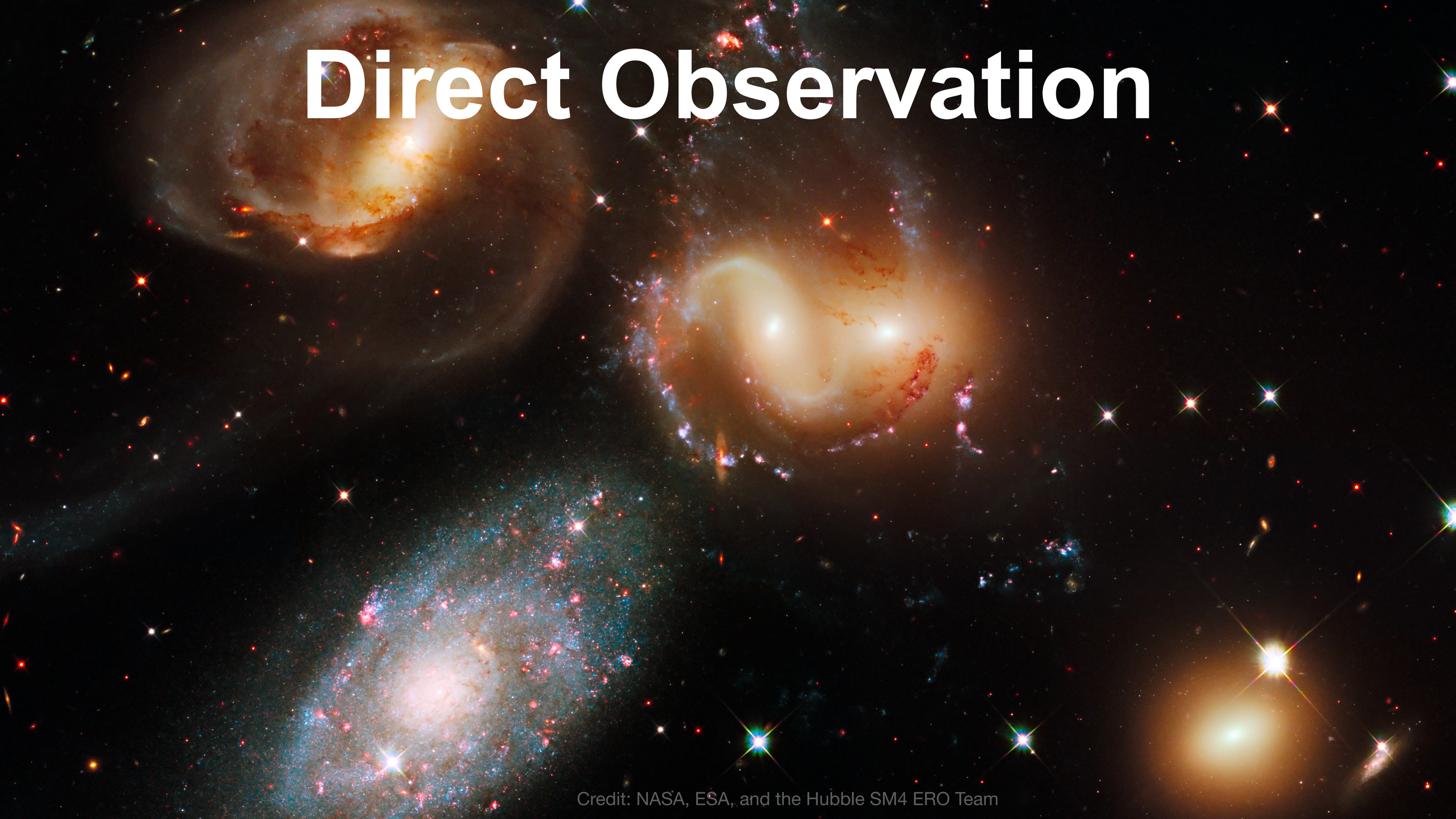
Revealing Galaxy Formation with Empirical Models

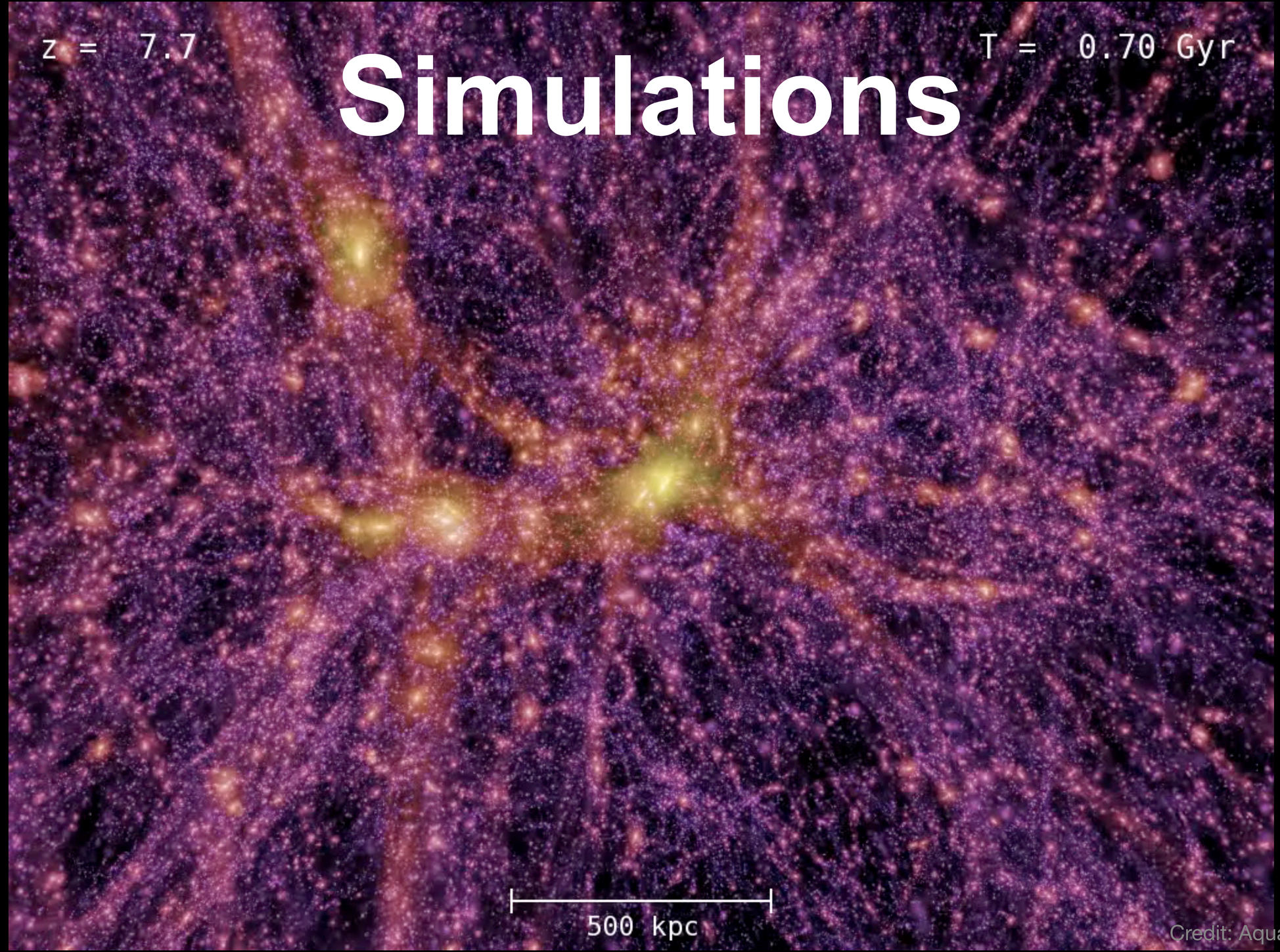
Peter Behroozi (University of Arizona)

YITP School, 6/3/25

Direct Observation



Credit: NASA, ESA, and the Hubble SM4 ERO Team

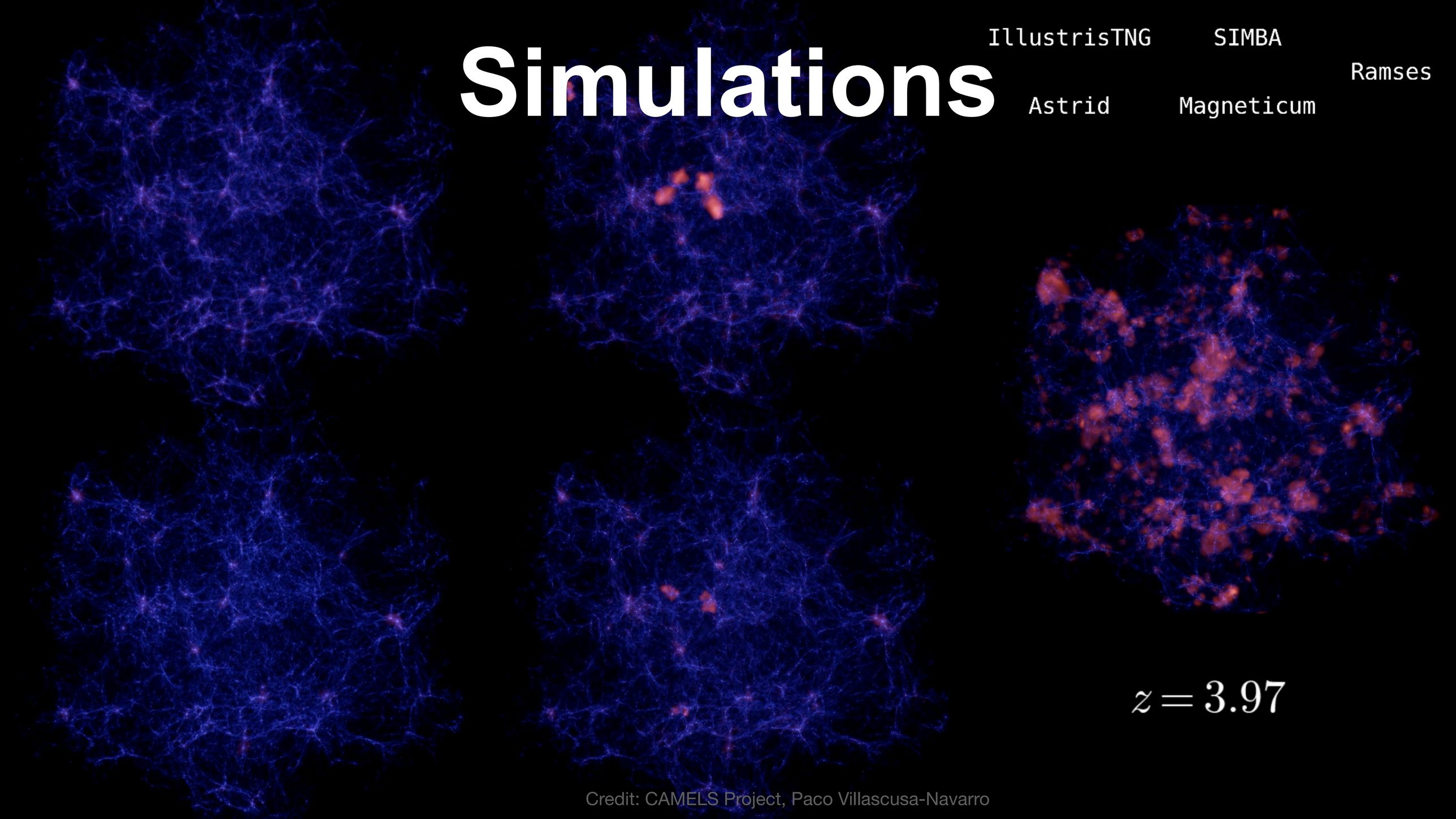


Credit: Aquarius Simulation Team

Simulations

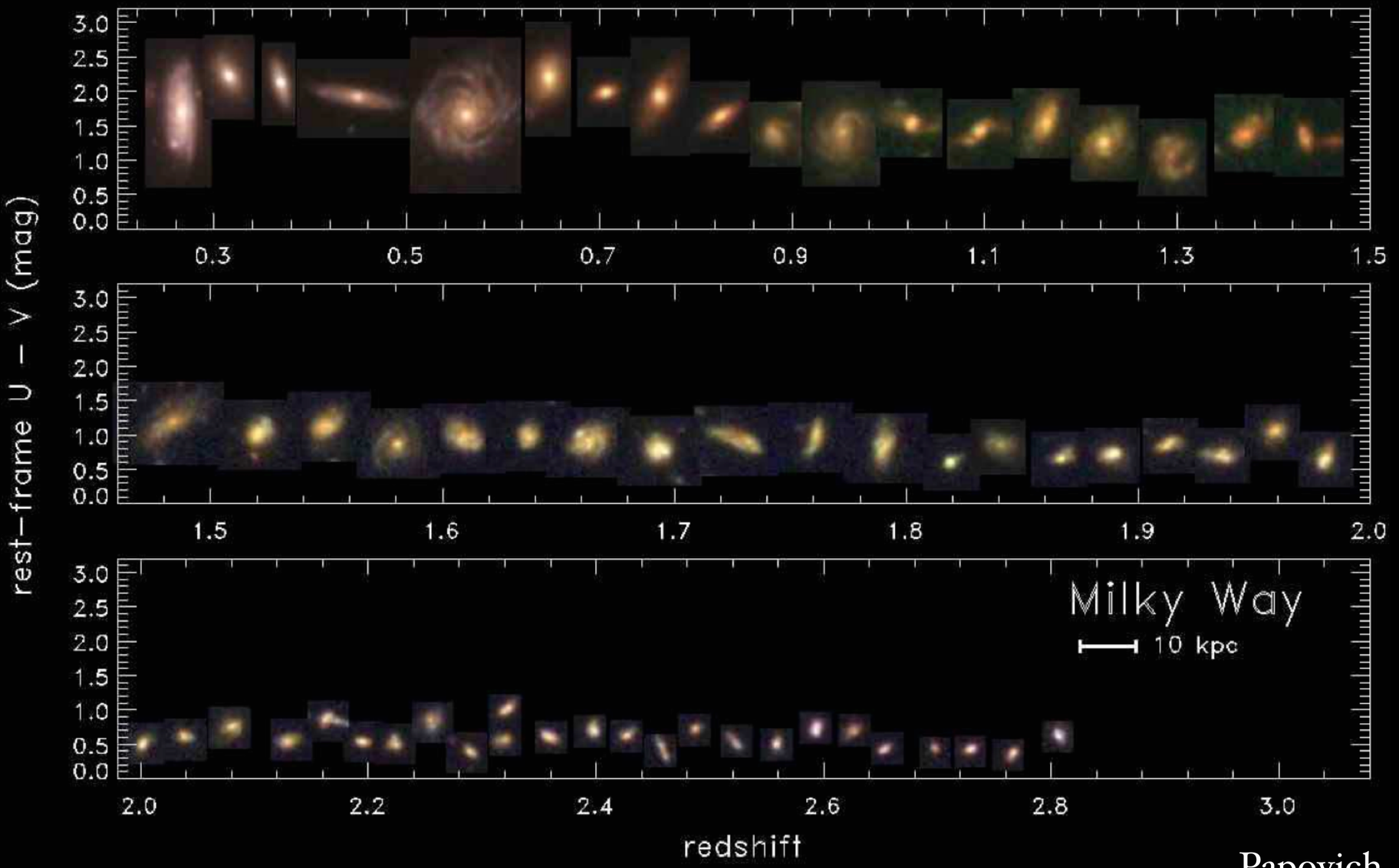
IllustrisTNG SIMBA Ramses

Astrid Magneticum

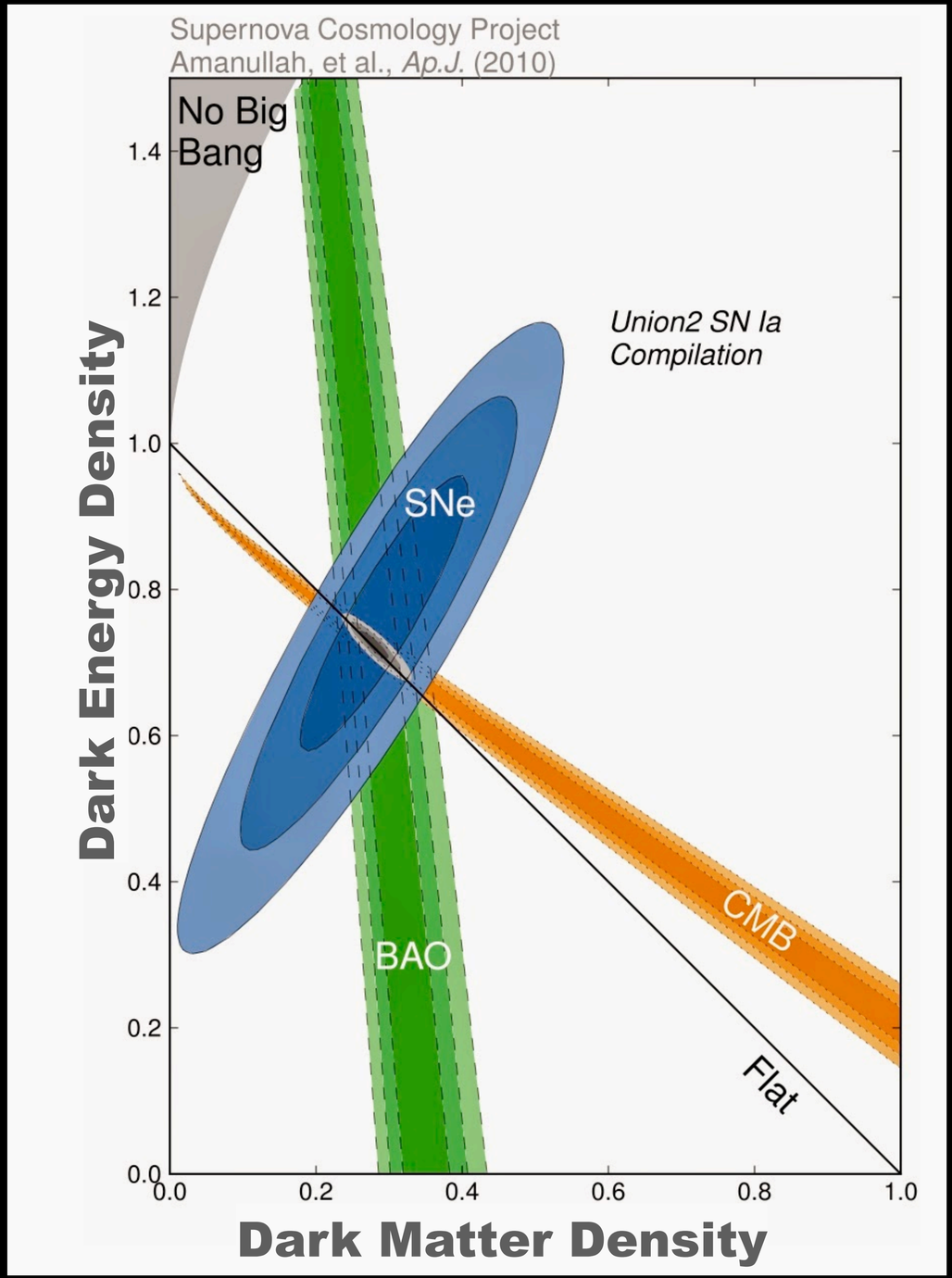


Credit: CAMELS Project, Paco Villascusa-Navarro

Empirical Modeling



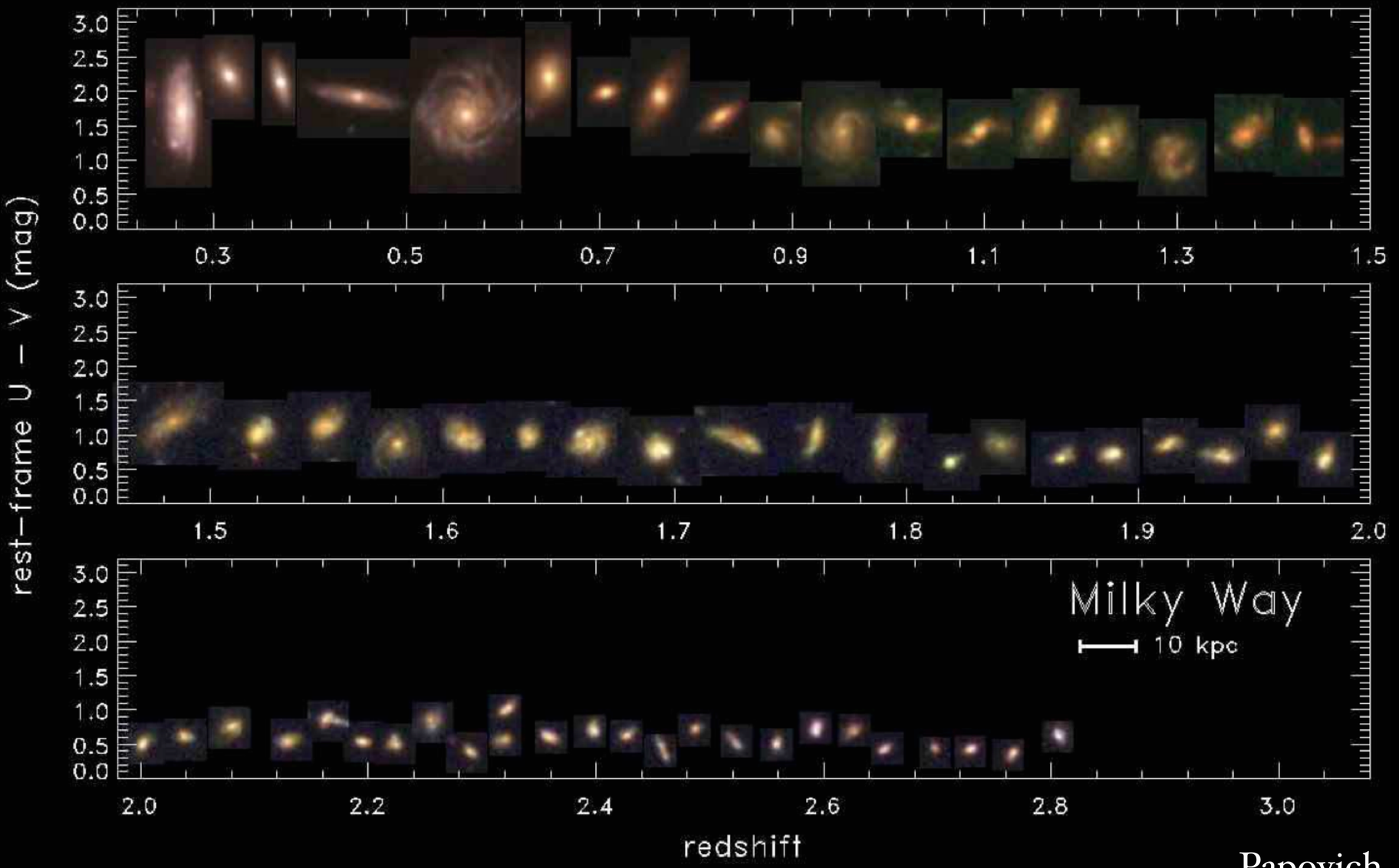
Papovich, ..., PB et al. (2015)


Supernova Cosmology Project
Amanullah, et al., Ap.J. (2010)
No Big Bang
Union2 SN Ia Compilation
SNe
BAO
CMB
Flat
Dark Energy Density
Dark Matter Density
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1.4

Empirical Models

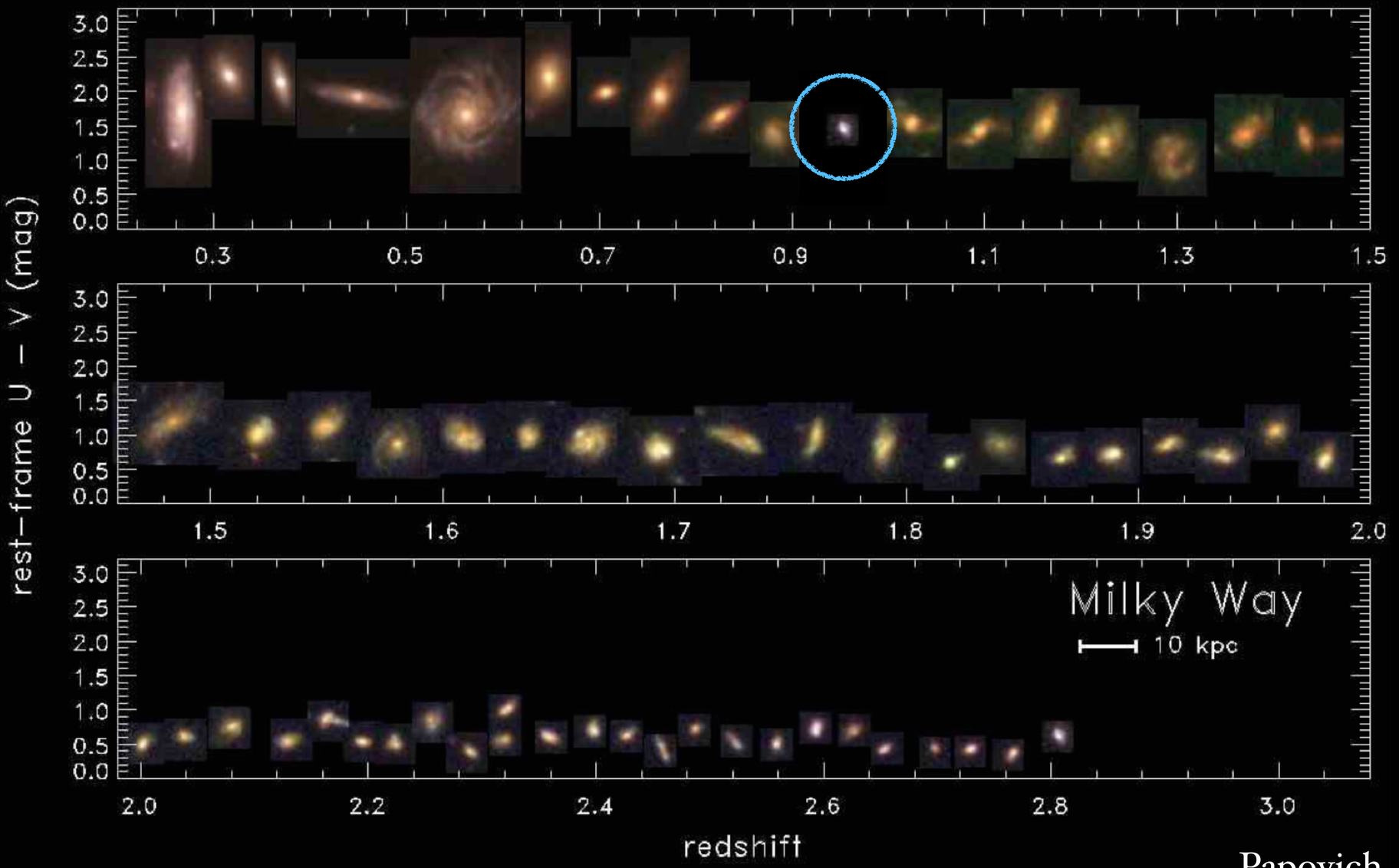
The actual growth of galaxies & SMBHs must be consistent with all available observations.

