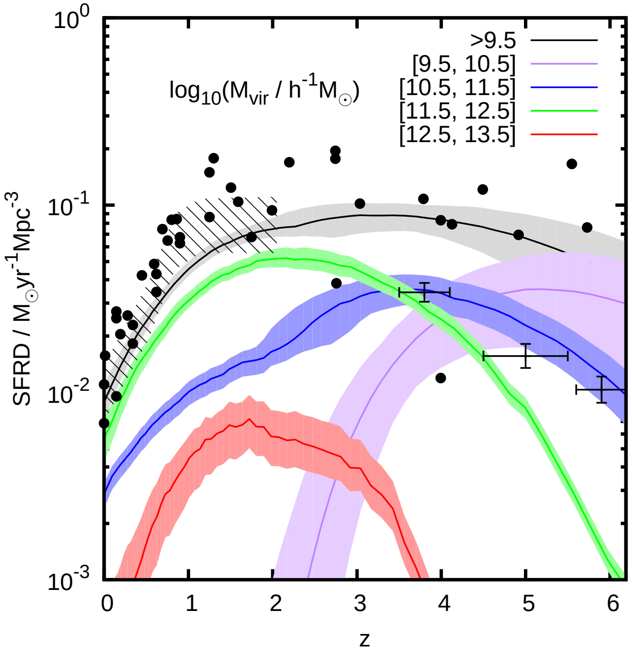
Cosmic Star Formation History



- Model II predicts a cosmic SFR that drops at $z > 3$ while Model III predicts it remains almost constant.
- Both models predict $\sim 10^{12}$ halos dominate star formation at $z < 3$.
- Model III: Ever decreasing halo masses dominate at $z > 3$.

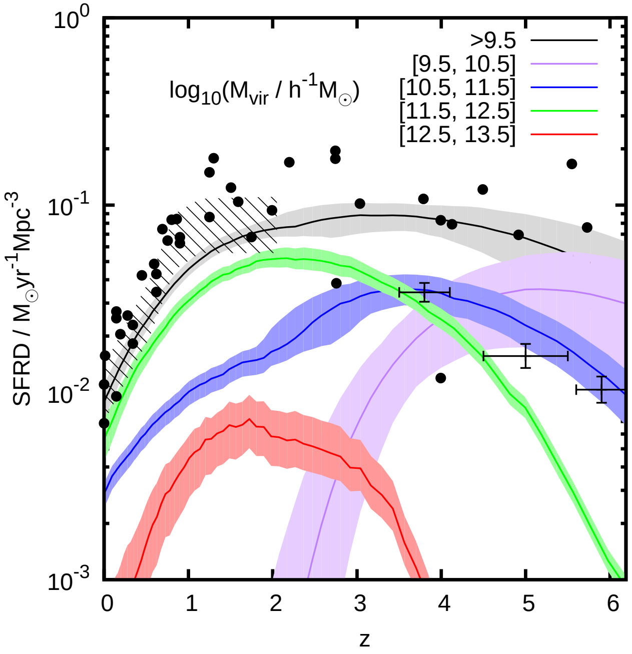
Star Formation in Massive Halos

- In halos $> 10^{12.5} h^{-1} \mathrm{M}_\odot$ galaxies only ever contribute $< 10\%$ to the cosmic star formation rate.



Star Formation in Massive Halos

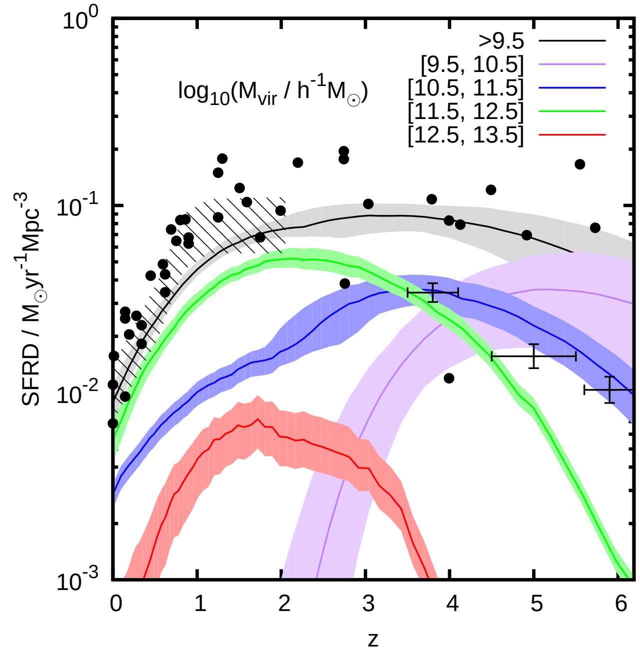
- In halos $> 10^{12.5} h^{-1} \mathrm{M}_{\odot}$ galaxies only ever contribute $< 10\%$ to the cosmic star formation rate.