Empirical Modeling



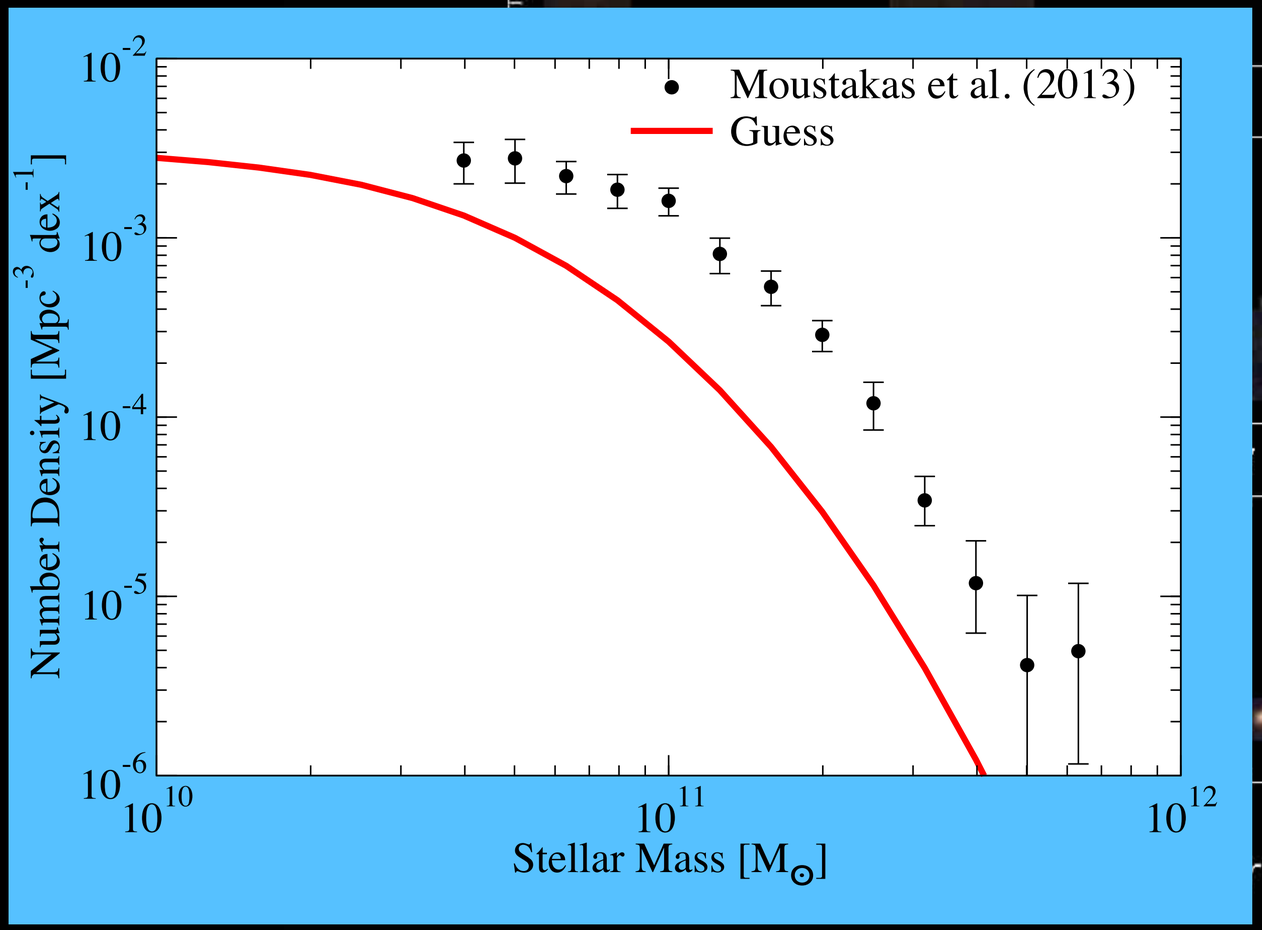
Papovich, ..., PB et al. (2015)

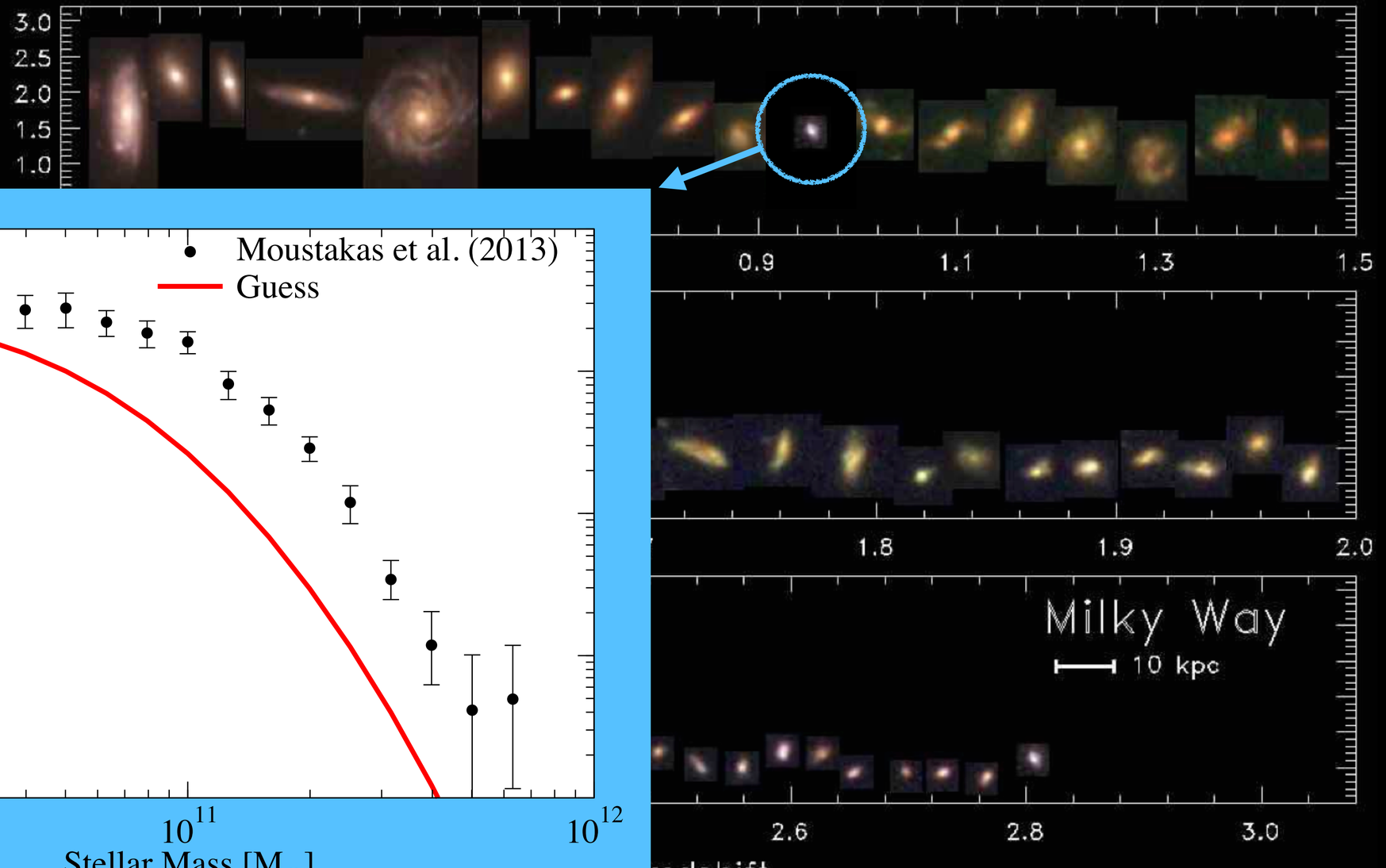
Empirical Modeling



Papovich, ..., PB et al. (2015)

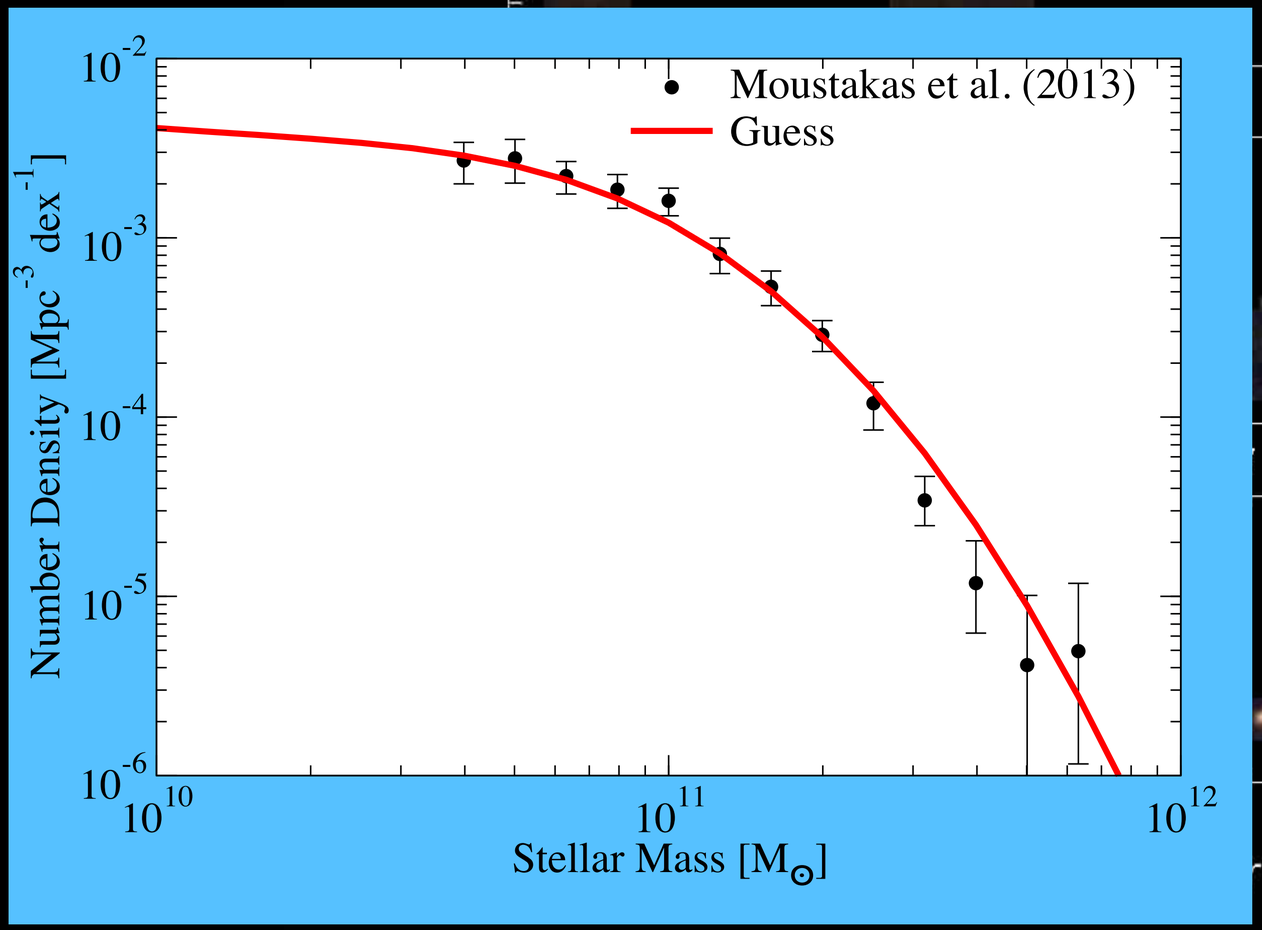
Empirical Modeling

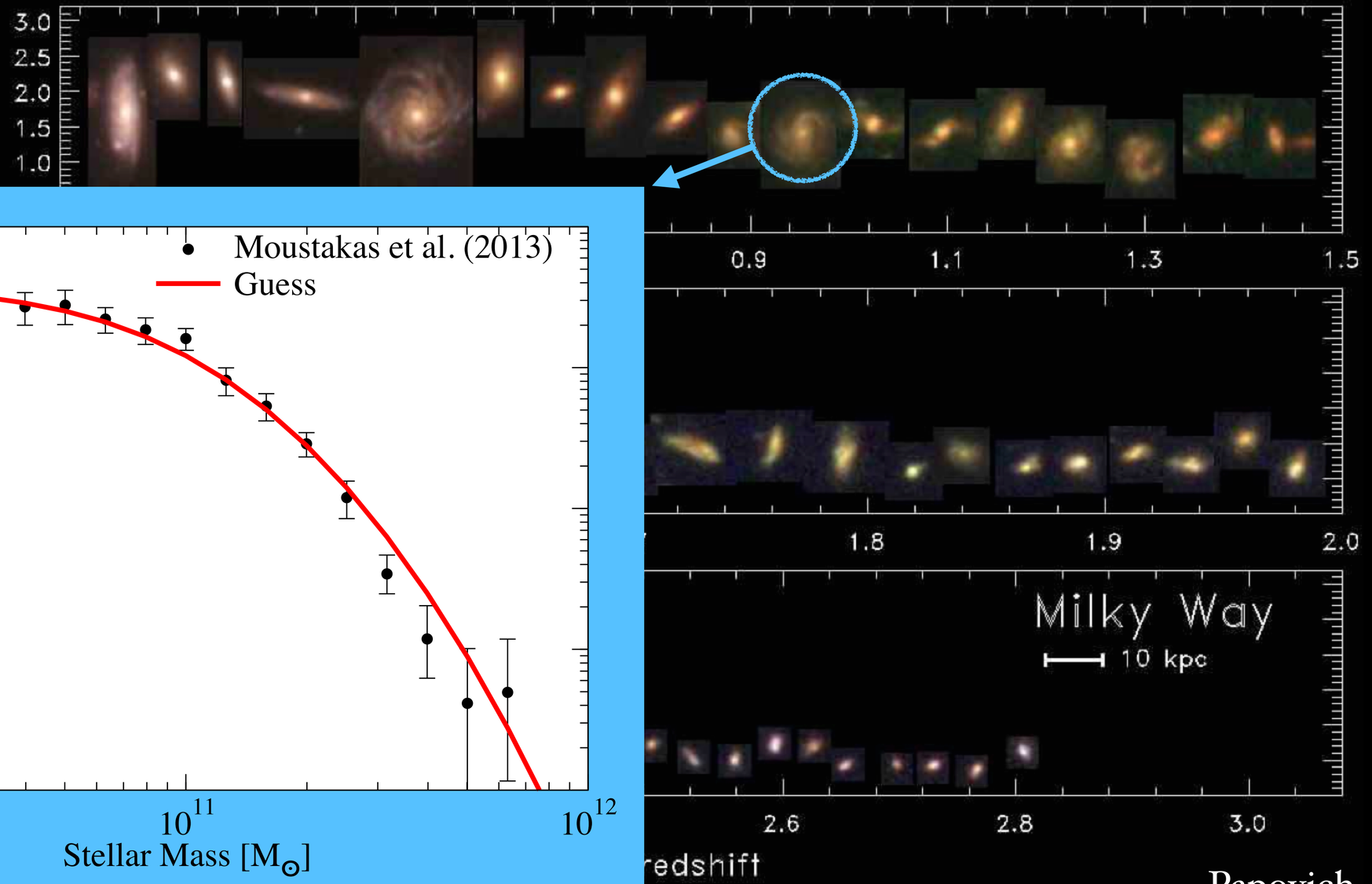




Papovich, ..., PB et al. (2015)

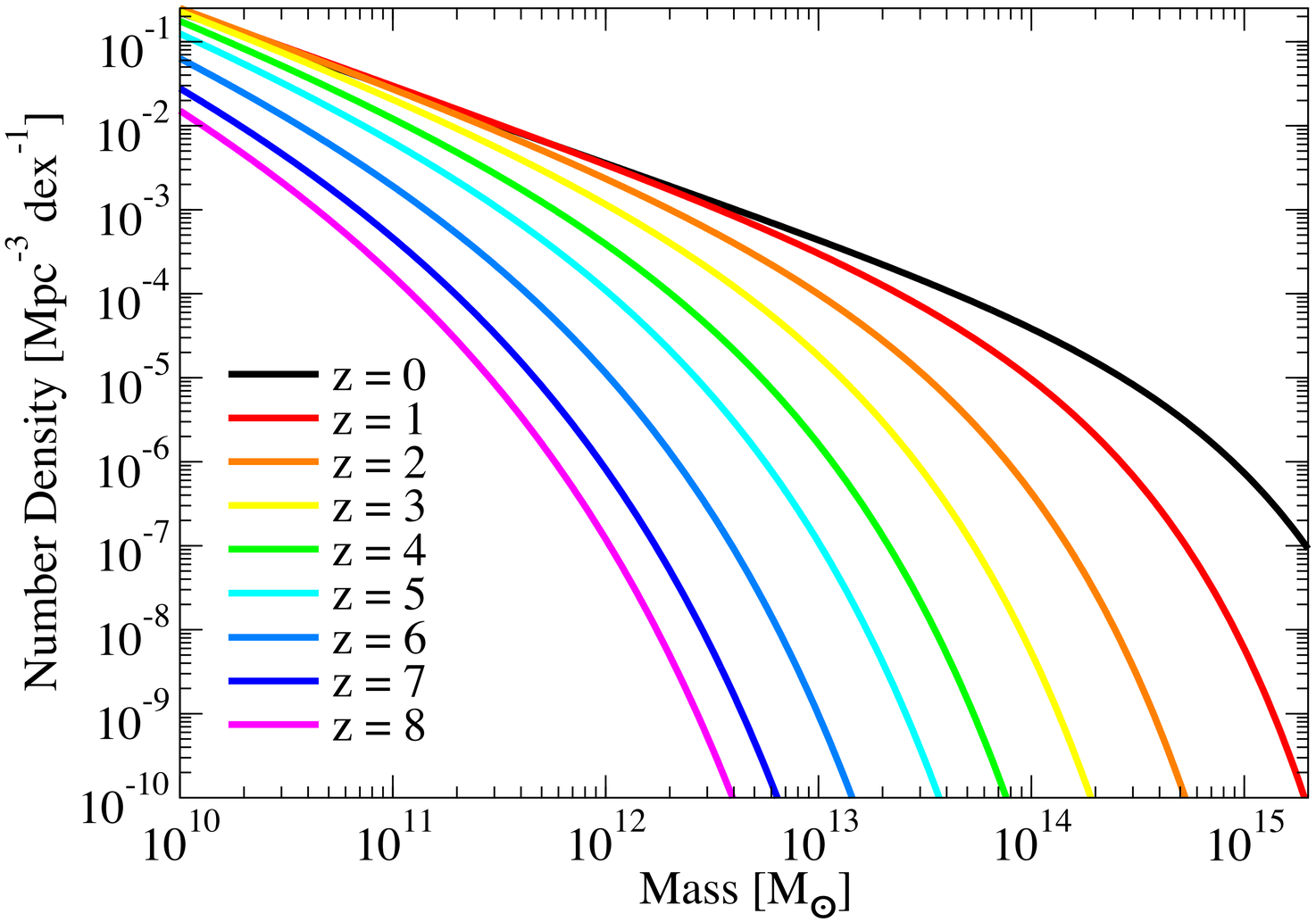
Empirical Modeling





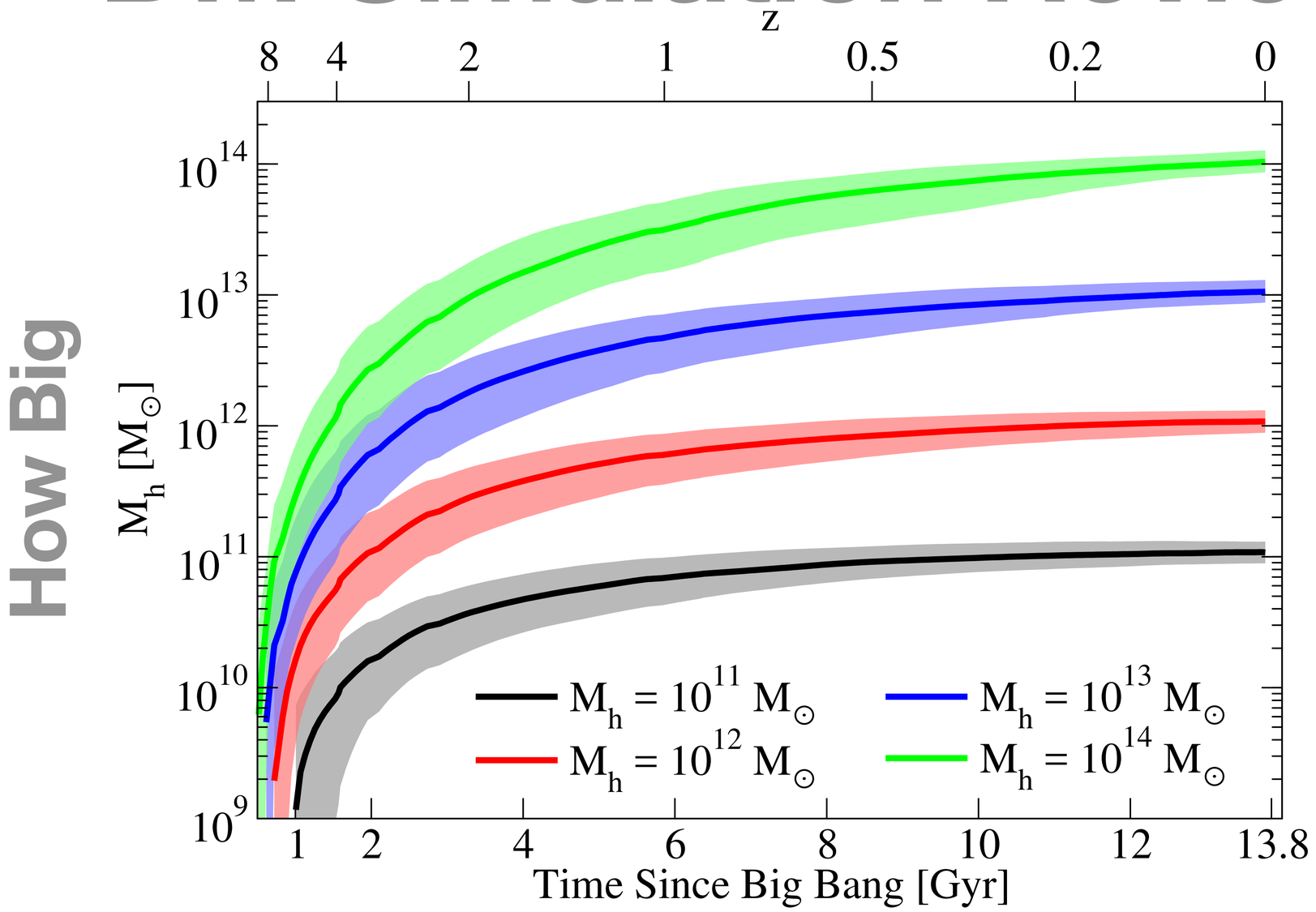
Papovich, ..., PB et al. (2015)

DM Simulation Review



PB+ (2013)

DM Simulation Review

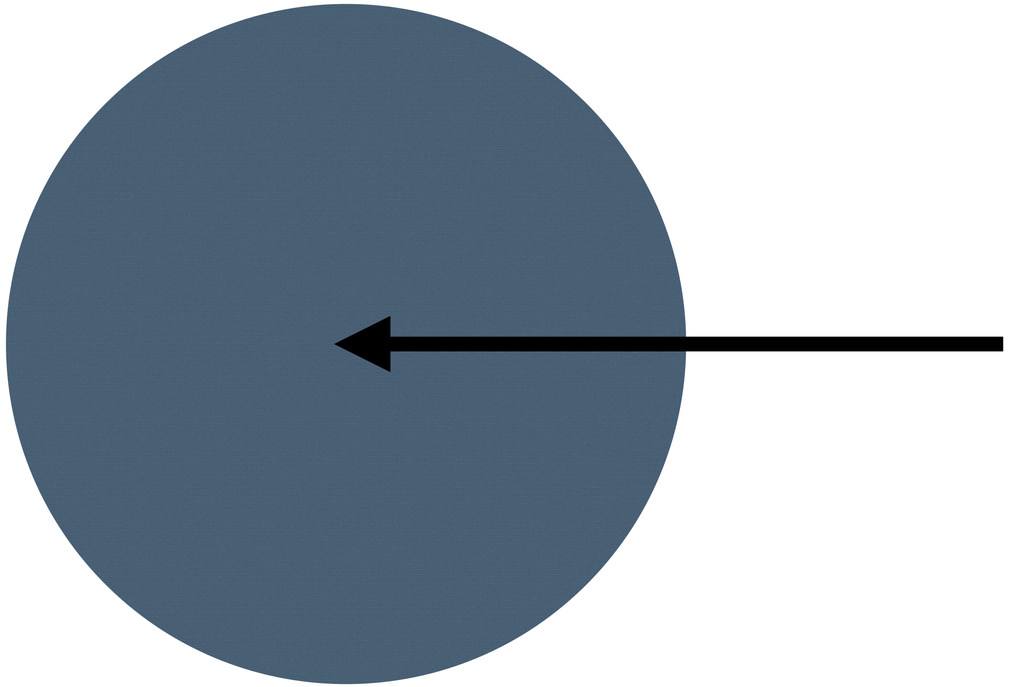


How Big

Cosmic Time

PB+ (2013)

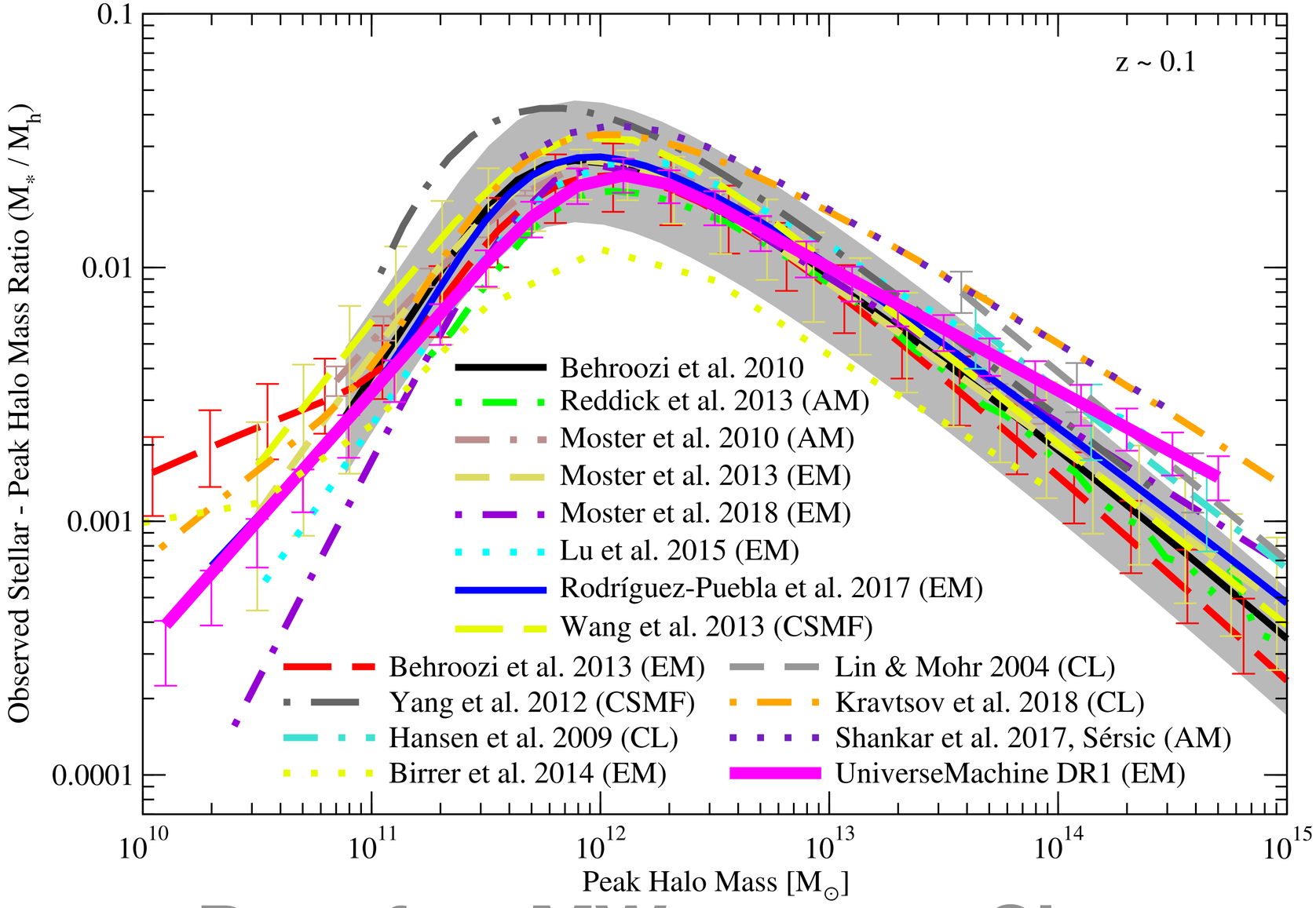
Empirically Modeling Galaxies



$$M_*(M_h, z)$$

Lots of Measurements at z=0

Efficiency of Star Formation



Dwarfs MW Clusters

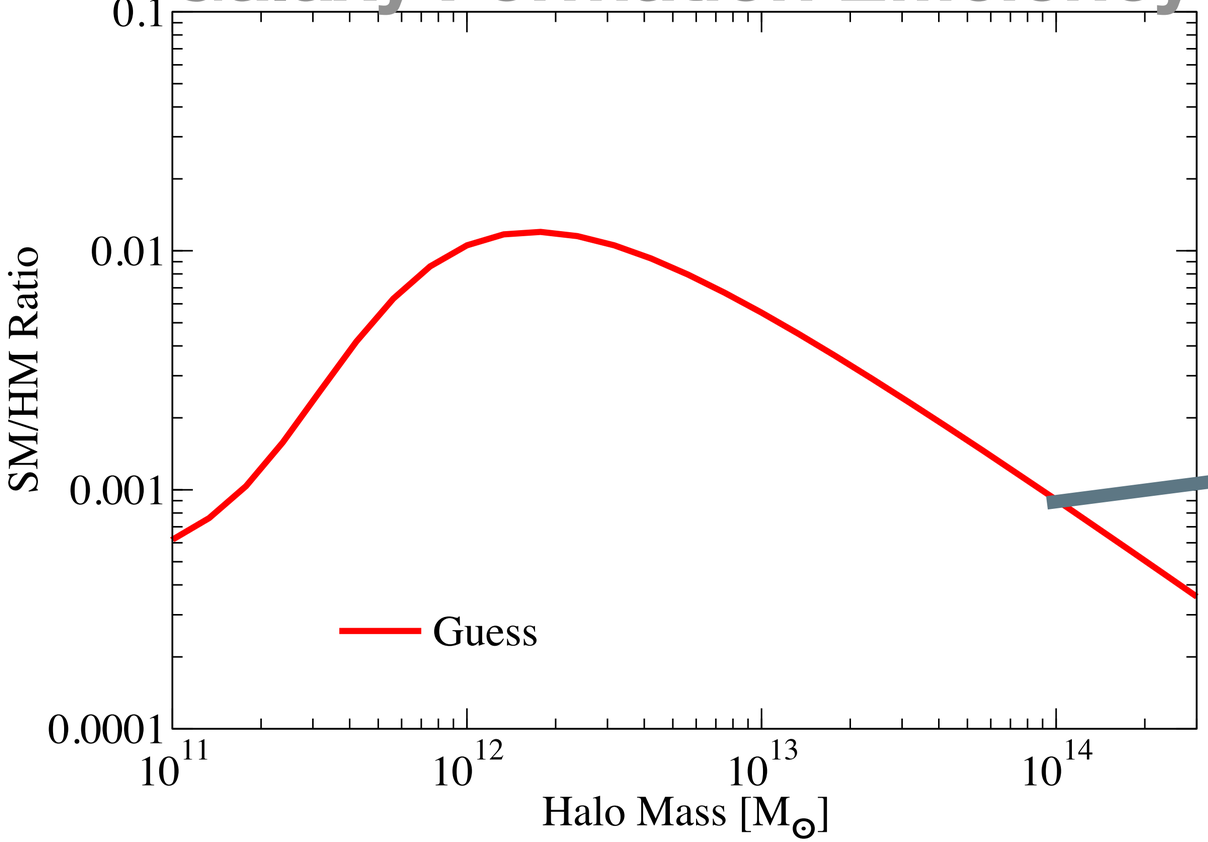
PB+ (2019)

At z>0?

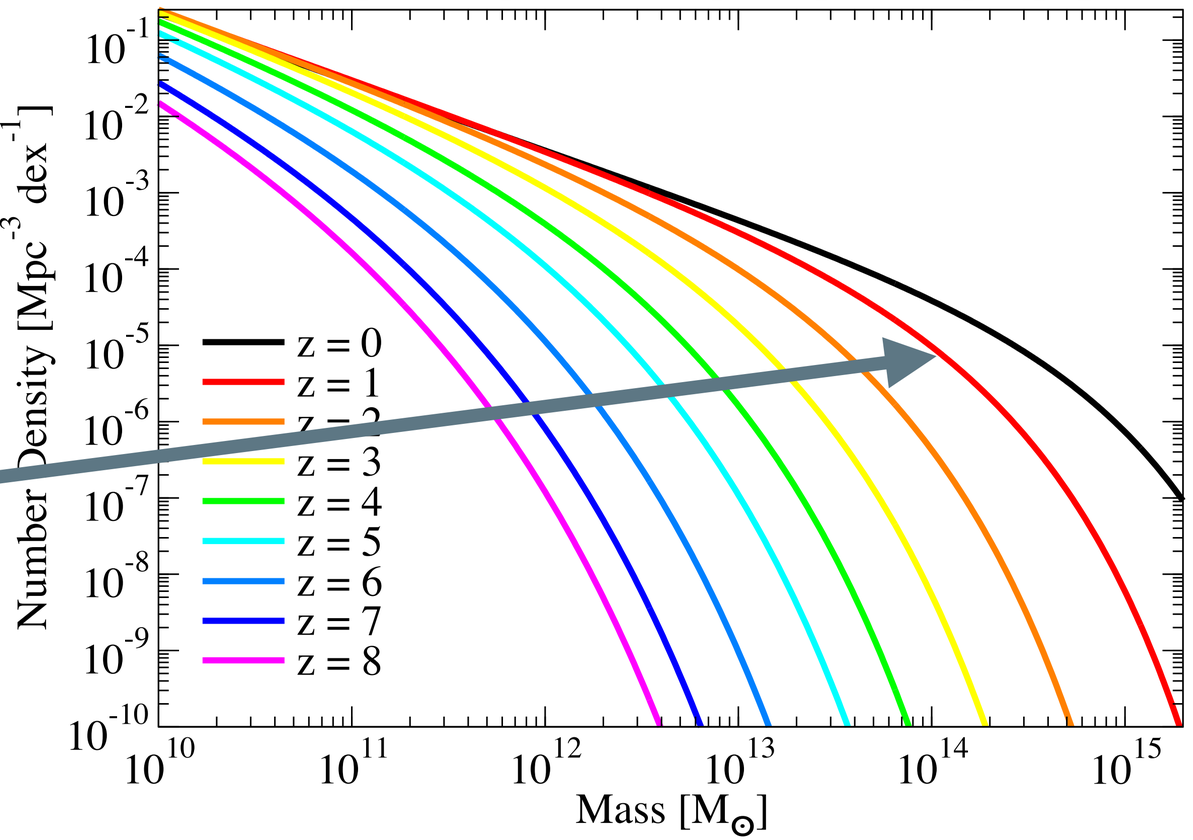
At z>0?

Guess and Check!

Galaxy Formation Efficiency



of Dark Matter Halos

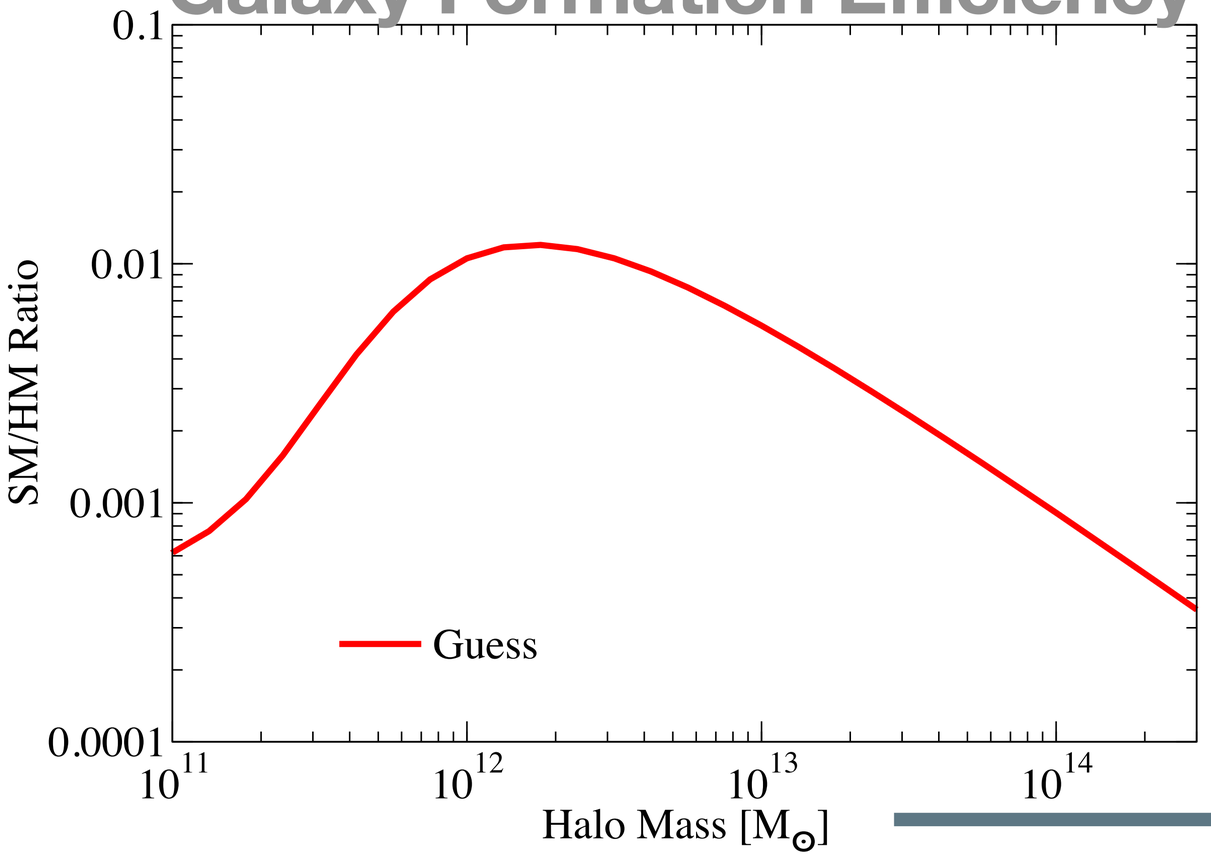


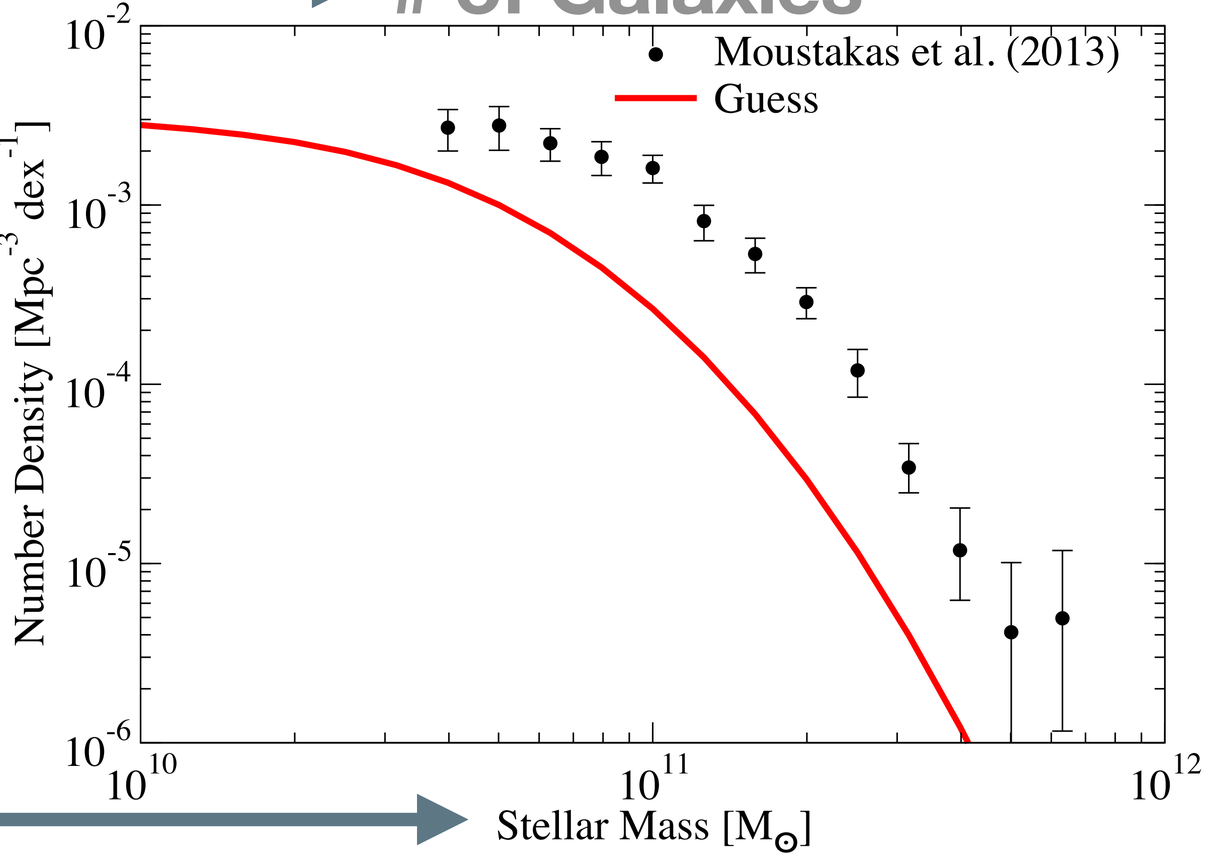
z=1

At z>0?

Guess and Check!

Galaxy Formation Efficiency → # of Galaxies



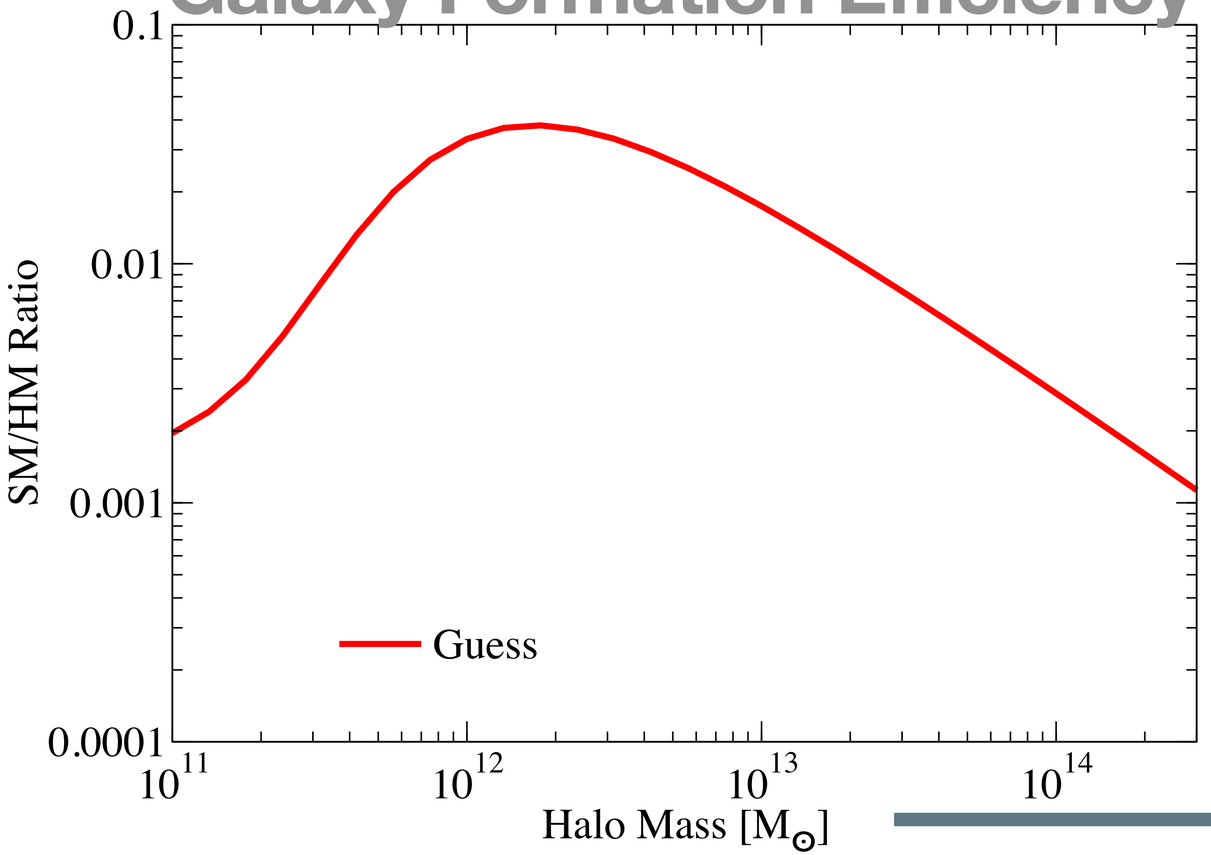


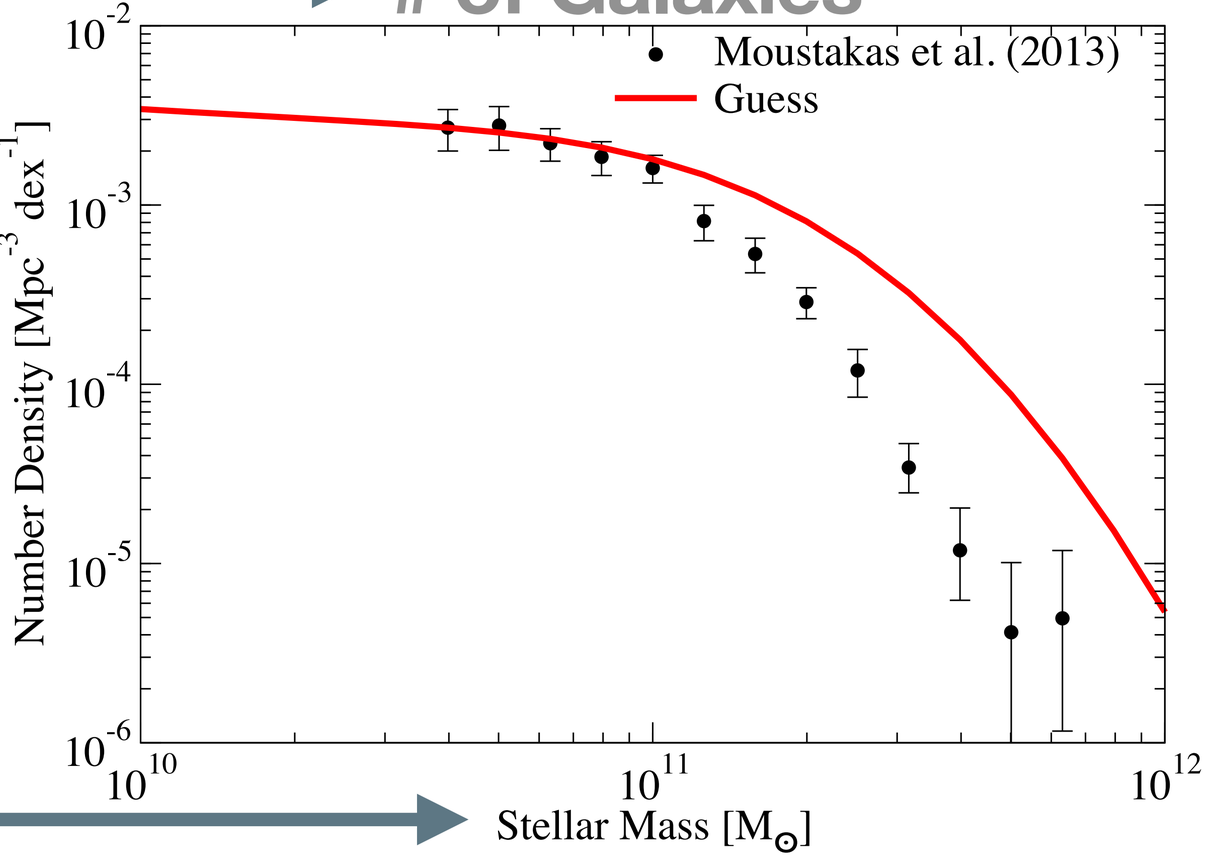
z=1

At z>0?

Guess and Check!

Galaxy Formation Efficiency → # of Galaxies



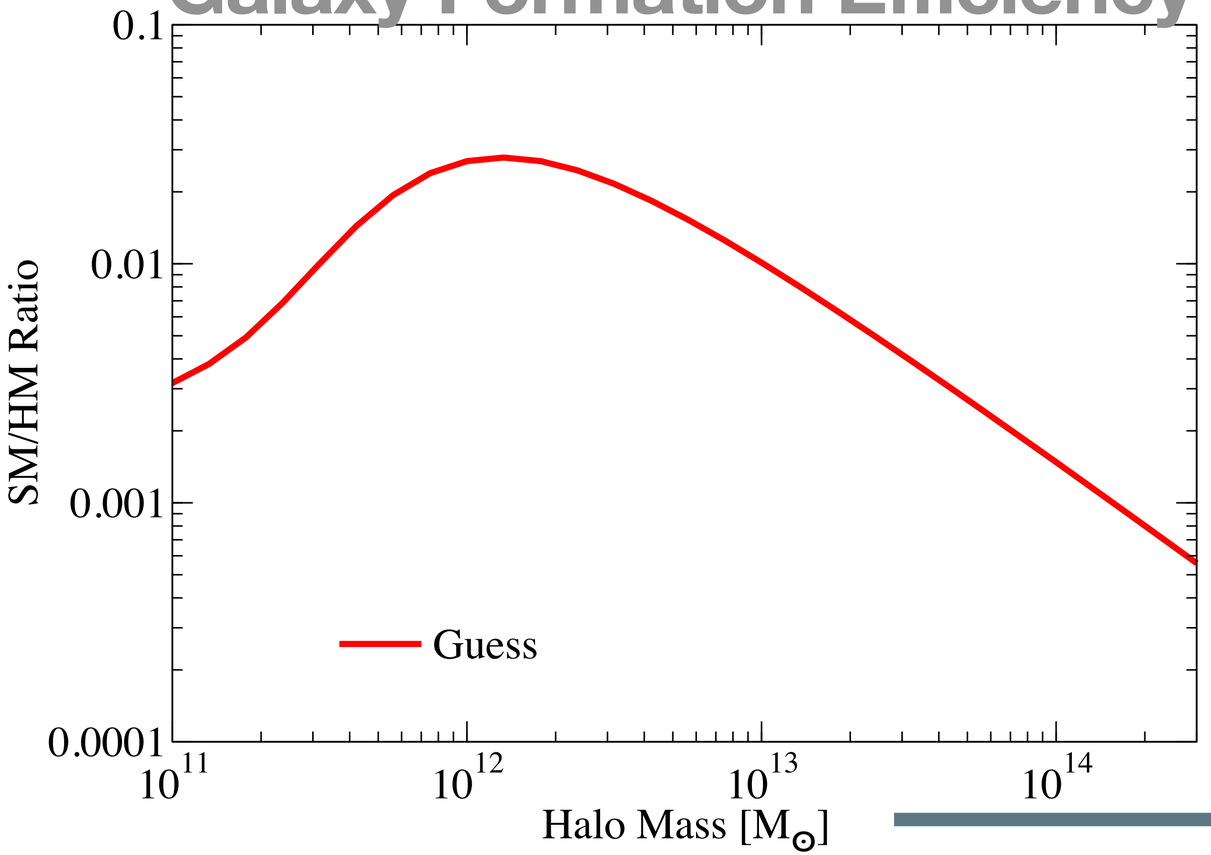


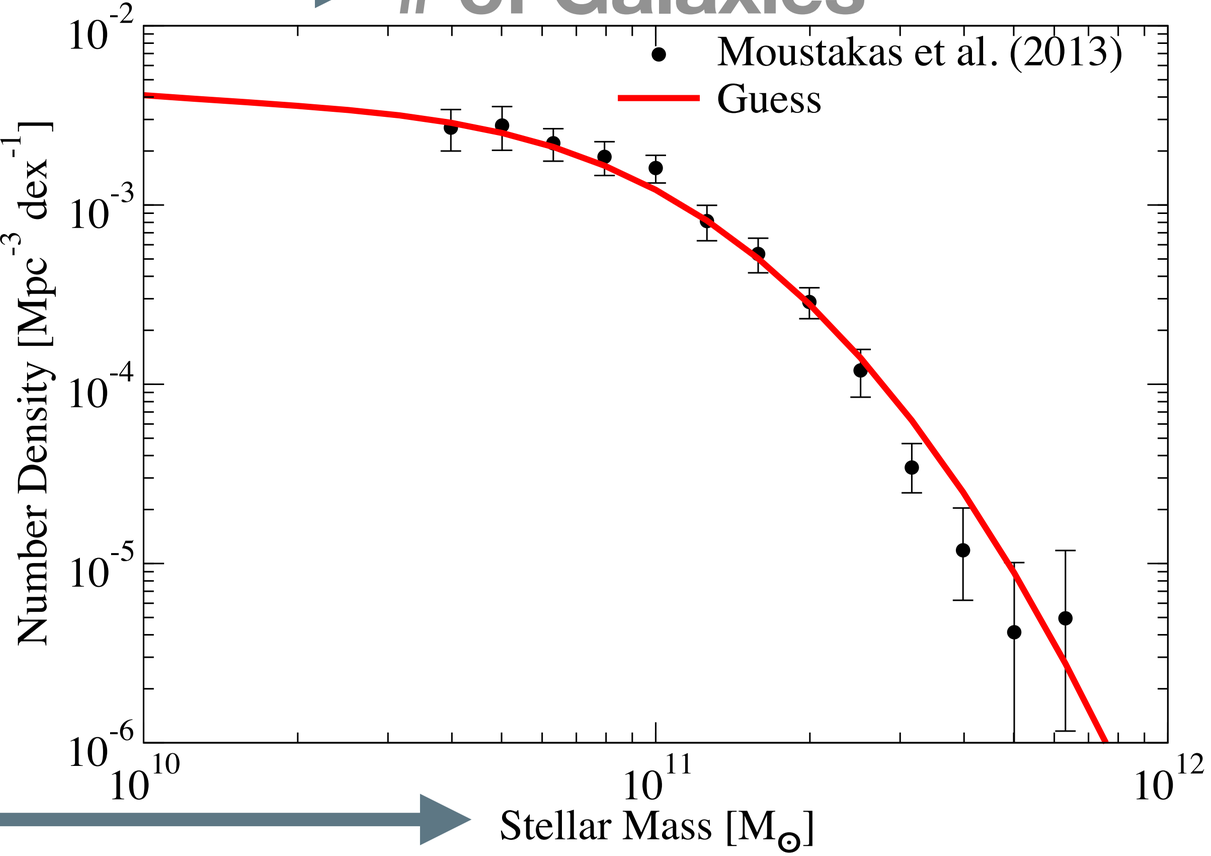
z=1

At z>0?