- Using their inferred SFRs, Submm galaxies are claimed to contribute 1/3 to 1/2 of the cosmic SFR.
- The clustering strength of submm galaxies implies that they are in massive halos.

Star Formation in Massive Halos

- In halos $> 10^{12.5} h^{-1} \mathrm{M}_\odot$ galaxies only ever contribute $< 10\%$ to the cosmic star formation rate.



- Using their inferred SFRs, Submm galaxies are claimed to contribute 1/3 to 1/2 of the cosmic SFR.
- The clustering strength of submm galaxies implies that they are in massive halos.
  - $\Rightarrow$ The star formation rates of submm galaxies must be overestimated.

Conclusions

- We investigated what aspects of galaxy formation are required by the observed galaxy stellar mass functions at $z = 0, 1.15, 2.5$ & 4.0 and the low redshift galaxy luminosity functions in clusters.

Conclusions

- We investigated what aspects of galaxy formation are required by the observed galaxy stellar mass functions at $z = 0, 1.15, 2.5$ & 4.0 and the low redshift galaxy luminosity functions in clusters.
- High mass halos ($\geq 10^{12}\,h^{-1}\mathrm{M}_{\odot}$) must have more efficient star formation at higher redshifts.
    - Perhaps owing to the filamentary cold mode accretion seen in many simulations.

Conclusions

- We investigated what aspects of galaxy formation are required by the observed galaxy stellar mass functions at $z = 0, 1.15, 2.5$ & 4.0 and the low redshift galaxy luminosity functions in clusters.
- High mass halos ($\geq 10^{12}\, h^{-1}\mathrm{M}_{\odot}$) must have more efficient star formation at higher redshifts.
  - Perhaps owing to the filamentary cold mode accretion seen in many simulations.
- Low mass halos ($\leq 10^{11}\, h^{-1}\mathrm{M}_{\odot}$) must have less efficient star formation below $z \sim 2$.
  - Perhaps owing to a preheating process that occurs at $z \sim 2$.
  - e.g. supernova heating, blasar heating, or pancake heating.

Conclusions

- Progenitor Milky Way galaxies have only assembled a few percent of their stellar mass by $z = 2$ but 25% of their dark halo mass.
  - Galaxy formation has to be delayed relative to dark halo formation perhaps owing to galactic wind recycling.
  - To observe Milky Way progenitors at $z \geq 2$ requires observing galaxies with stellar masses $< 10^9\, h^{-1}\mathrm{M}_\odot$, something that is very difficult in present surveys.

Conclusions

- Progenitor Milky Way galaxies have only assembled a few percent of their stellar mass by $z = 2$ but 25% of their dark halo mass.
  - Galaxy formation has to be delayed relative to dark halo formation perhaps owing to galactic wind recycling.
  - To observe Milky Way progenitors at $z \geq 2$ requires observing galaxies with stellar masses $< 10^9\,h^{-1}\mathrm{M}_\odot$, something that is very difficult in present surveys.
- Dwarf galaxies formed $> 60\%$ of their stars by $z = 2$ regardless of morphological type but many continue forming stars today.
  - Consistent with the ages inferred from observed star counts.

Conclusions

- Progenitor Milky Way galaxies have only assembled a few percent of their stellar mass by $z = 2$ but 25% of their dark halo mass.
  - Galaxy formation has to be delayed relative to dark halo formation perhaps owing to galactic wind recycling.
  - To observe Milky Way progenitors at $z \geq 2$ requires observing galaxies with stellar masses $< 10^9\,h^{-1}\mathrm{M}_\odot$, something that is very difficult in present surveys.
- Dwarf galaxies formed $> 60\%$ of their stars by $z = 2$ regardless of morphological type but many continue forming stars today.
  - Consistent with the ages inferred from observed star counts.
- Massive galaxies form most of their stars at $z > 3$ but half of their final mass is assembled at $z < 2$.
  - Massive galaxies assemble through merging.

Conclusions

- Galaxies in halos $> 10^{12.5}\,h^{-1}\mathrm{M}_{\odot}$ contribute very little to the global star formation of the Universe at $z > 2$ in apparent contradiction to their large inferred SFRs.