Guess and Check!

Galaxy Formation Efficiency → # of Galaxies



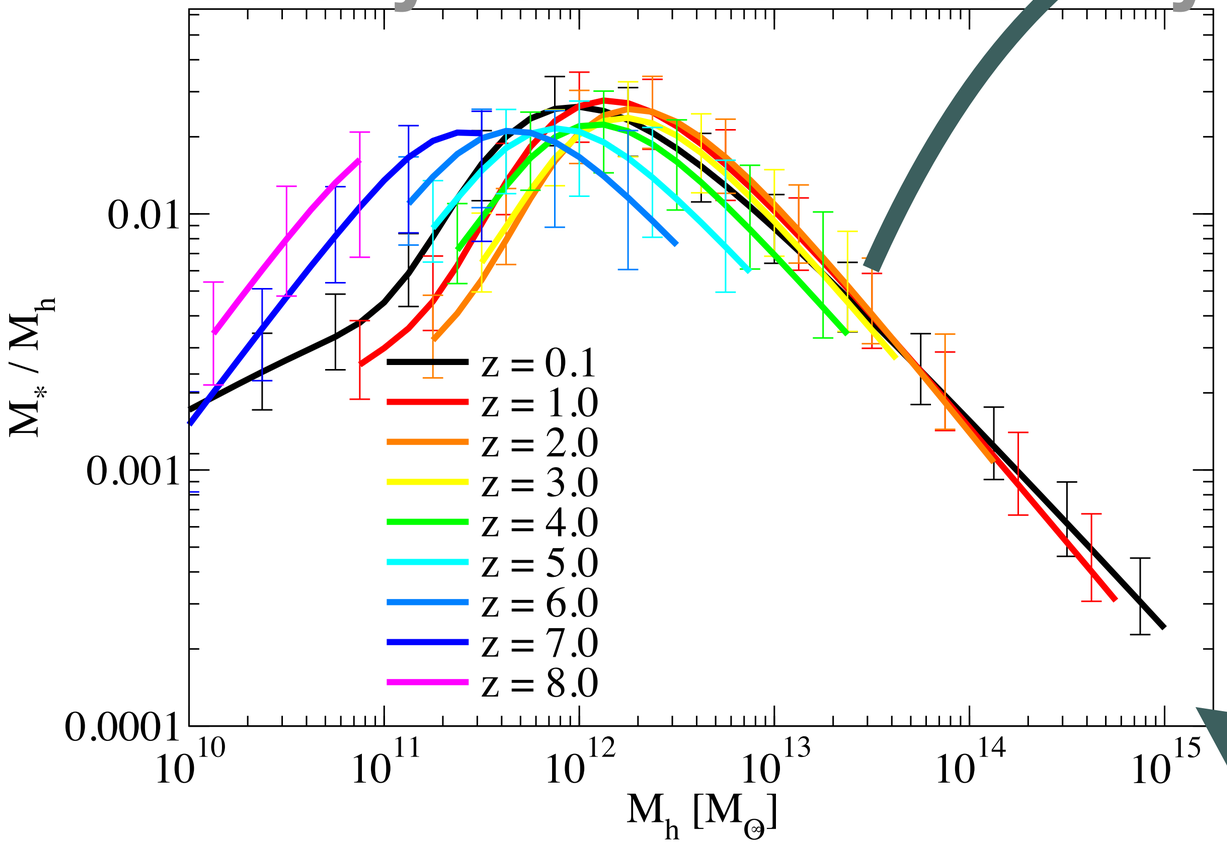


z=1

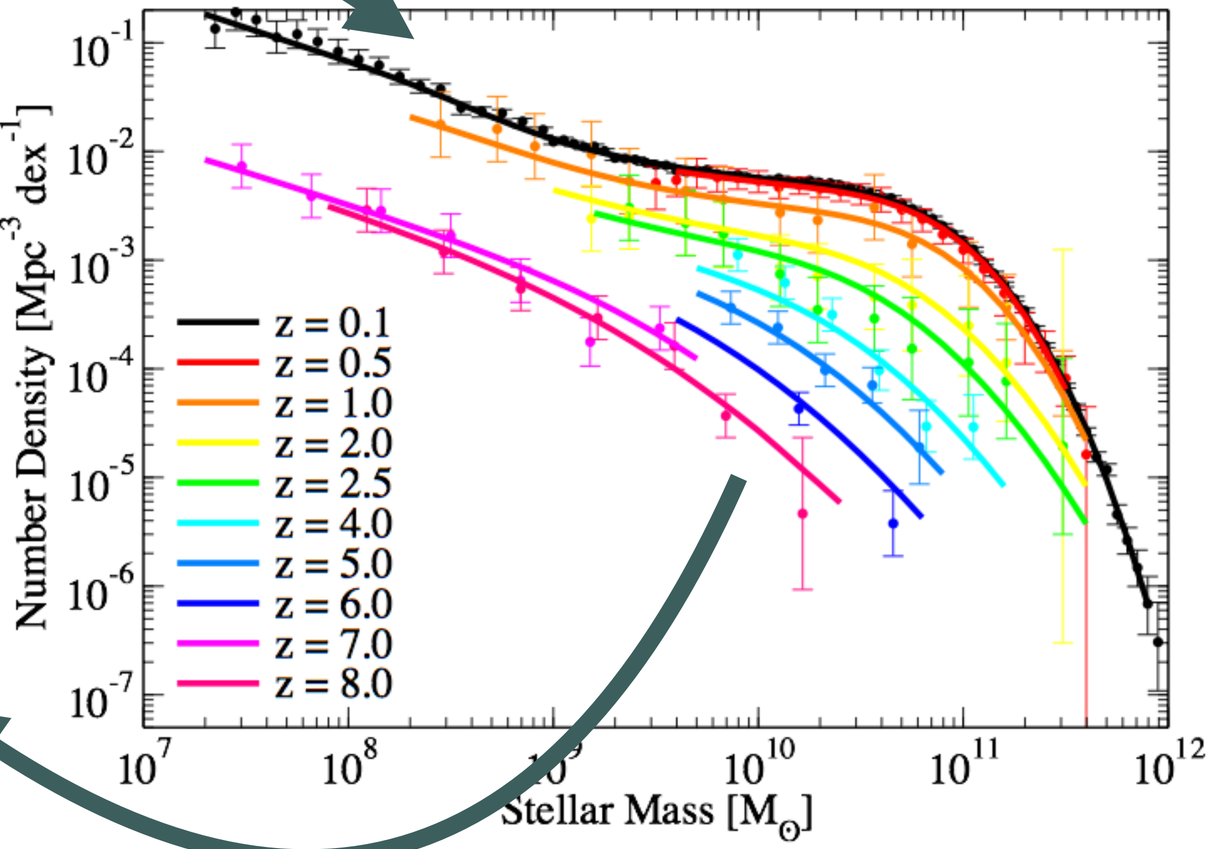
At z>0?

MCMC

Galaxy Formation Efficiency



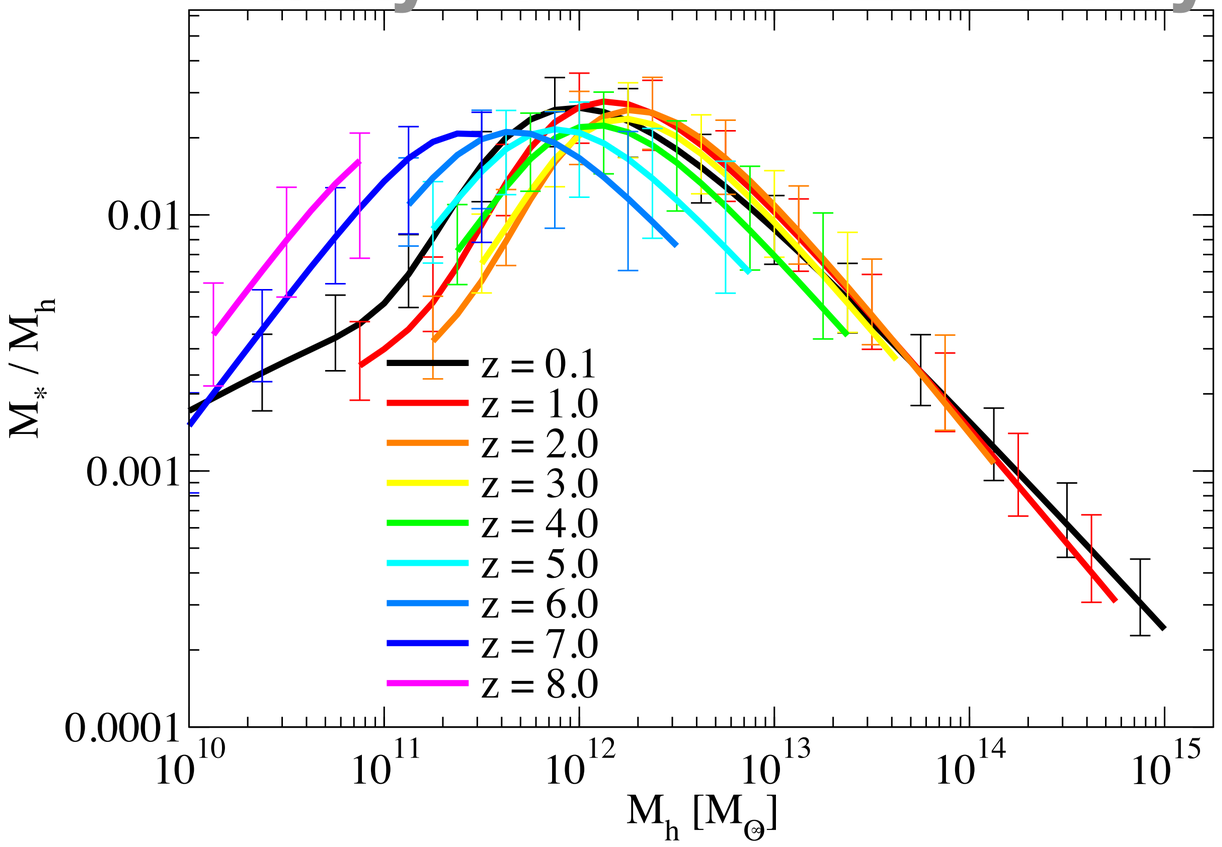
of Galaxies

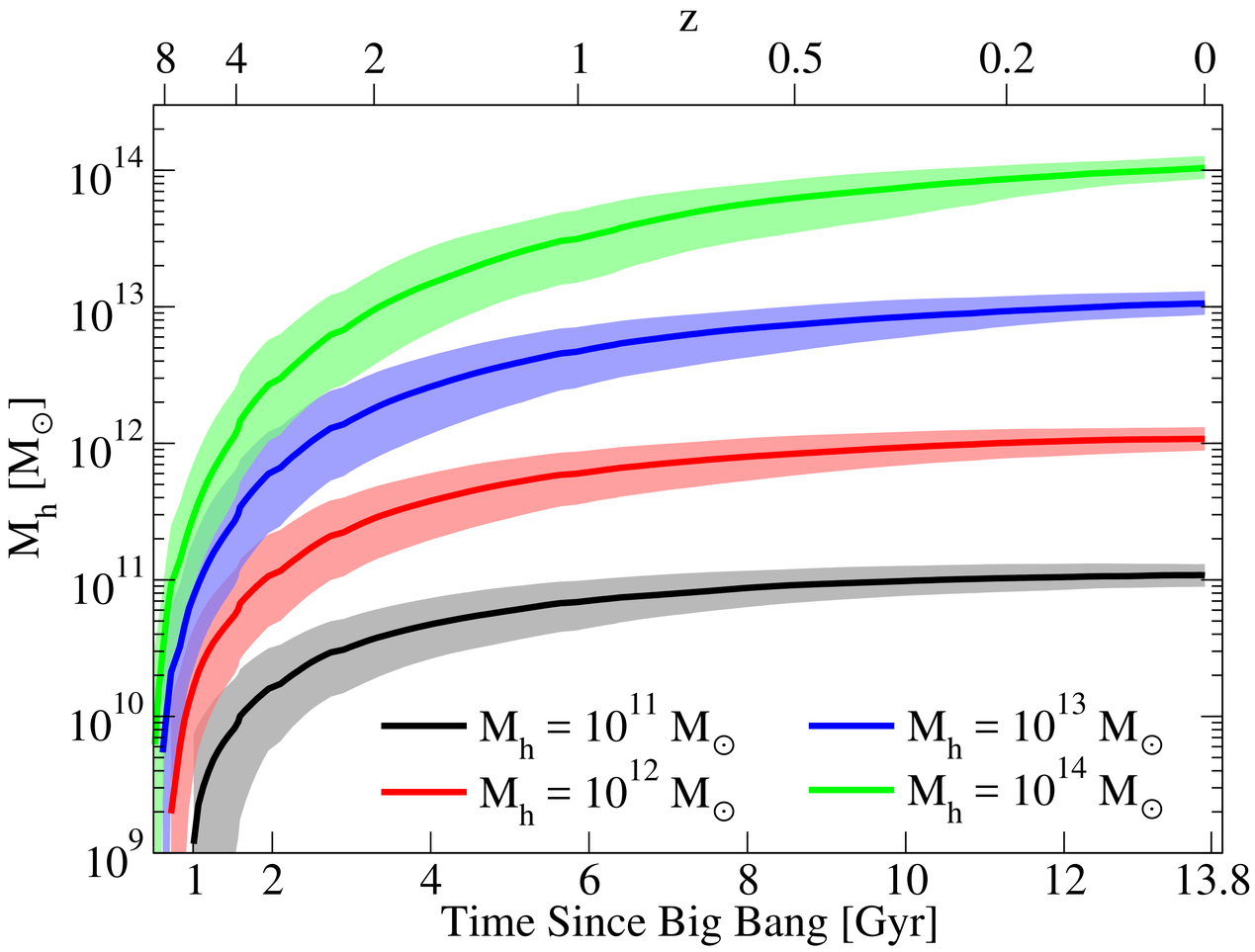


PB et al. (2013)

Galaxy Formation Histories

Galaxy Formation Efficiency

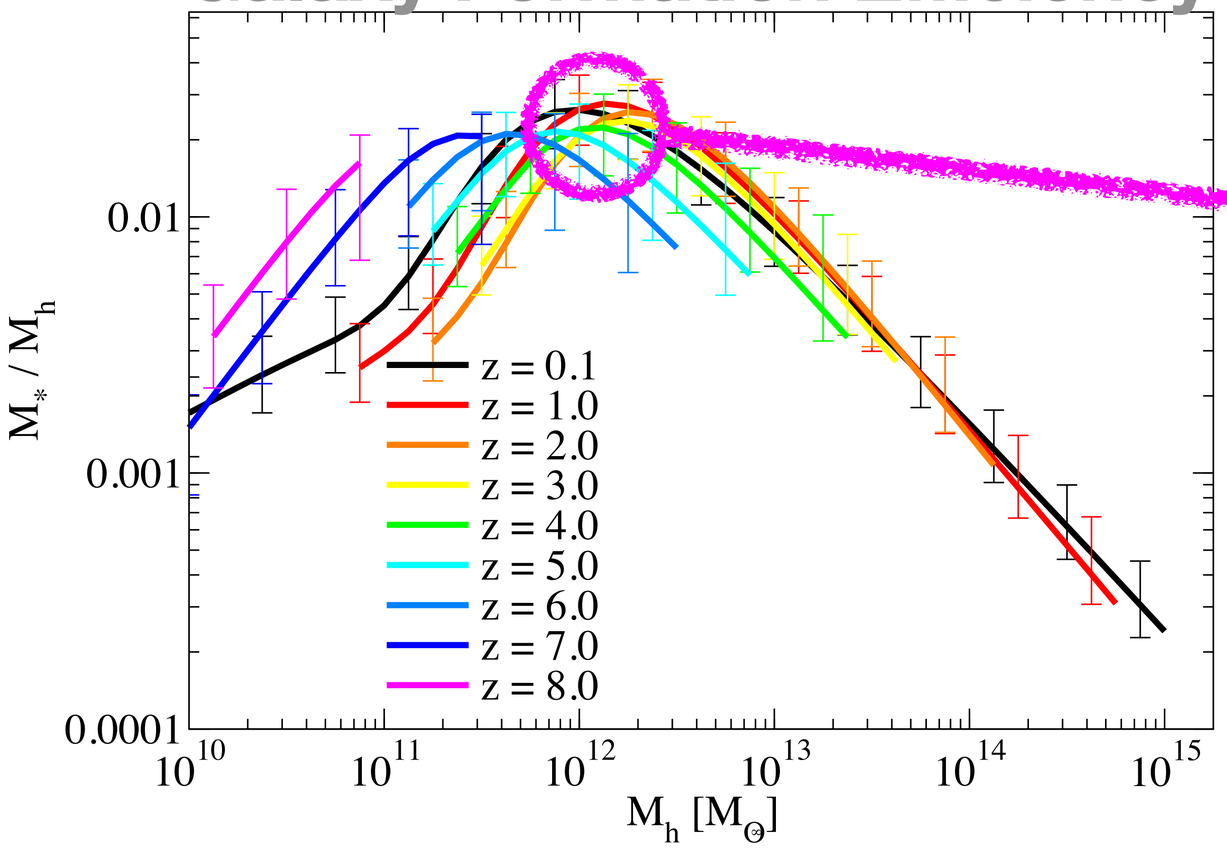


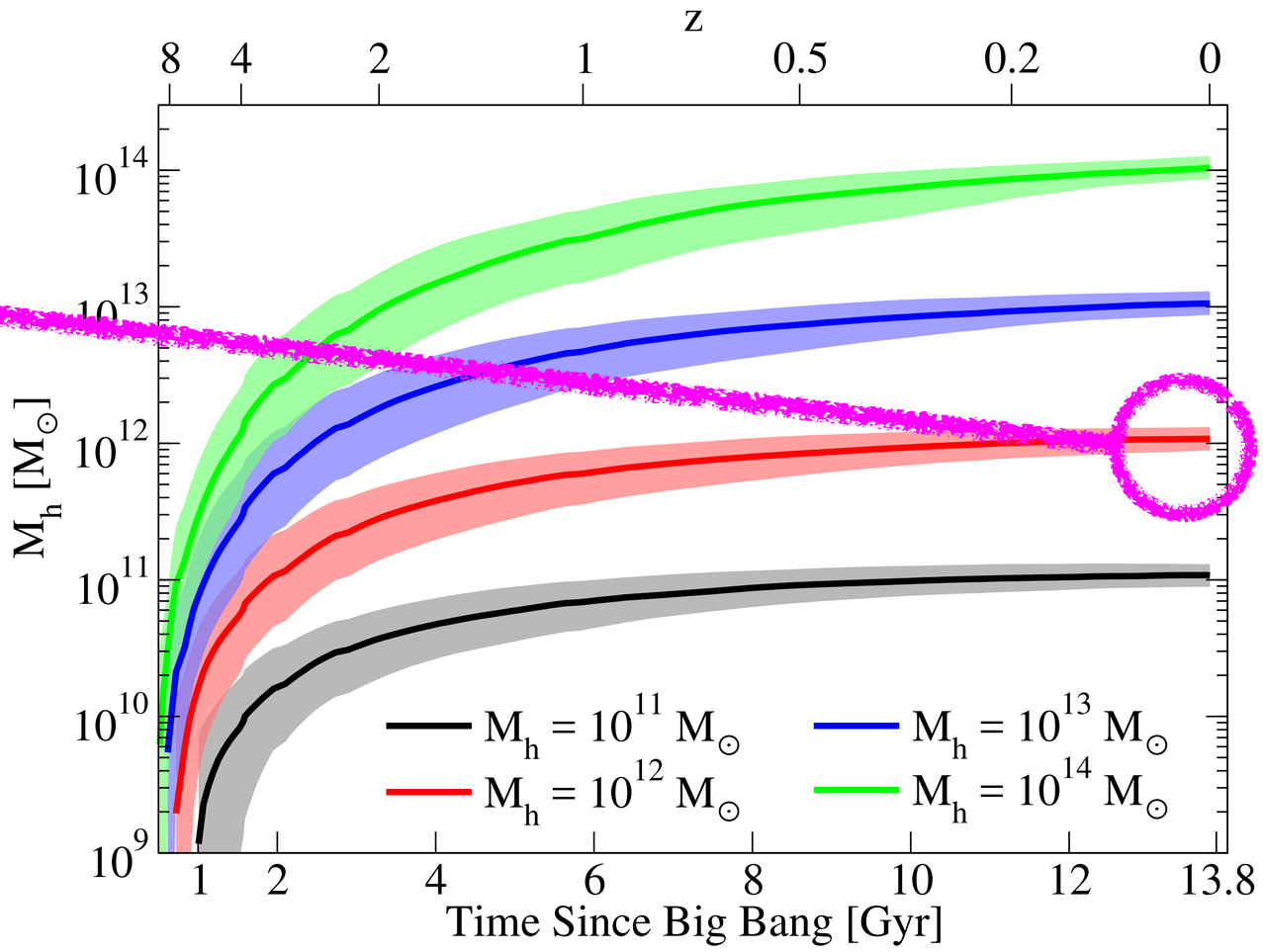


PB et al. (2013)

Galaxy Formation Histories

Galaxy Formation Efficiency

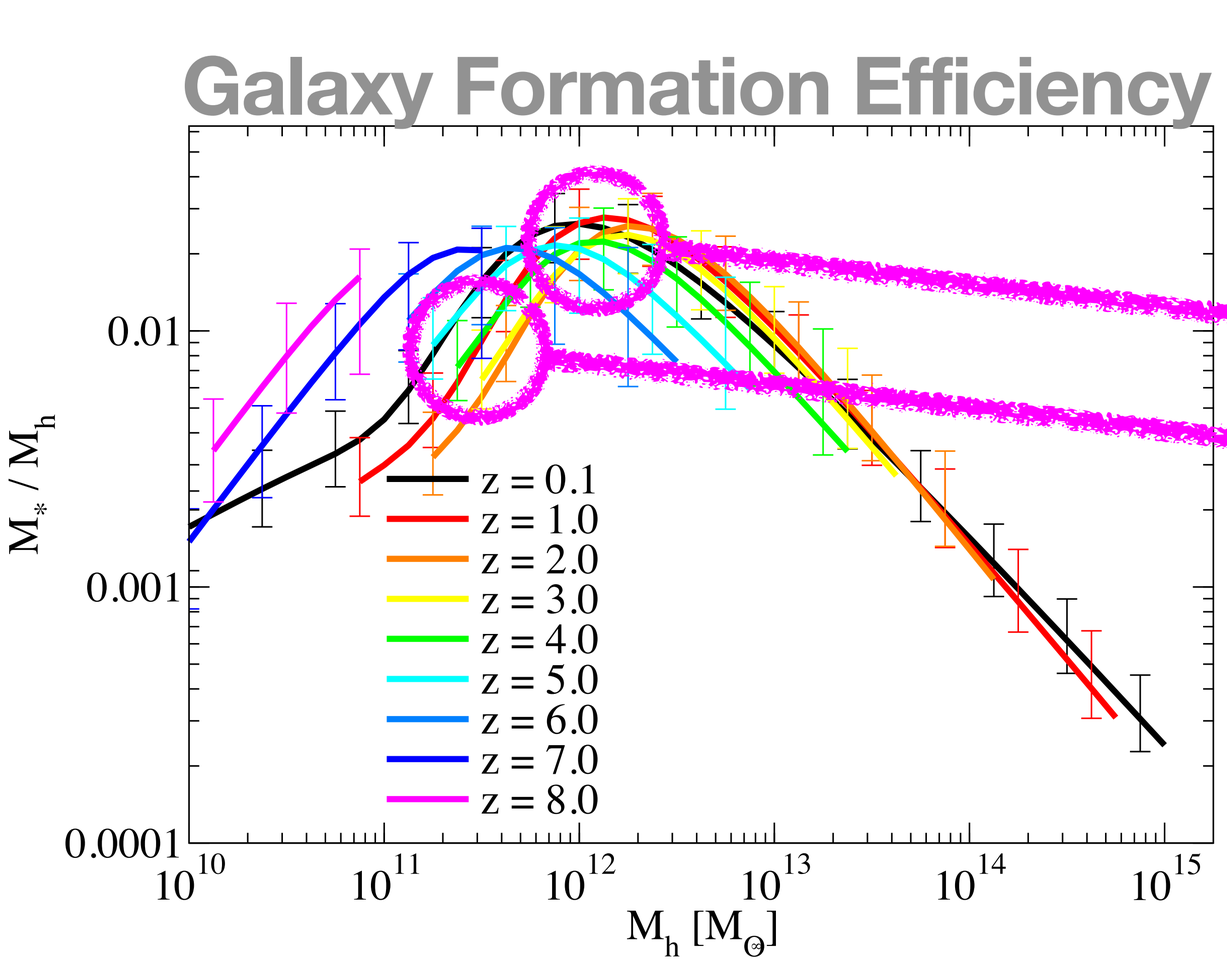


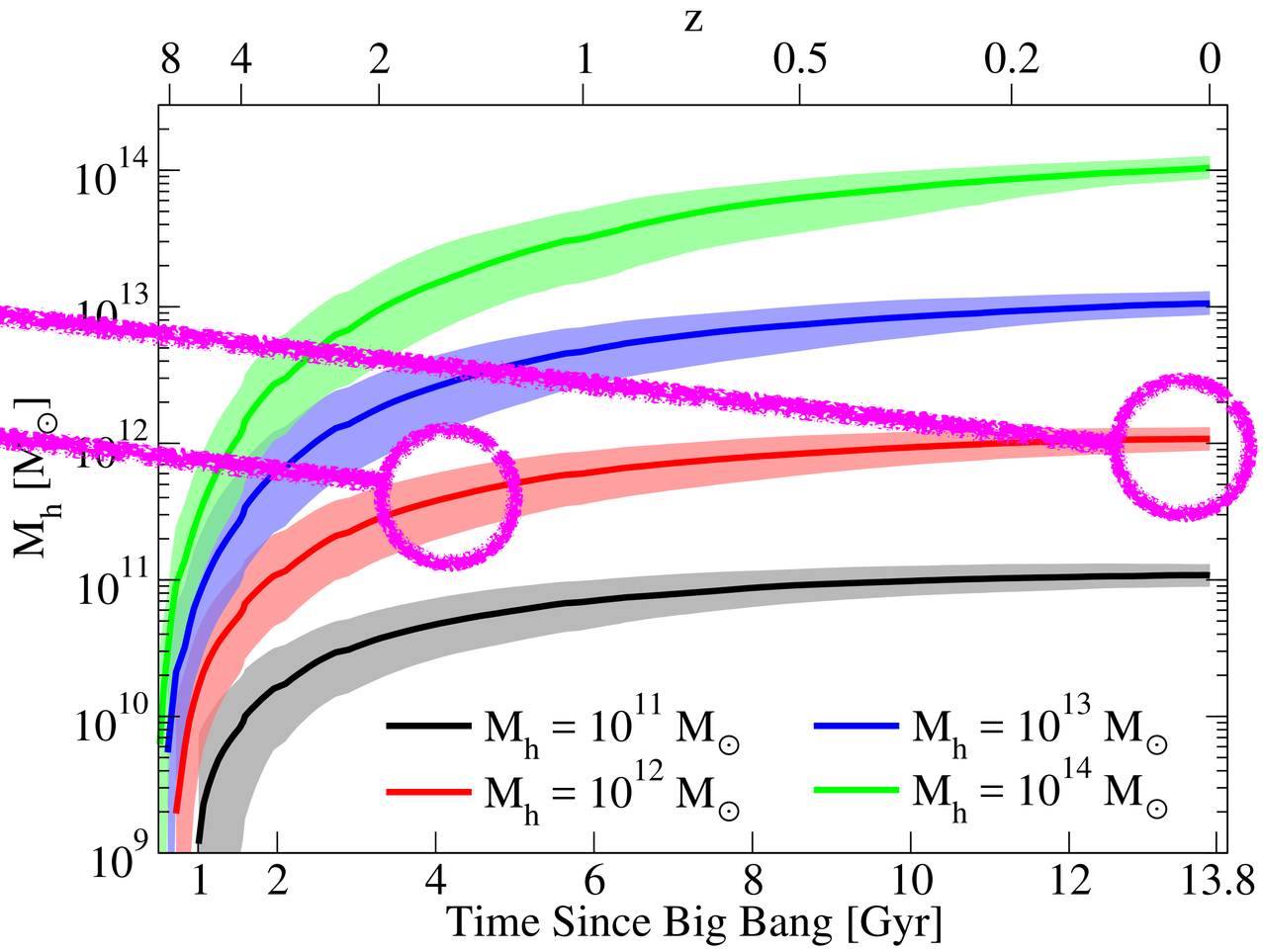


PB et al. (2013)

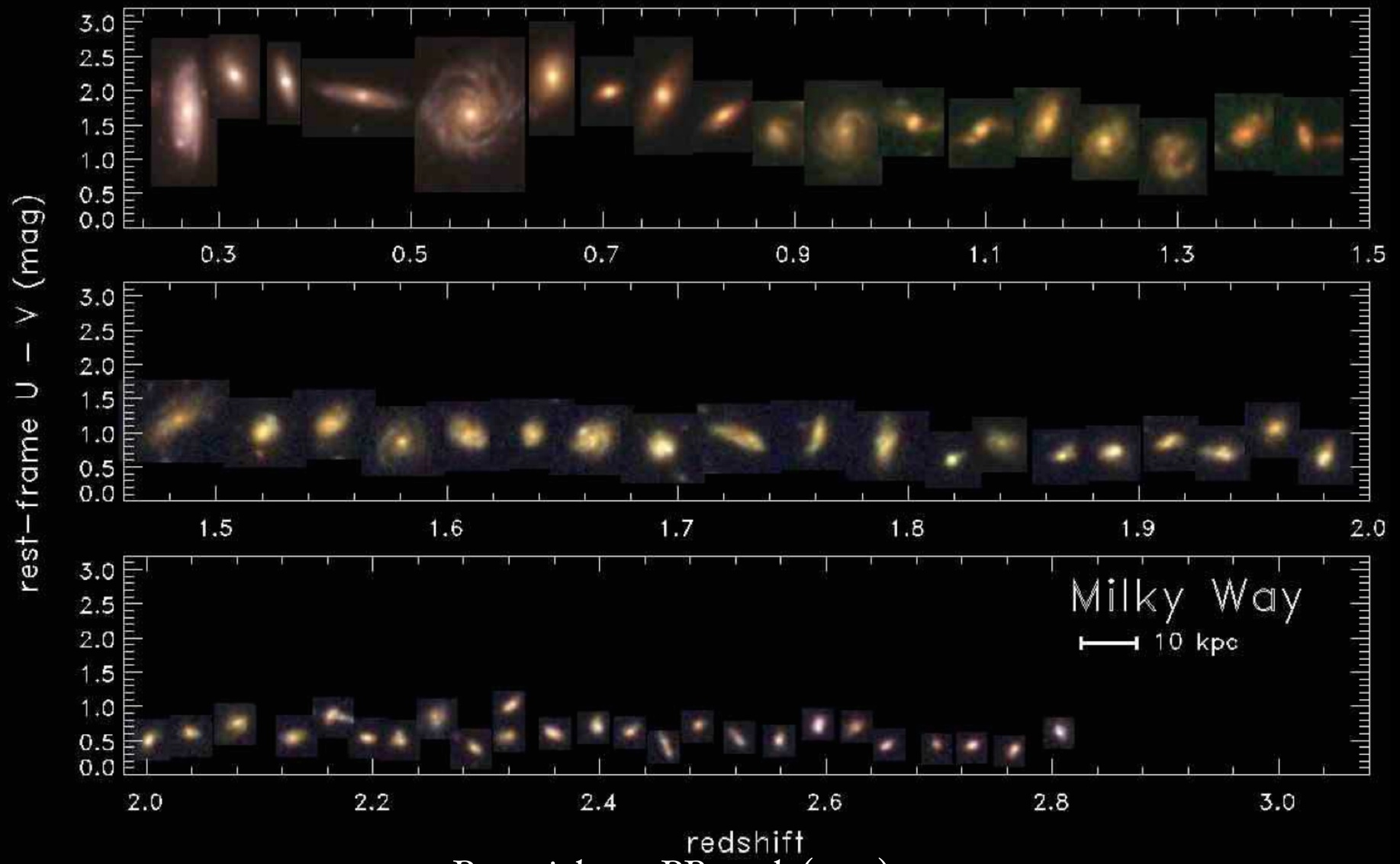
Galaxy Formation Histories

Galaxy Formation Efficiency

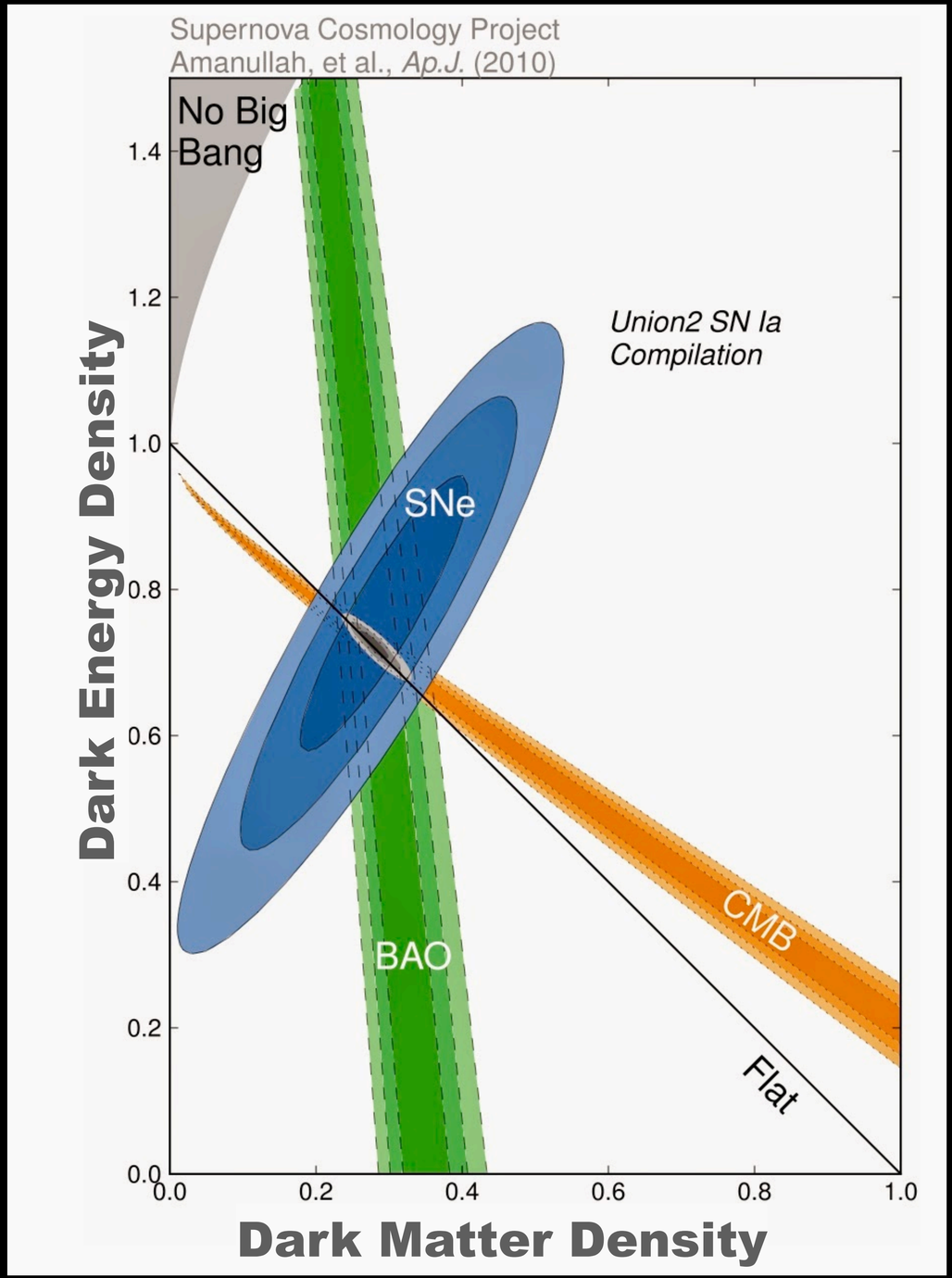




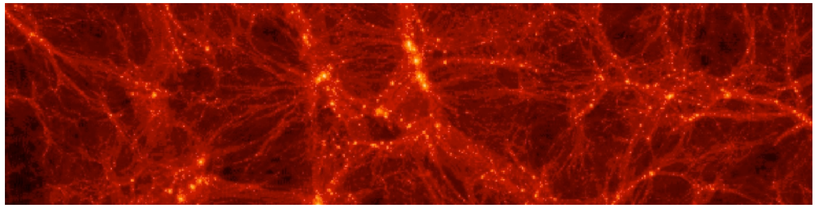
PB et al. (2013)



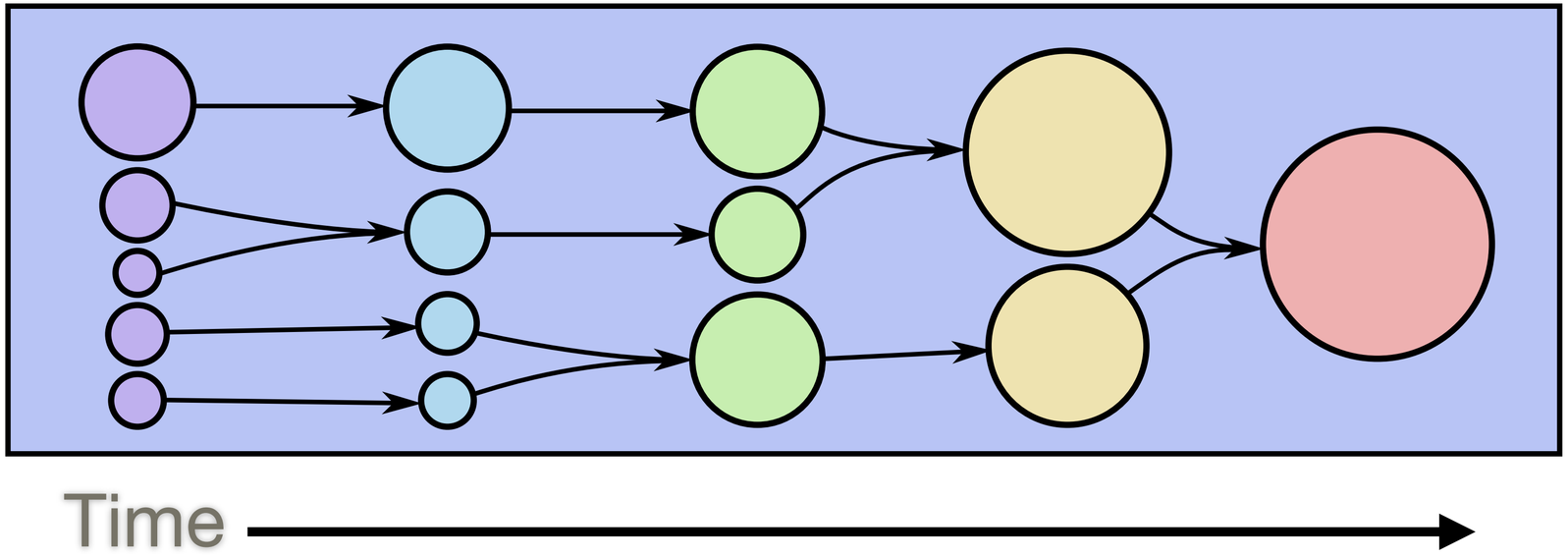
Papovich, ..., PB et al. (2015)


Supernova Cosmology Project
Amanullah, et al., Ap.J. (2010)
No Big Bang
Union2 SN Ia Compilation
SNe
BAO
CMB
Flat
Dark Energy Density
Dark Matter Density
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1.4

UniverseMachine

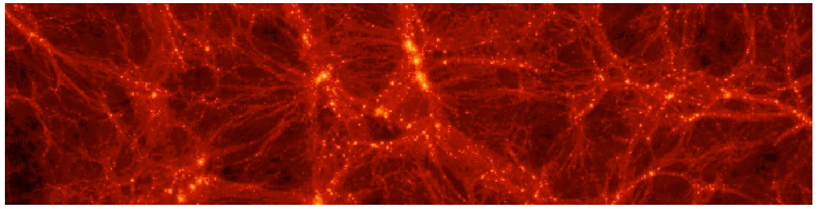


DM Simulation



PB+ (2019)

UniverseMachine



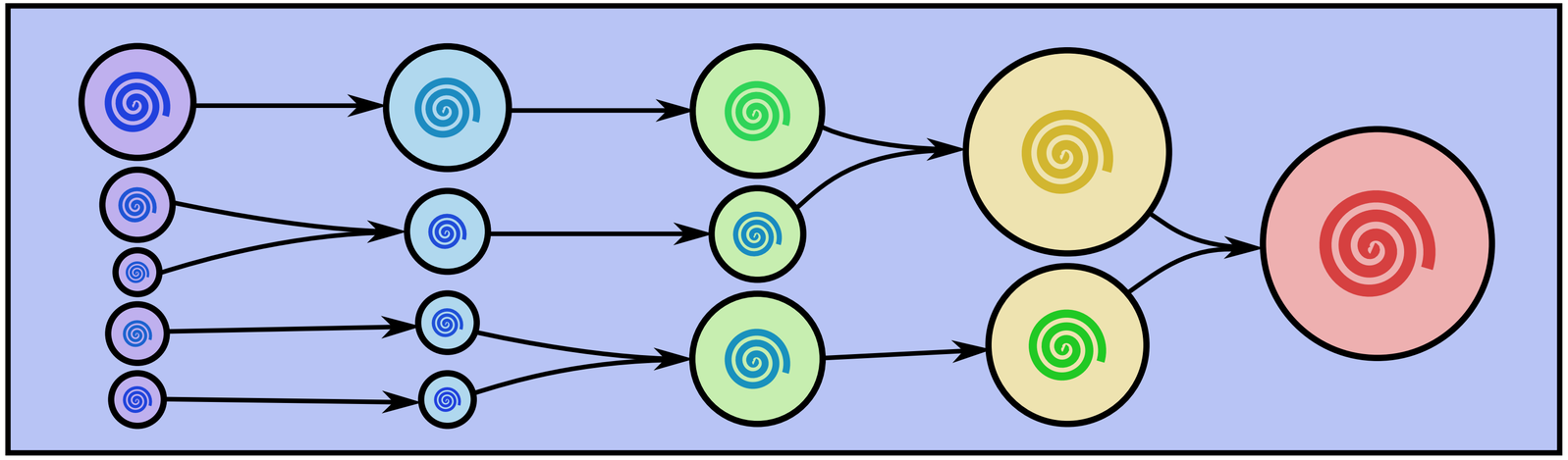
DM Simulation

+

$$\mathrm{SFR} = f(M_h, \dot{M}_h, z)$$

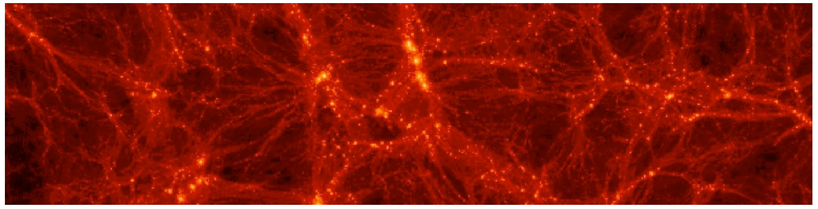
Flexible Fitting Function

=



PB+ (2019)

UniverseMachine



DM Simulation

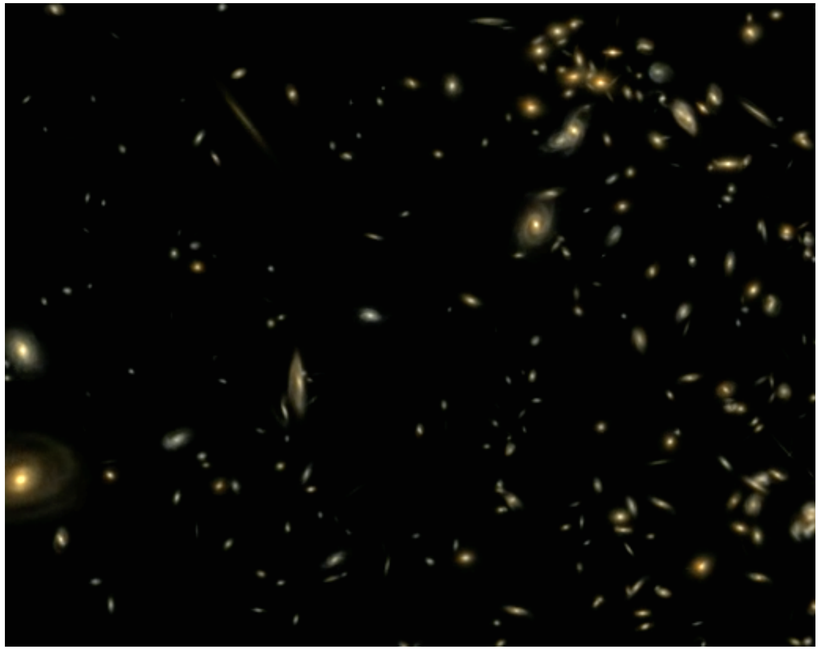
+

$$\mathrm{SFR} = f(M_h, \dot{M}_h, z)$$

Flexible Fitting Function

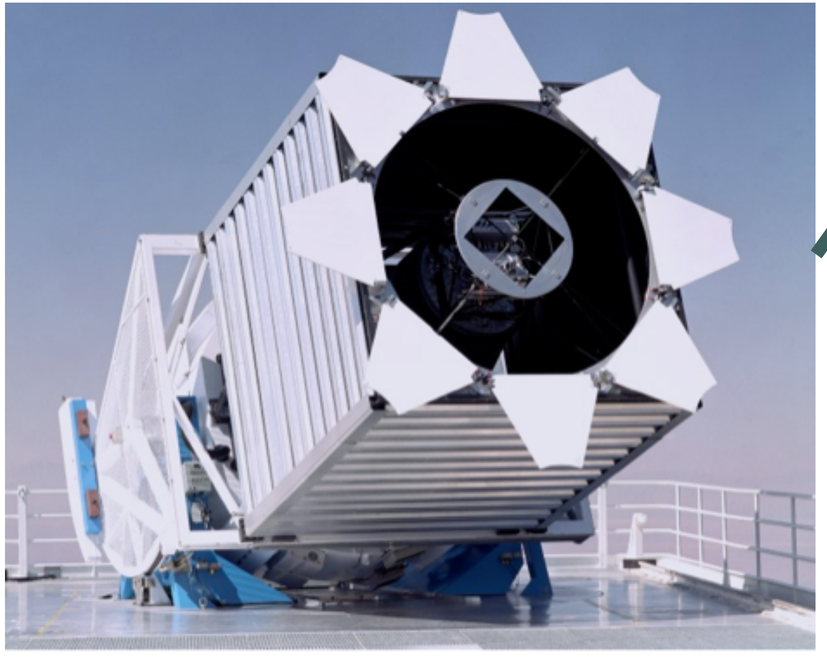
=

MCMC



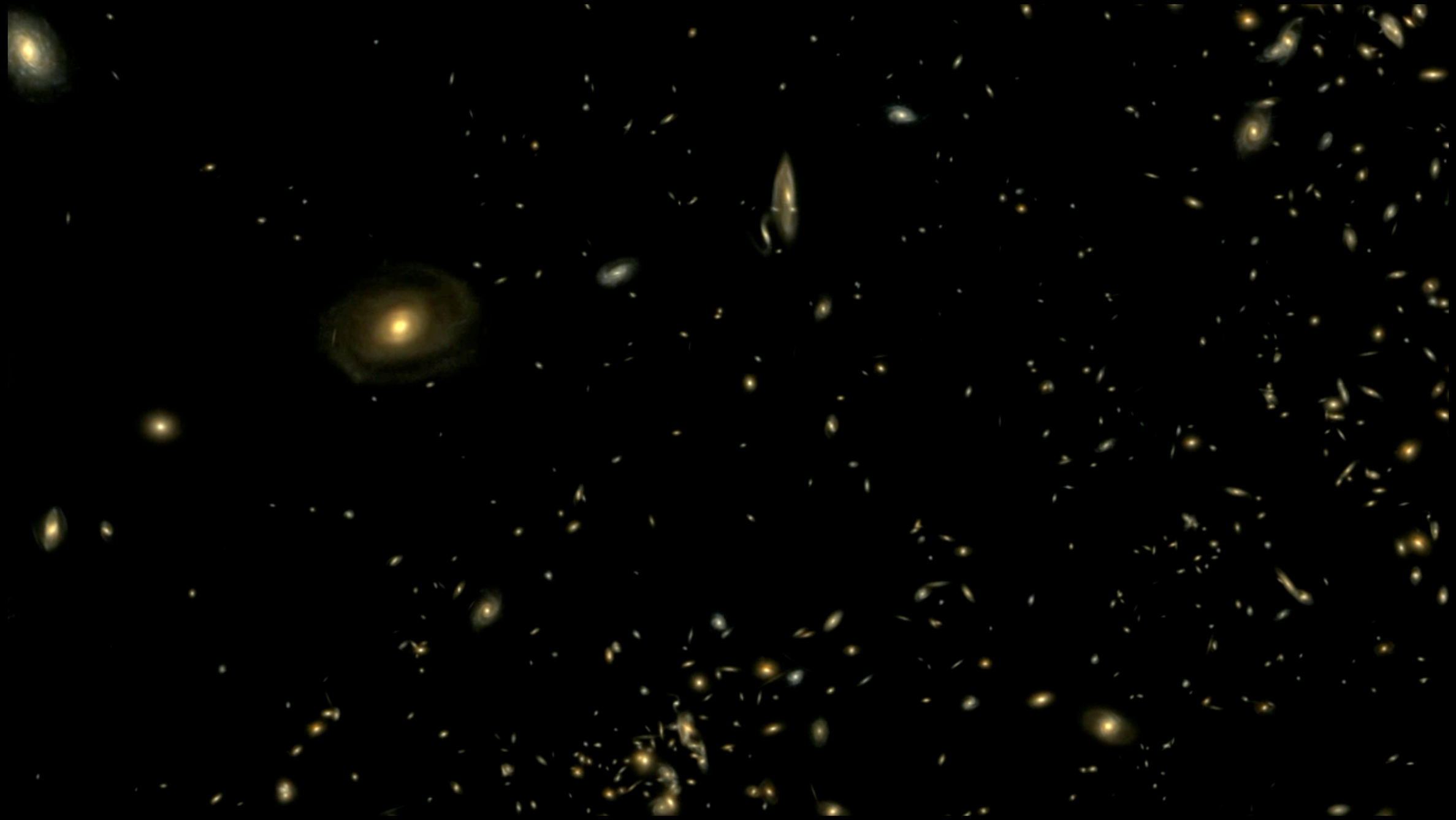
Mock Universe

?
=

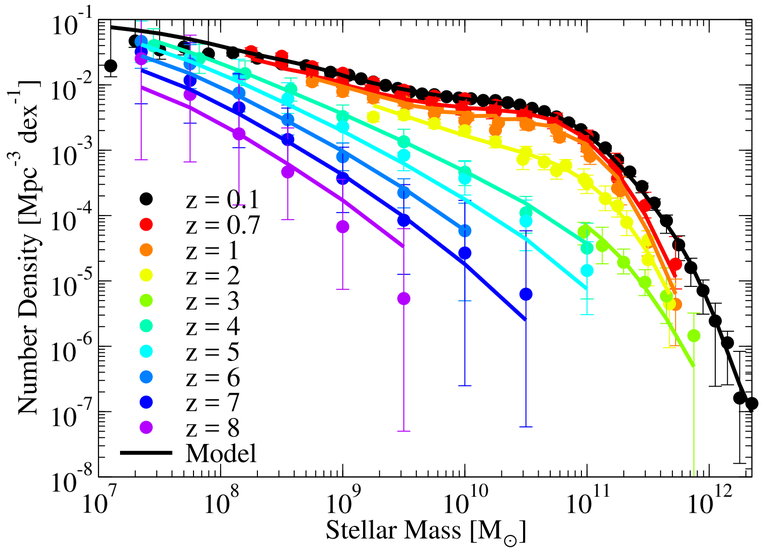


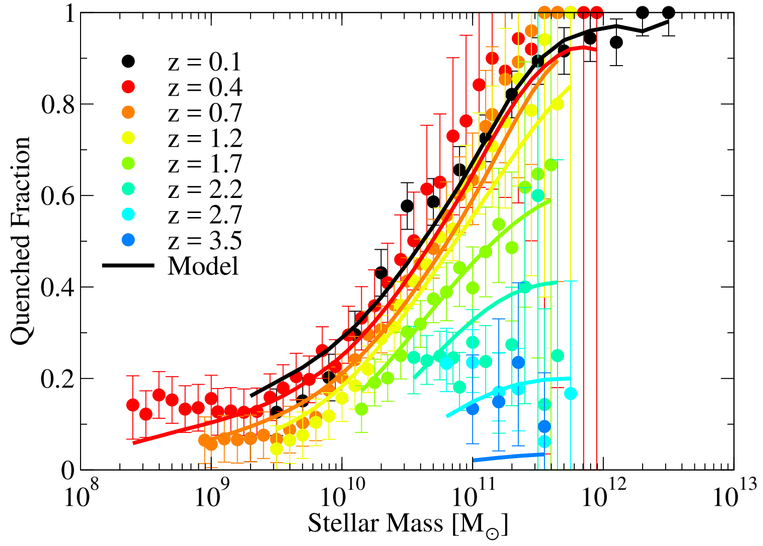
Real Universe

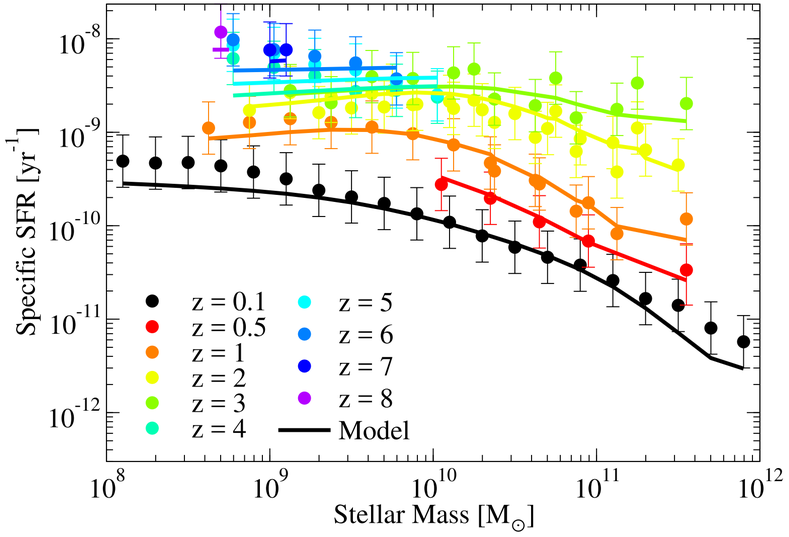
PB+ (2019)

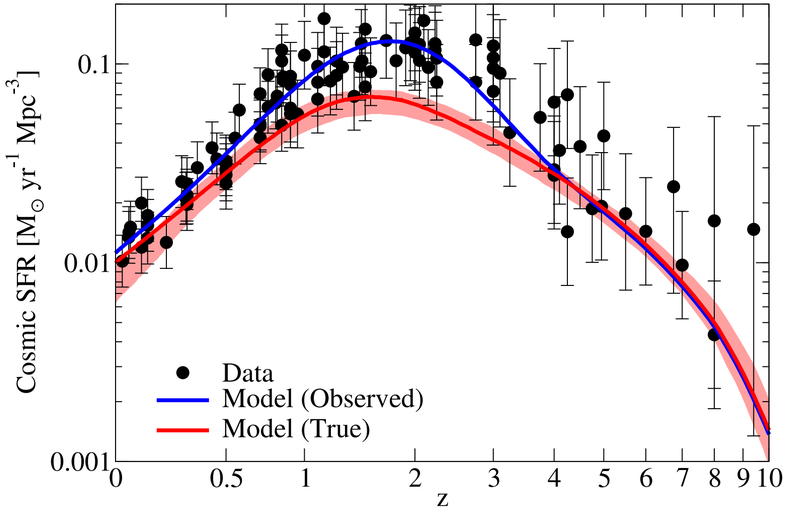


UniverseMachine



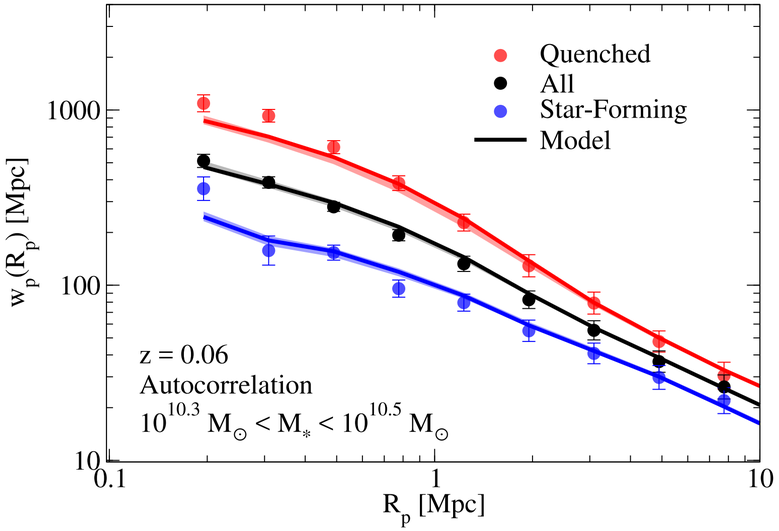


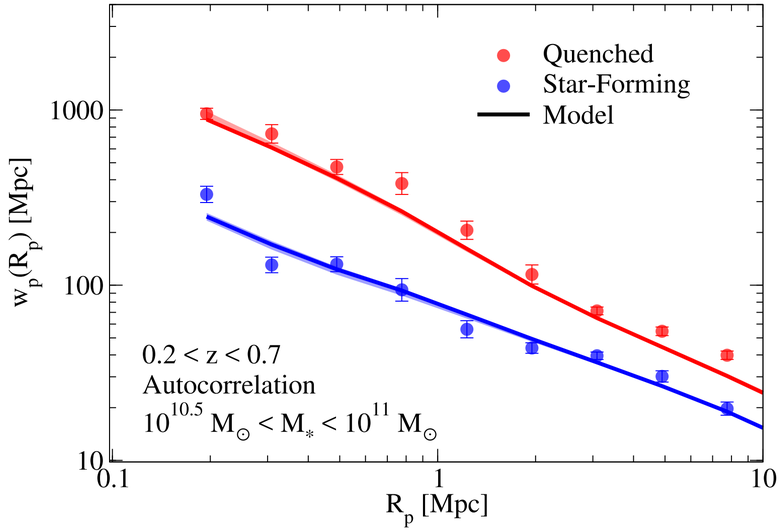


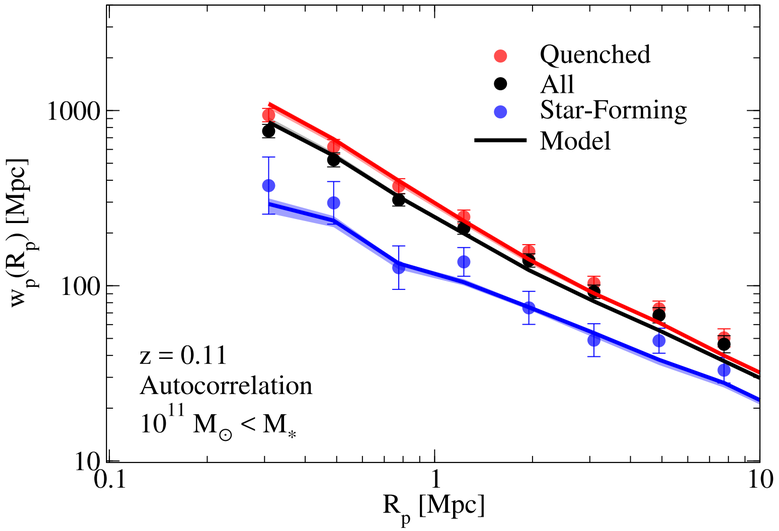


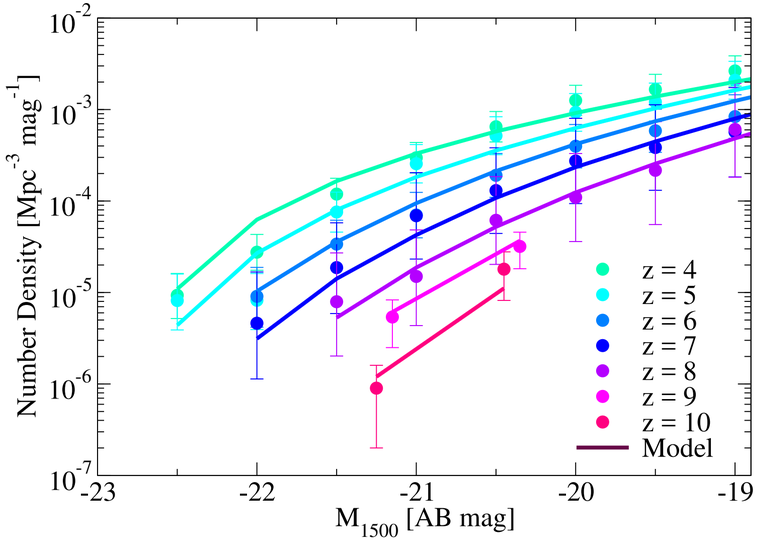
PB+ (2019)

UniverseMachine





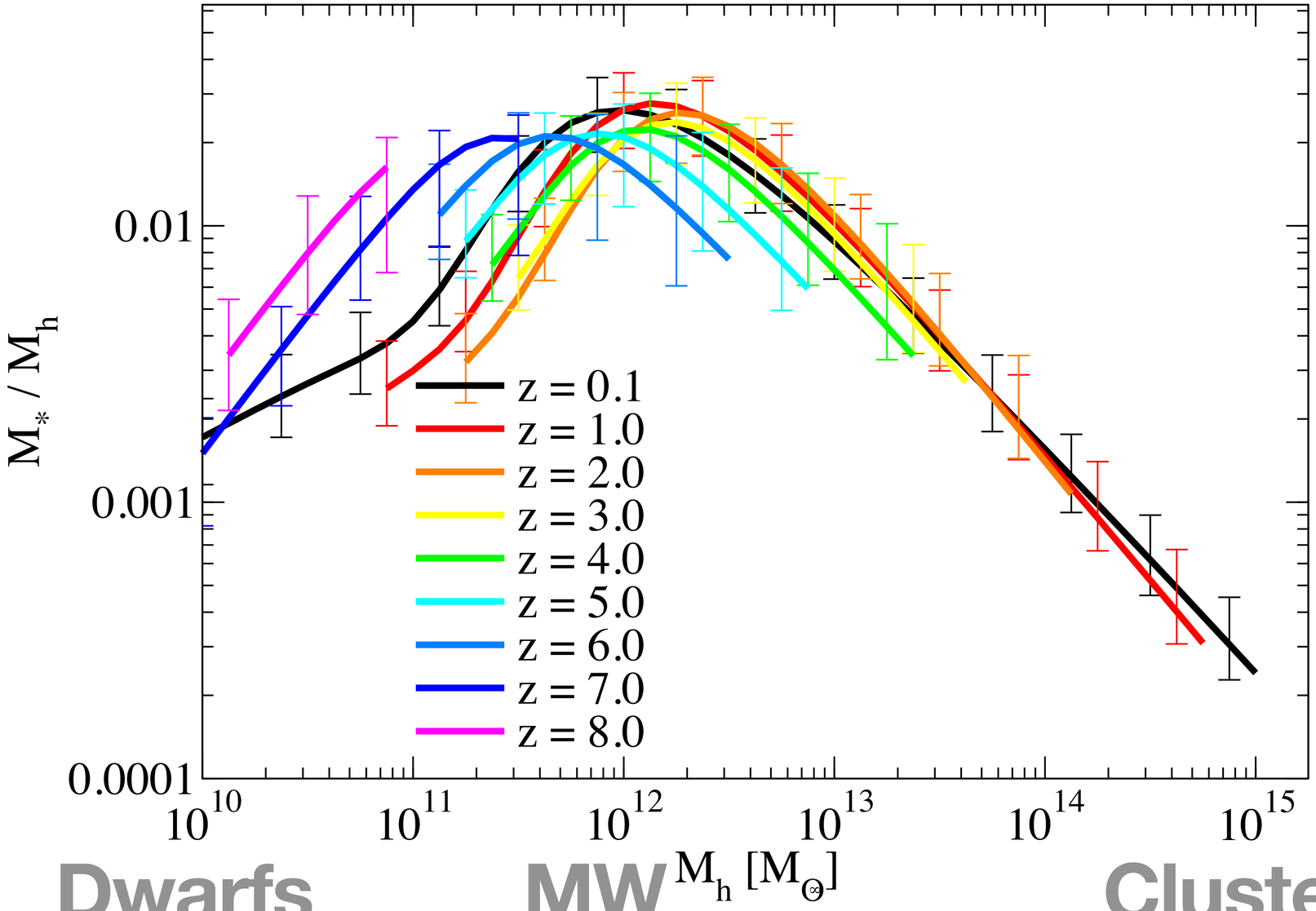




PB+ (2019)

Unique Constraints

Efficiency of Star Formation



Dwarfs MW Clusters

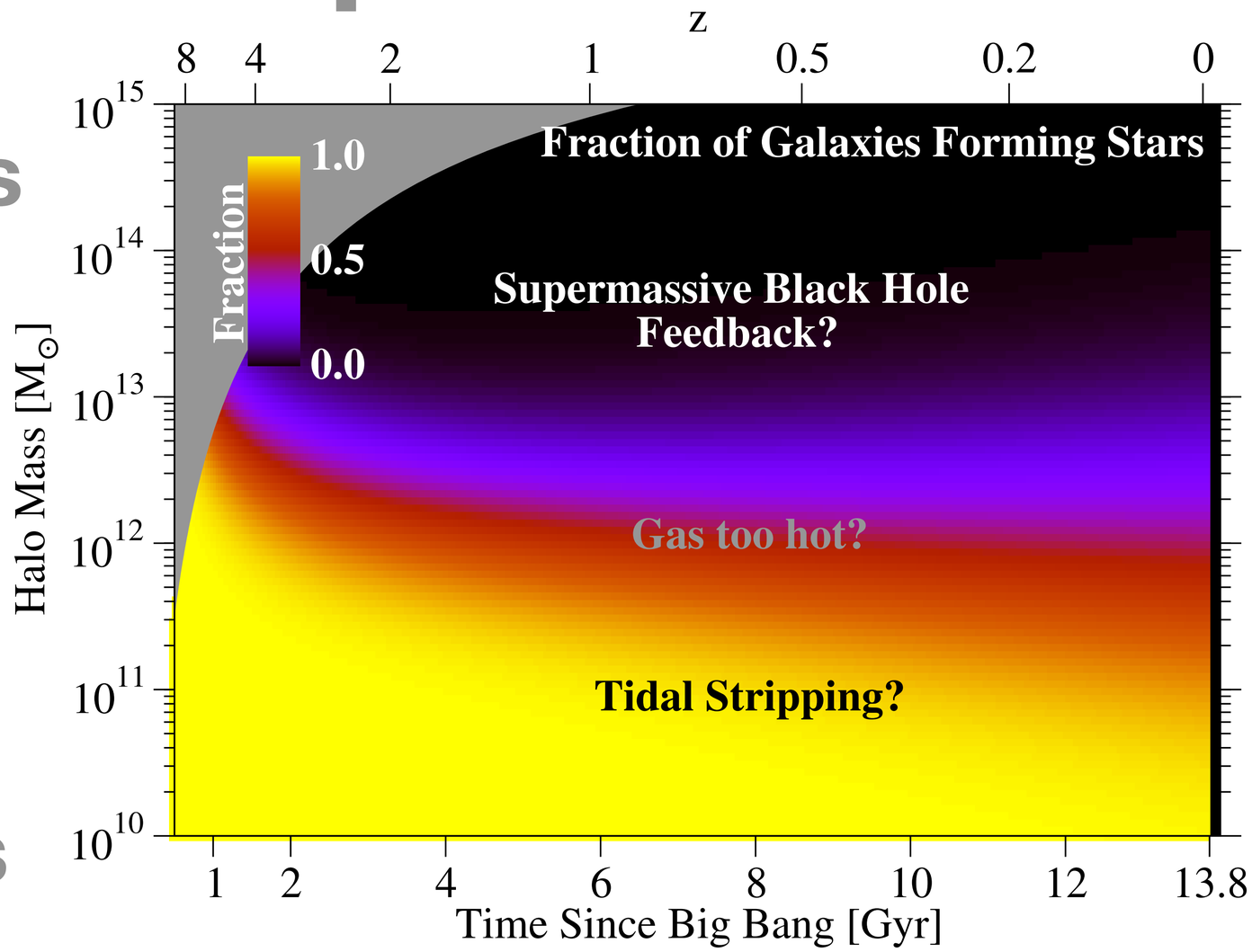
PB+ (2013)

Unique Constraints

Clusters

MW

Dwarfs



PB+ (2019)

Biggest Open Questions

How does gas cycle in/out?

How do galaxy shapes/sizes evolve?

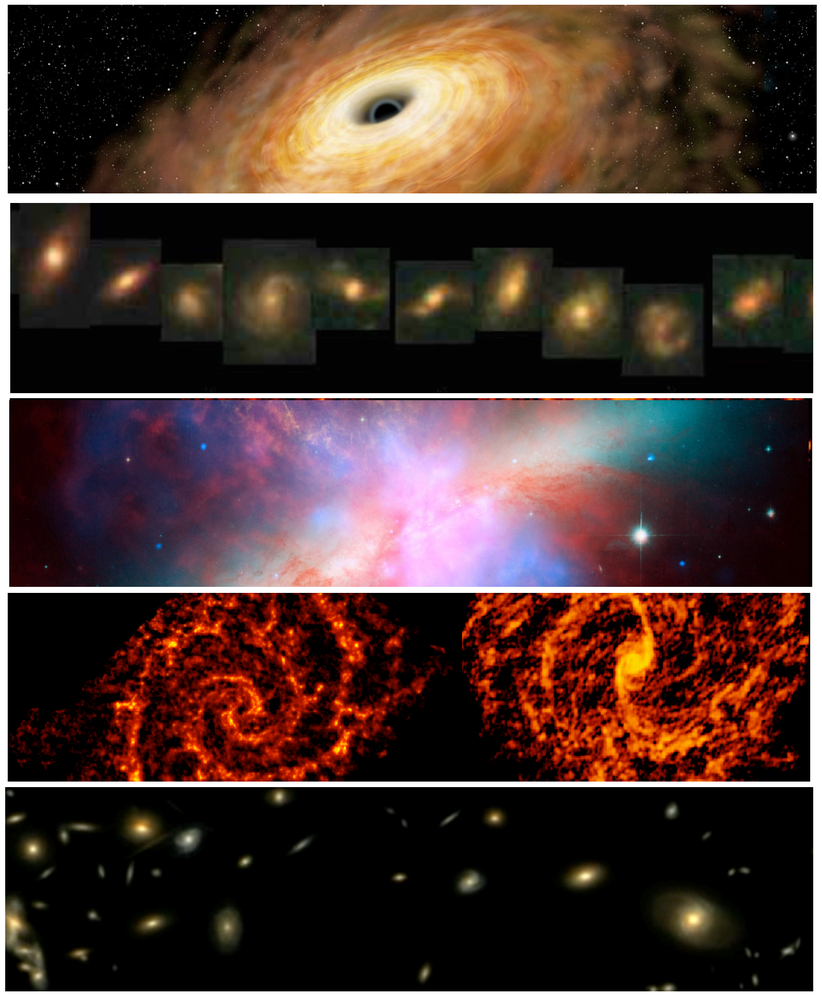
How do SMBHs grow?

Why do galaxies stop forming stars?

How does dust influence what we see?

What will future telescopes see?

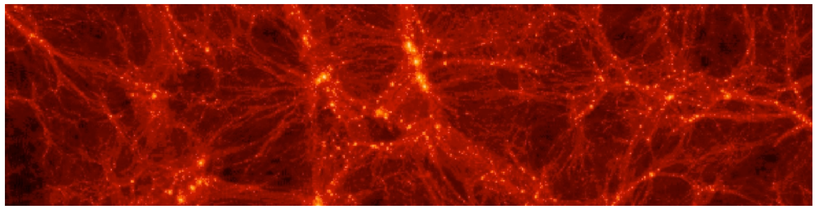
Empirical Galaxy Models



SMBHs	Trinity (Zhang et al. 2023ab, 2024ab)
Dust/Color	UniverseMachine 2 (Cooray et al. in prep.)
Metallicity	ChemicalUniverseMachine (Nishigaki et al., 2025)
HI, H_2 Masses	NeutralUniverseMachine (Guo et al. 2023)
Galaxy Masses	UniverseMachine (PB+ 2019); UM-SAGA (Wang et al. 2024)

Image credits: https://www.universetoday.com/108702/what-fuels-the-engine-of-a-supermassive-black-hole/;
Papovich et al. 2015; Armus et al. 2020; HST/Spitzer/Chandra; SDSS

Trinity



DM Simulation

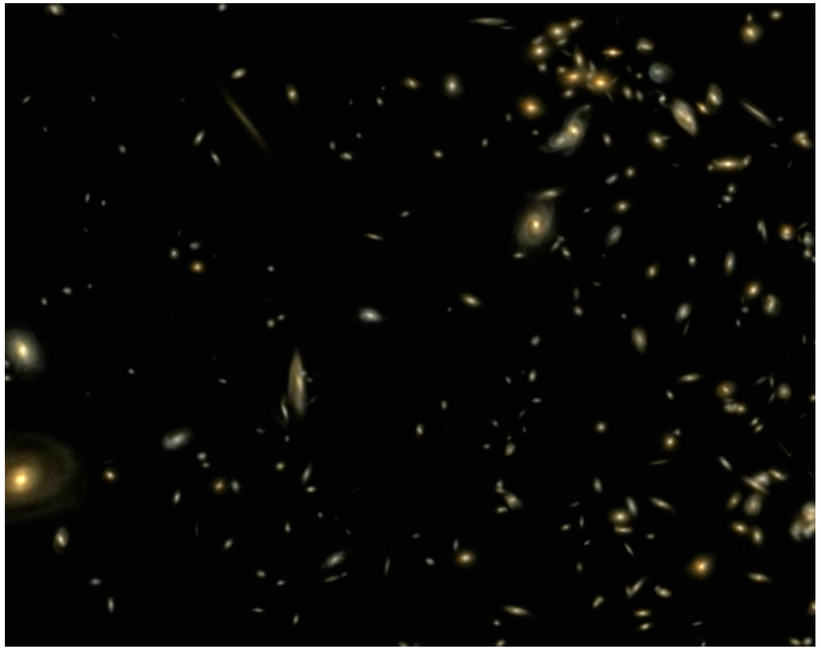
+

$$\mathrm{SFR} = f(M_h, \dot{M}_h, z)$$
$$P(\dot{M}_\bullet) = g(M_*, z)$$

Galaxy+BH Fit. Function

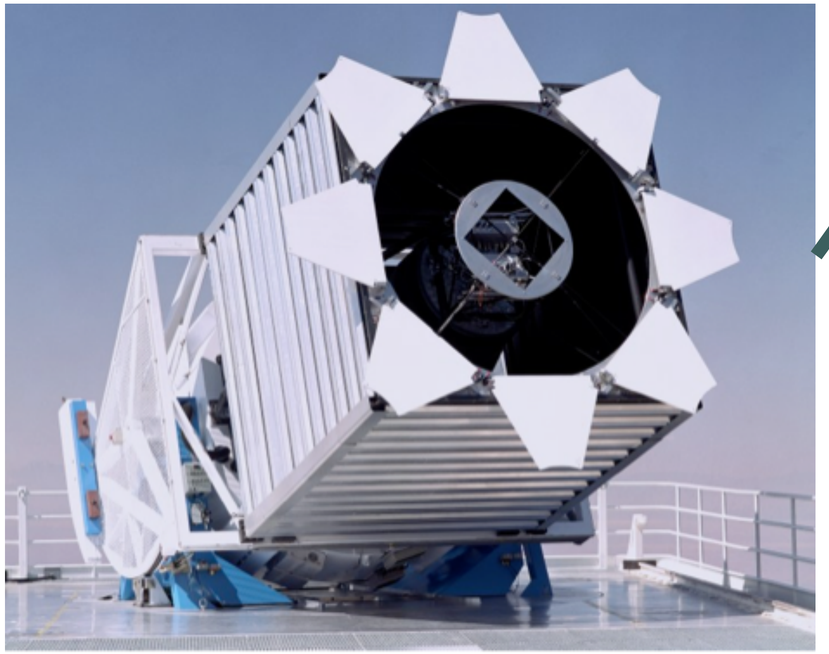
=

MCMC



Mock Universe

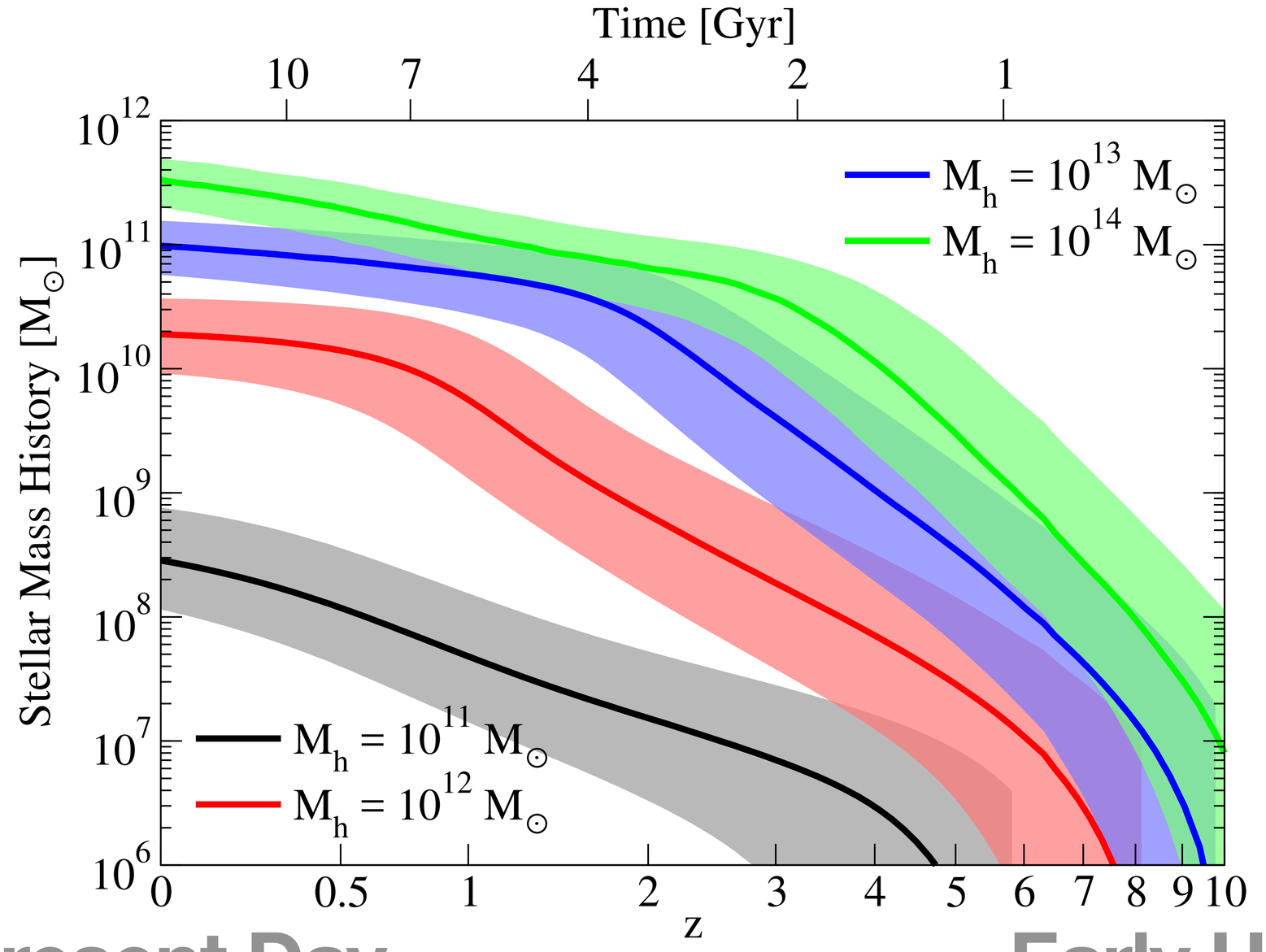
?
=



Real Universe

HZ, PB, et al.: https://arxiv.org/abs/2105.10474

UniverseMachine Results



Mass in Galaxy Progenitor

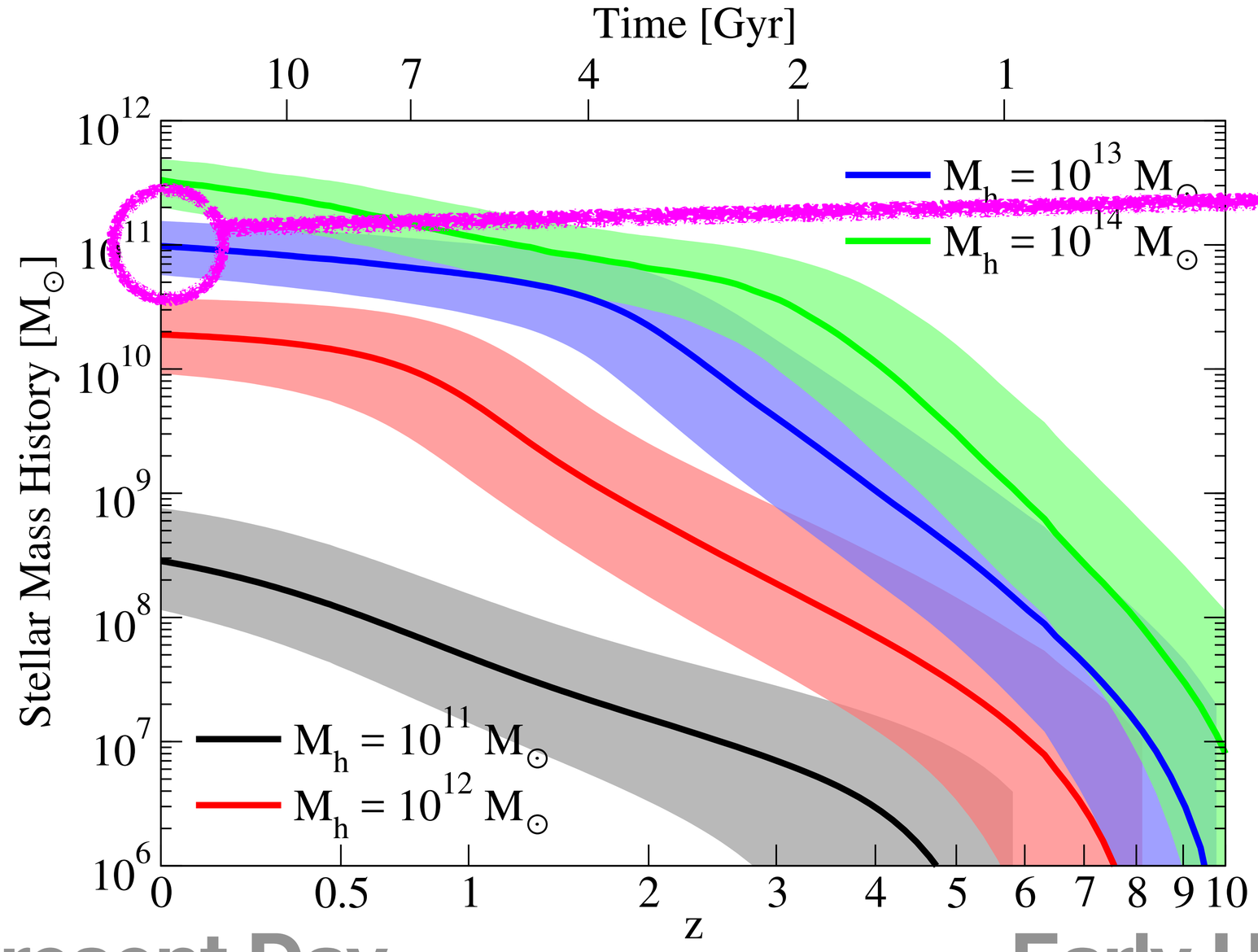
Present Day

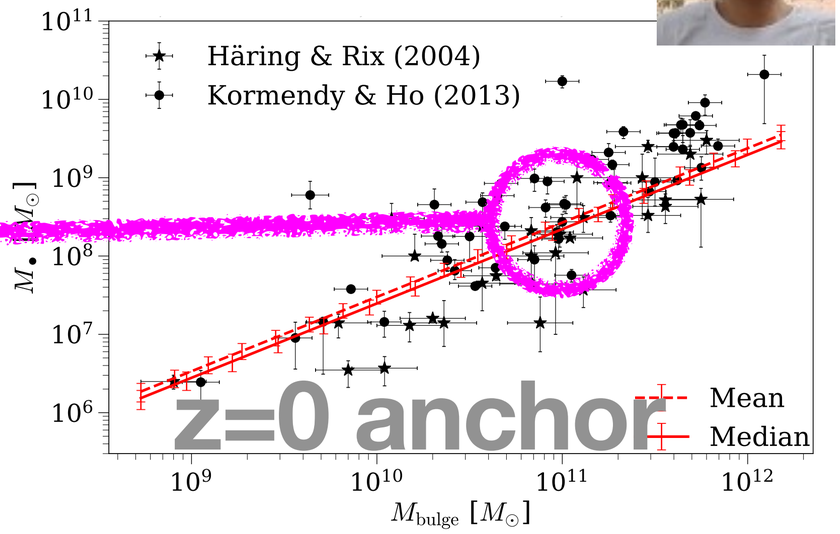
Early Universe

PB+ (2019)

UniverseMachine Results

Mass in Galaxy Progenitor





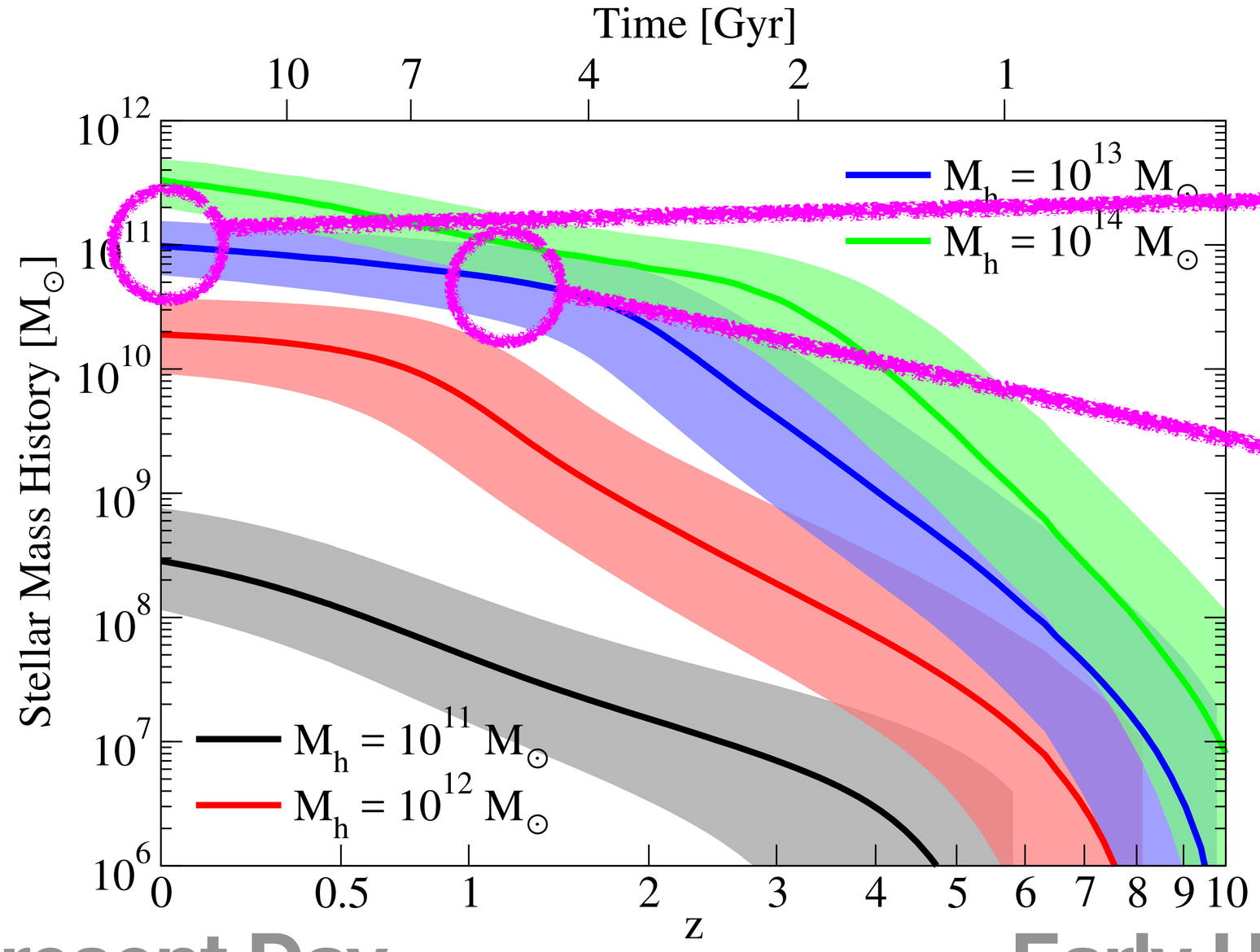
Present Day

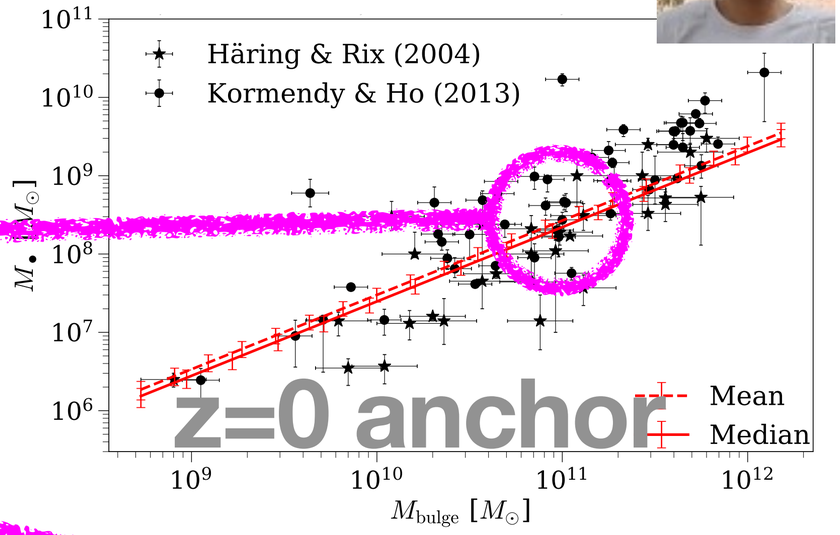
Early Universe

PB+ (2019)

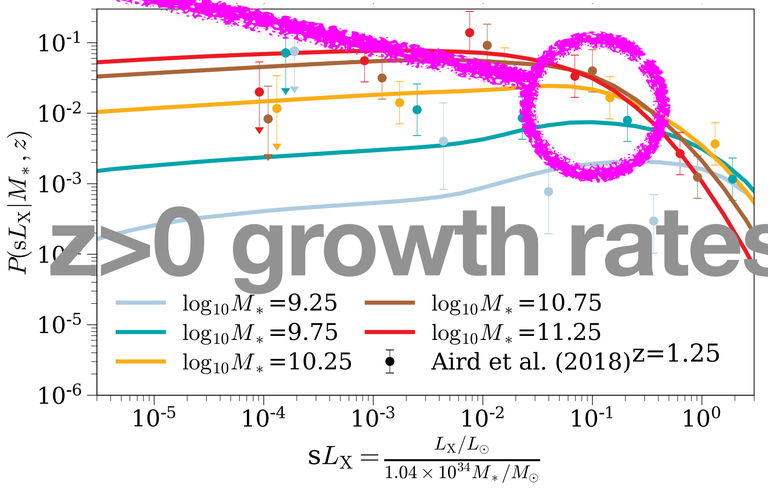
UniverseMachine Results

Mass in Galaxy Progenitor





z=0 anchor



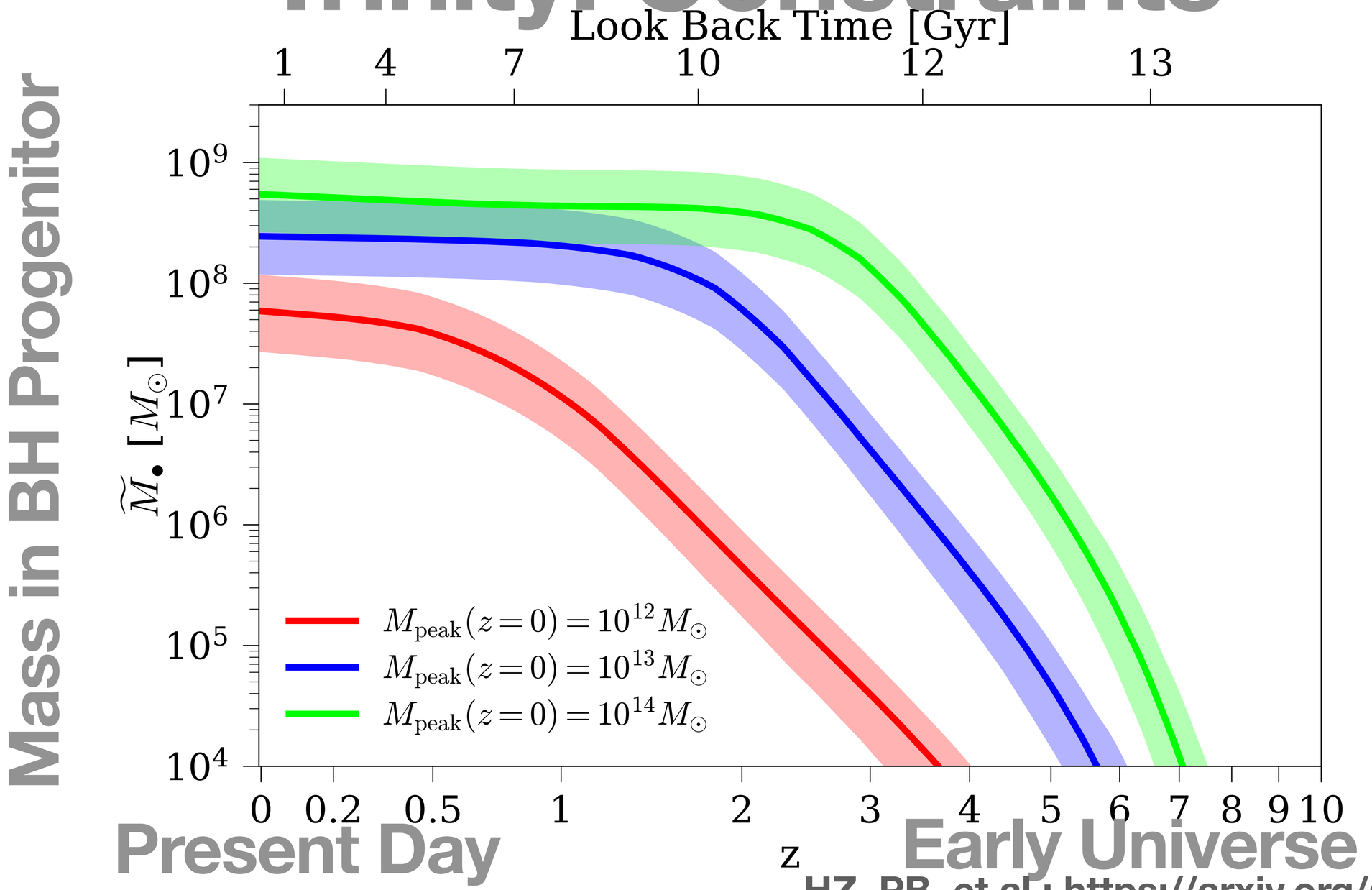
z>0 growth rates

Present Day

Early Universe

PB+ (2019)

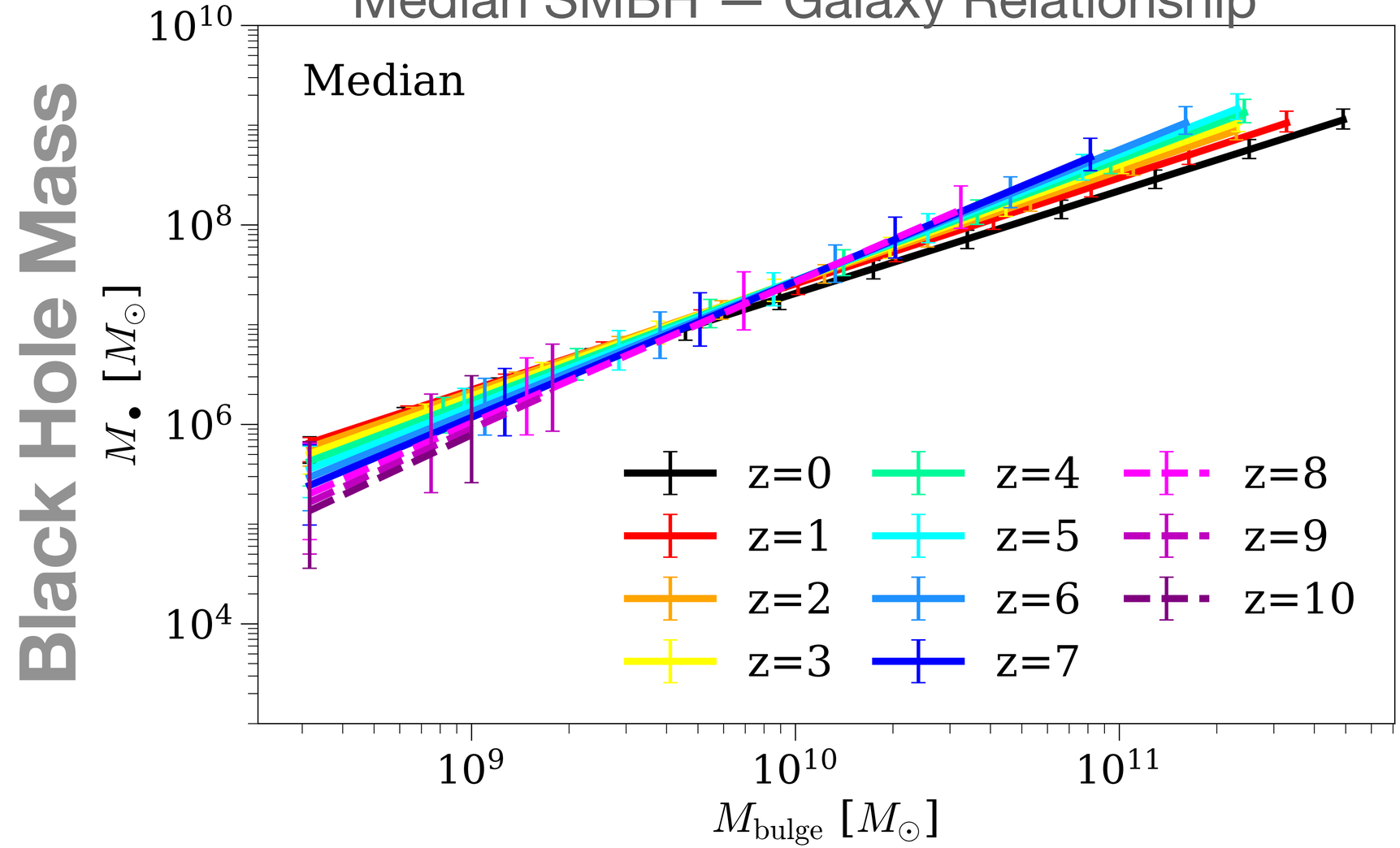
Trinity: Constraints



HZ, PB, et al.: https://arxiv.org/abs/2105.10474

Trinity: Results

Median SMBH — Galaxy Relationship

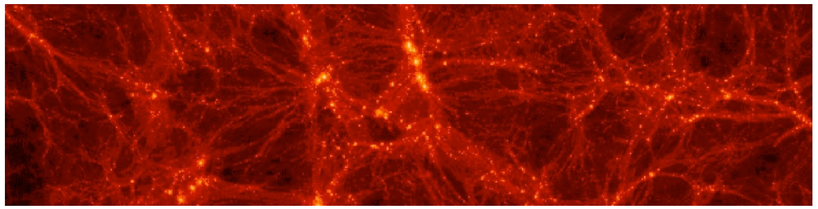


Black Hole Mass

Galaxy Central Mass

HZ, PB, et al.: https://arxiv.org/abs/2105.10474

Trinity



DM Simulation

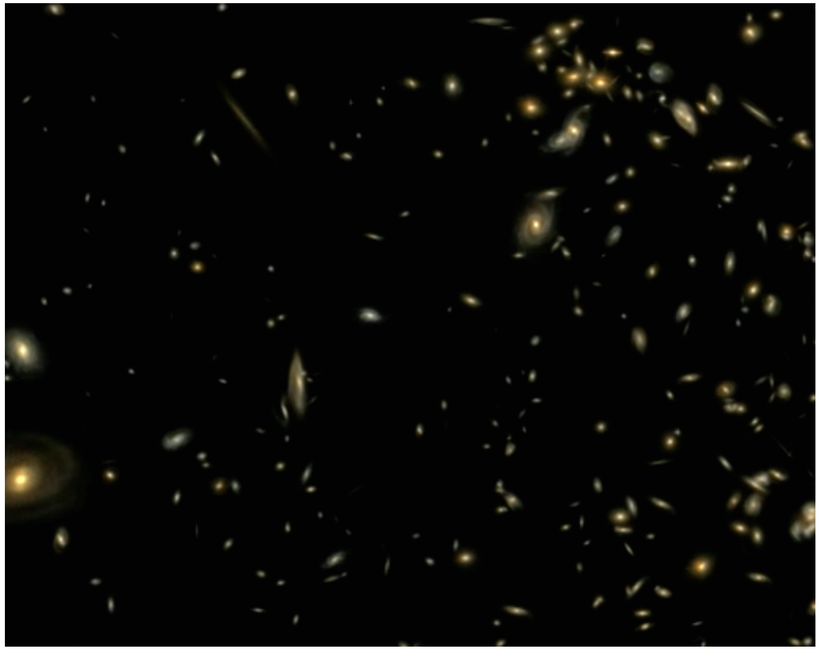
+

$$\mathrm{SFR} = f(M_h, \dot{M}_h, z)$$
$$P(\dot{M}_\bullet) = g(M_*, z)$$

Galaxy+BH Fit. Function

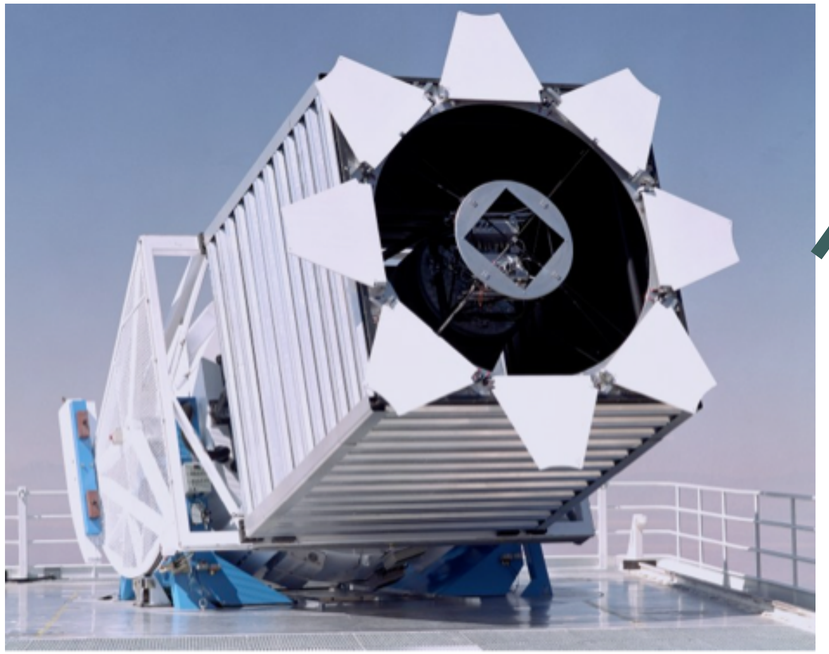
=

MCMC



Mock Universe

?
=



Real Universe

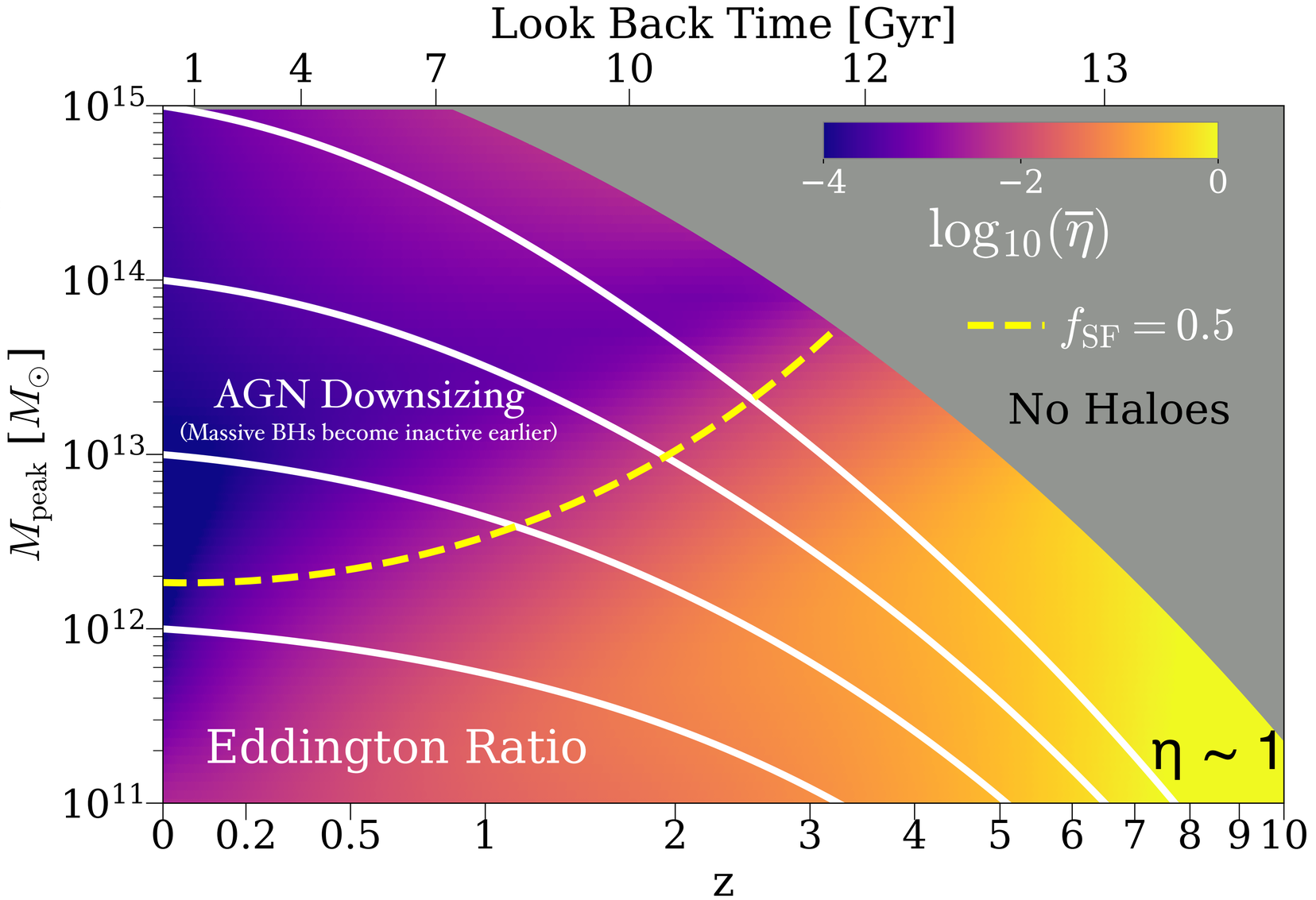
HZ, PB, et al.: https://arxiv.org/abs/2105.10474

Trinity: Constraints

Clusters

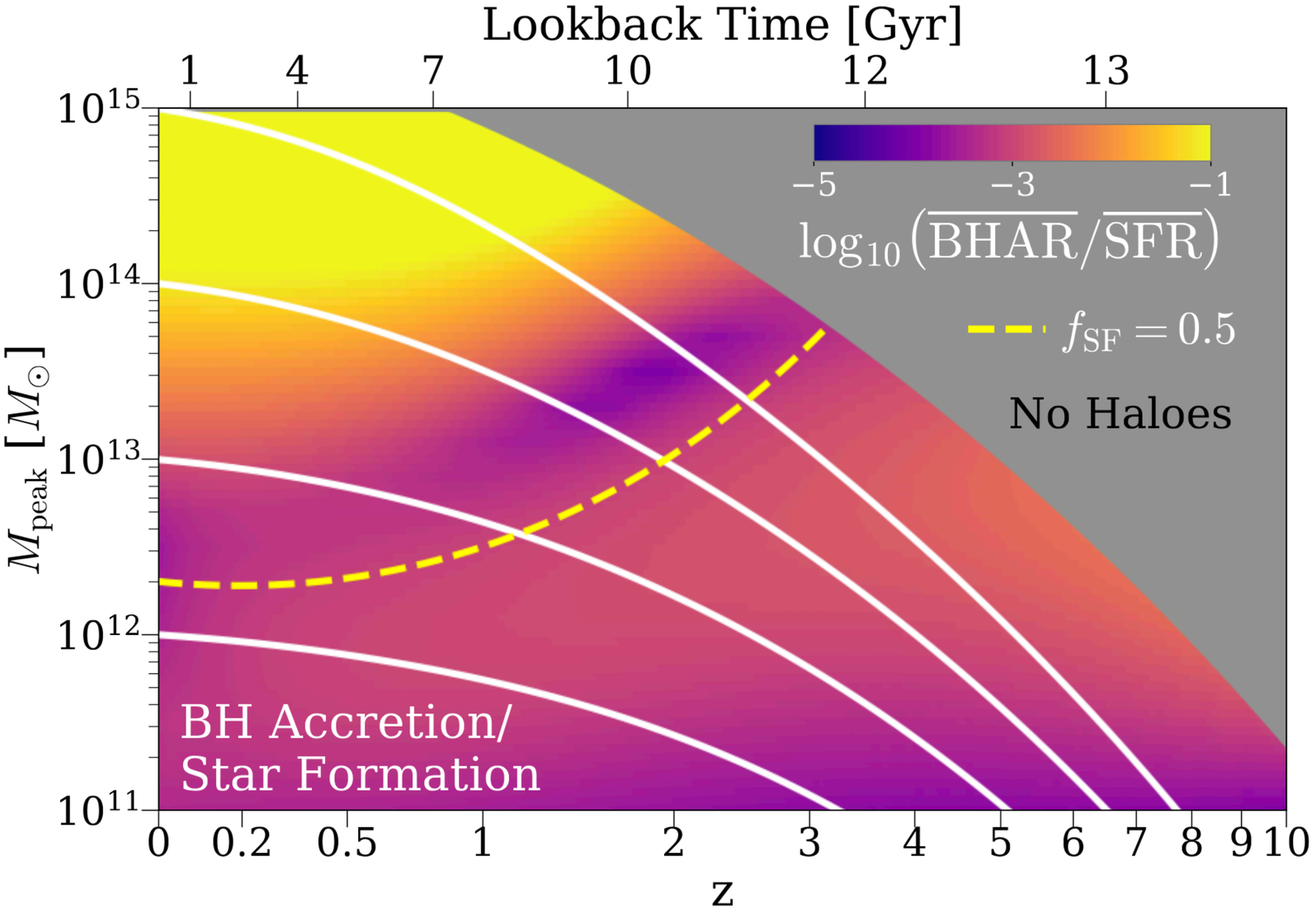
MW

Dwarfs



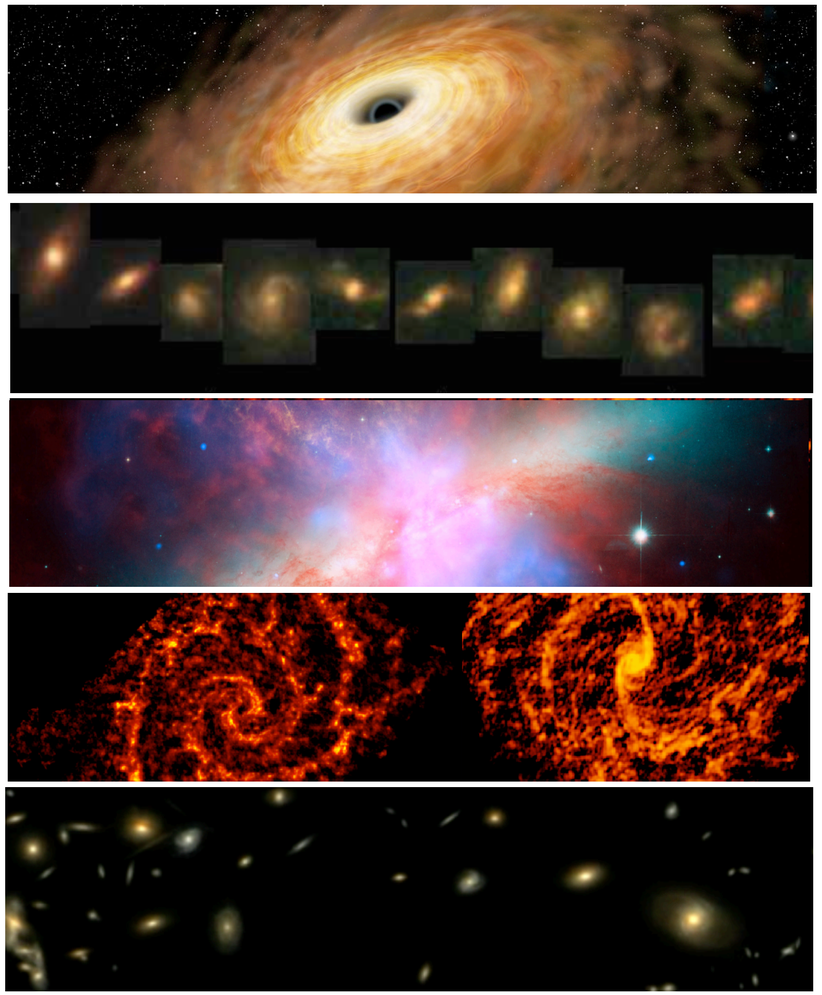
HZ, PB, et al.: https://arxiv.org/abs/2105.10474

Trinity: Results



HZ, PB, et al.: https://arxiv.org/pdf/2105.10474

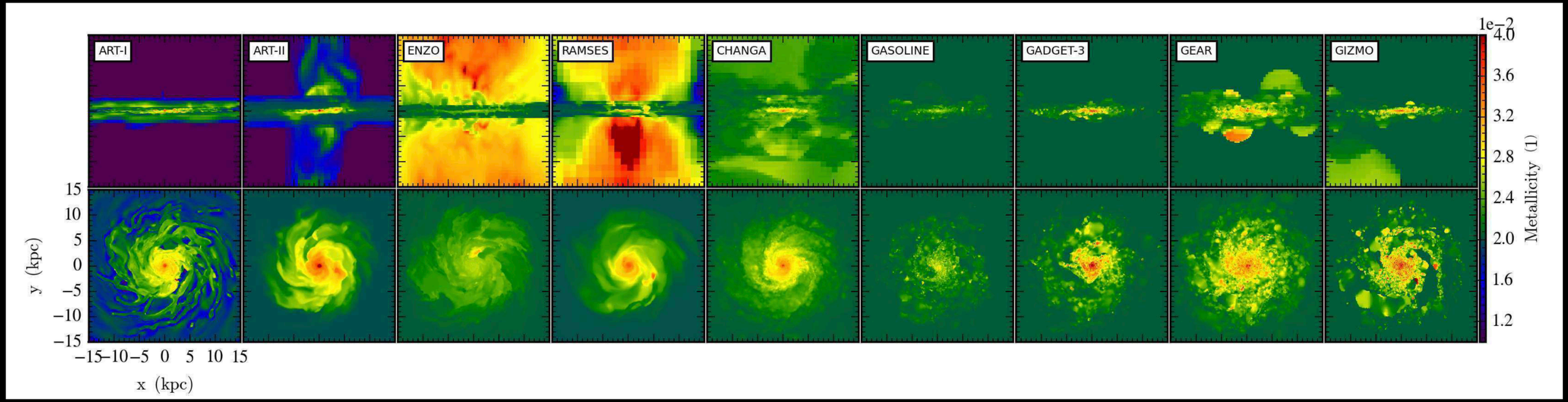
Empirical Galaxy Models



SMBHs	Trinity (Zhang et al. 2023ab, 2024ab)
Dust/Color	UniverseMachine 2 (Cooray et al. in prep.)
Metallicity	ChemicalUniverseMachine (Nishigaki et al., submitted)
HI, H_2 Masses	NeutralUniverseMachine (Guo et al. 2023)
Galaxy Masses	UniverseMachine (PB+ 2019); UM-SAGA (Wang et al. 2024)

Image credits: https://www.universetoday.com/108702/what-fuels-the-engine-of-a-supermassive-black-hole/;
Papovich et al. 2015; Armus et al. 2020; HST/Spitzer/Chandra; SDSS

No clear understanding yet:



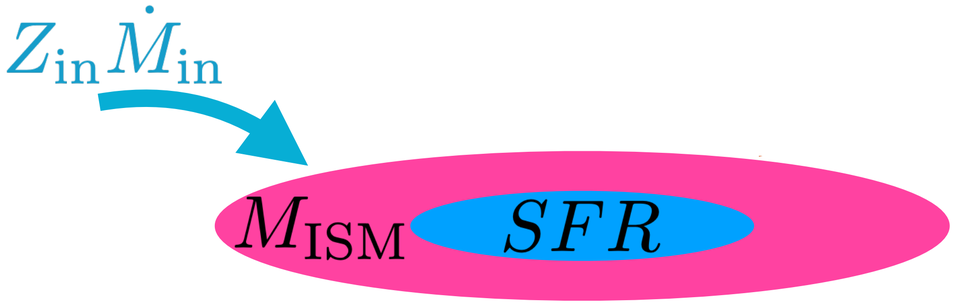
Credit: Kim et al., 2016

Chemical UniverseMachine

$$\frac{d}{dt}(\underbrace{Z_{\rm ISM}}_{\text{ISM Metallicity}} \cdot \underbrace{M_{\rm ISM}}_{\text{HI + HII Mass}}) = \overbrace{\quad\quad}^{\text{Total Metal Mass in ISM}}$$

Nishigaki et al., accepted

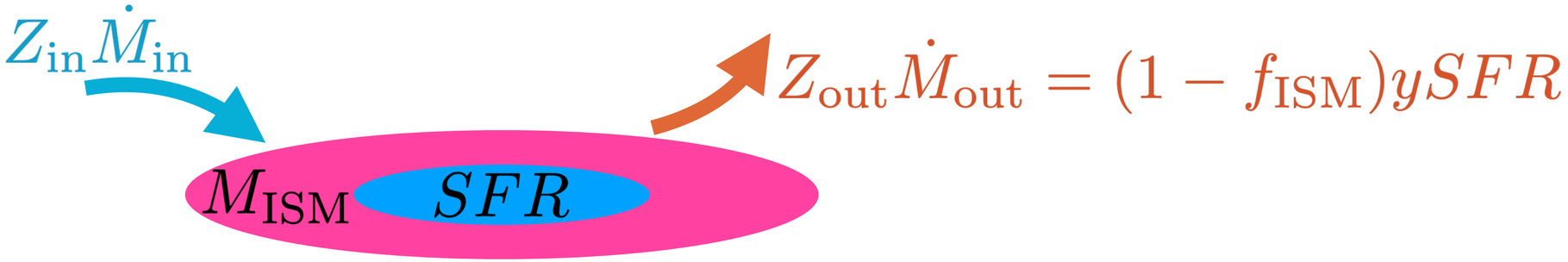
Chemical UniverseMachine



$$\frac{d}{dt}(\underbrace{Z_{\rm ISM}}_{\text{ISM Metallicity}} \cdot \underbrace{M_{\rm ISM}}_{\text{HI + HII Mass}}) = \underbrace{Z_{\rm in}}_{\substack{\text{Inflow} \\ \text{Metallicity}}} \underbrace{\dot{M}_{\rm in}}_{\text{Inflow Rate}} + \quad \text{Total Metal Mass in ISM}$$

Nishigaki et al., accepted

Chemical UniverseMachine



$$\frac{d}{dt}(Z_{\rm ISM} \cdot M_{\rm ISM}) = Z_{\rm in}\dot{M}_{\rm in} + (f_{\rm ISM} \cdot y)SFR$$

Total Metal Mass in ISM: $Z_{\rm ISM} \cdot M_{\rm ISM}$ — ISM Metallicity; HI + HII Mass

Inflow Metallicity: $Z_{\rm in}$

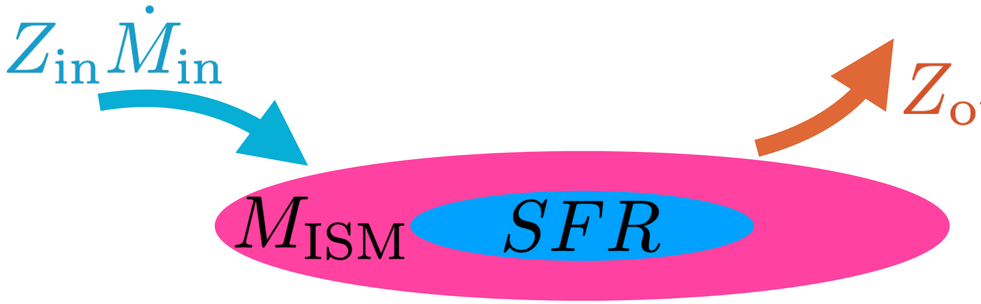
Inflow Rate: $\dot{M}_{\rm in}$

Effective Yield: $f_{\rm ISM} \cdot y$

Star Formation Rate: SFR

Nishigaki et al., accepted

Chemical UniverseMachine

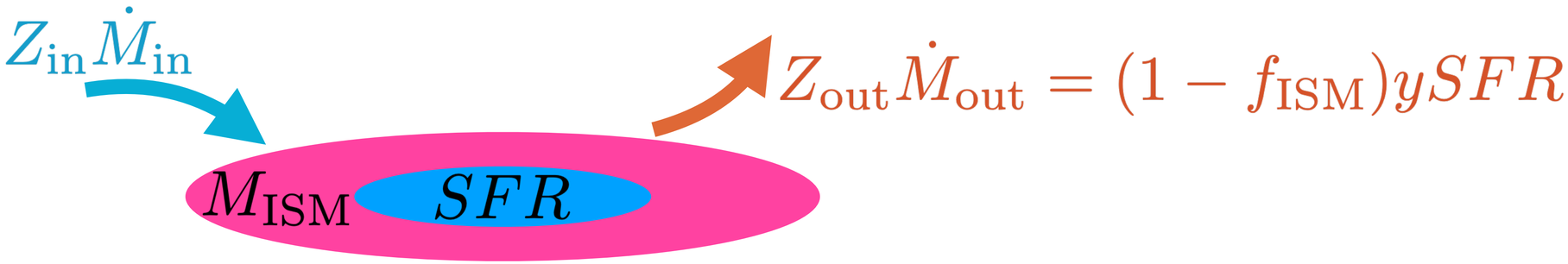


$$\frac{d}{dt}(Z_{\rm ISM} \cdot M_{\rm ISM}) = Z_{\rm in}\dot{M}_{\rm in} + (f_{\rm ISM} \cdot y - Z_{\rm ISM} \cdot f_{\rm rem})SFR$$

- $Z_{\rm ISM} \cdot M_{\rm ISM}$: Total Metal Mass in ISM (ISM Metallicity × HI + HII Mass)
- $Z_{\rm in}$: Inflow Metallicity
- $\dot{M}_{\rm in}$: Inflow Rate
- $f_{\rm ISM} \cdot y$: Effective Yield
- $Z_{\rm ISM}$: ISM Metallicity
- $f_{\rm rem}$: Fraction of long-lived stellar mass
- SFR: Star Formation Rate

Nishigaki et al., accepted

Intuition

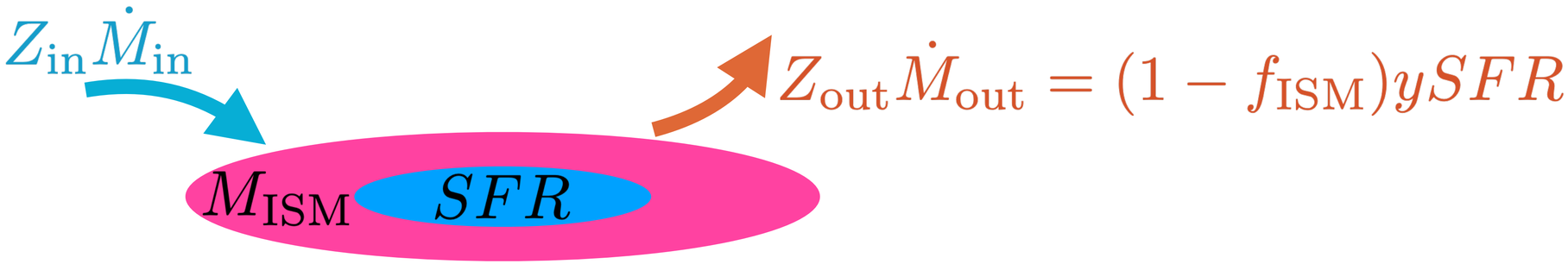


$$\frac{d}{dt}(Z_{\rm ISM} \cdot M_{\rm ISM}) = \sim 0 \; Z_{\rm in}\dot{M}_{\rm in} + (f_{\rm ISM} \cdot y - \sim 0 \; Z_{\rm ISM} \cdot f_{\rm rem})SFR$$

- $\frac{d}{dt}(Z_{\rm ISM} \cdot M_{\rm ISM})$: Total Metal Mass in ISM
- $Z_{\rm ISM}$: ISM Metallicity
- $M_{\rm ISM}$: HI + HII Mass
- $Z_{\rm in}$: Inflow Metallicity
- $\dot{M}_{\rm in}$: Inflow Rate
- $f_{\rm ISM} \cdot y$: Effective Yield
- $Z_{\rm ISM}$: ISM Metallicity
- $f_{\rm rem}$: Fraction of long-lived stellar mass
- SFR: Star Formation Rate

$$(Z_{\rm ISM} \cdot M_{\rm ISM}) \approx \langle f_{\rm ISM} \rangle y M_*$$

Intuition



$$\frac{d}{dt}(Z_{\rm ISM} \cdot M_{\rm ISM}) = \cancel{Z_{\rm in}\dot{M}_{\rm in}} + (f_{\rm ISM} \cdot y - Z_{\rm ISM} \cdot \cancel{f_{\rm rem}})SFR$$

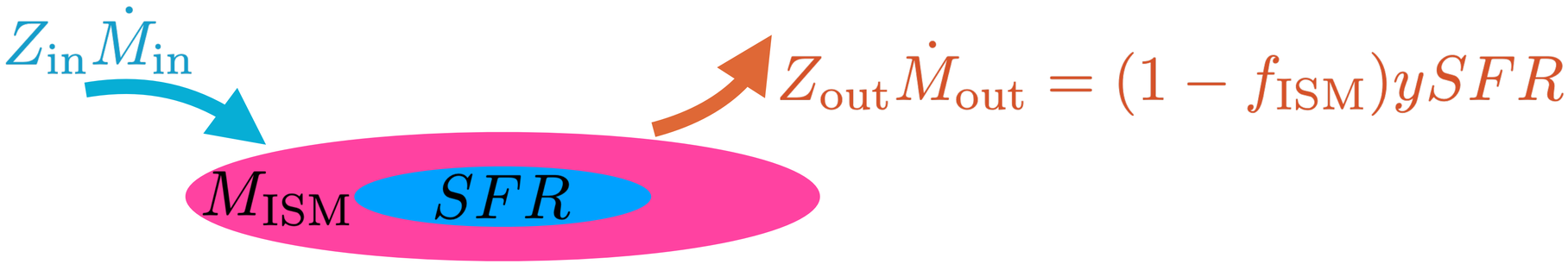
~ 0 ~ 0

- $\frac{d}{dt}(Z_{\rm ISM} \cdot M_{\rm ISM})$: Total Metal Mass in ISM
- $Z_{\rm ISM}$: ISM Metallicity
- $M_{\rm ISM}$: HI + HII Mass
- $Z_{\rm in}$: Inflow Metallicity
- $\dot{M}_{\rm in}$: Inflow Rate
- $f_{\rm ISM} \cdot y$: Effective Yield
- $Z_{\rm ISM}$: ISM Metallicity
- $f_{\rm rem}$: Fraction of long-lived stellar mass
- SFR: Star Formation Rate

$$(Z_{\rm ISM} \cdot M_{\rm ISM}) \approx \langle f_{\rm ISM} \rangle y M_*$$

$$\langle f_{\rm ISM} \rangle \approx \frac{Z_{\rm ISM} \cdot M_{\rm ISM}}{yM_*}$$

Intuition



$$\frac{d}{dt}(Z_{\rm ISM} \cdot M_{\rm ISM}) = \cancel{Z_{\rm in}\dot{M}_{\rm in}} + (f_{\rm ISM} \cdot y - \cancel{Z_{\rm ISM} \cdot f_{\rm rem}})SFR$$

$\frac{d}{dt}$: Total Metal Mass in ISM

$Z_{\rm ISM}$: ISM Metallicity

$M_{\rm ISM}$: HI + HII Mass

$Z_{\rm in}\dot{M}_{\rm in} \sim 0$: Inflow Metallicity, Inflow Rate

$f_{\rm ISM} \cdot y$: Effective Yield

$Z_{\rm ISM} \cdot f_{\rm rem} \sim 0$: ISM Metallicity, Fraction of long-lived stellar mass

SFR: Star Formation Rate

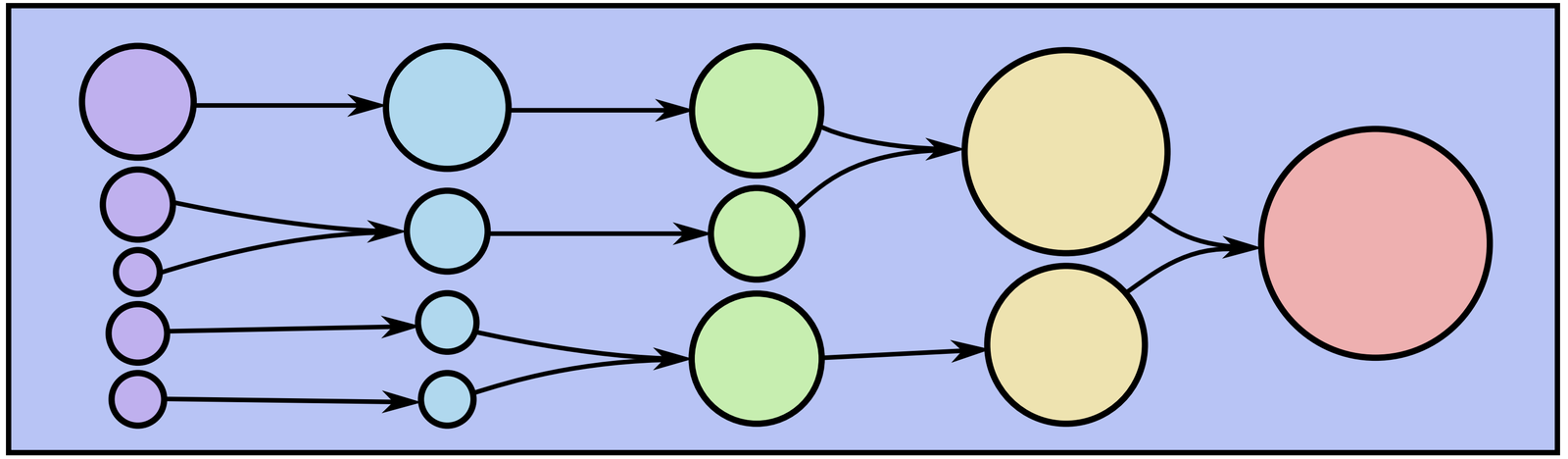
$$(Z_{\rm ISM} \cdot M_{\rm ISM}) \approx \langle f_{\rm ISM} \rangle y M_*$$

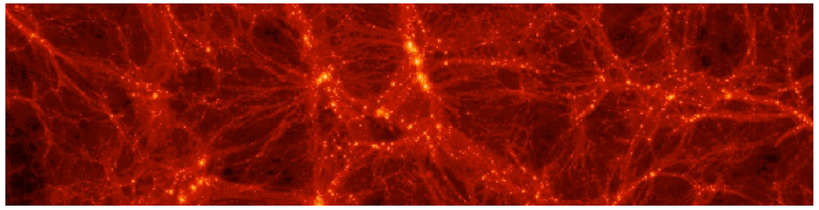
$$\langle f_{\rm ISM} \rangle \approx \frac{Z_{\rm ISM} \cdot M_{\rm ISM}}{y M_*}$$

Everything is observable/constrainable(!)

$$\langle f_{\rm ISM} \rangle \approx \frac{Z_{\rm ISM} \cdot M_{\rm ISM}}{y M_*}$$

Modeling Stages

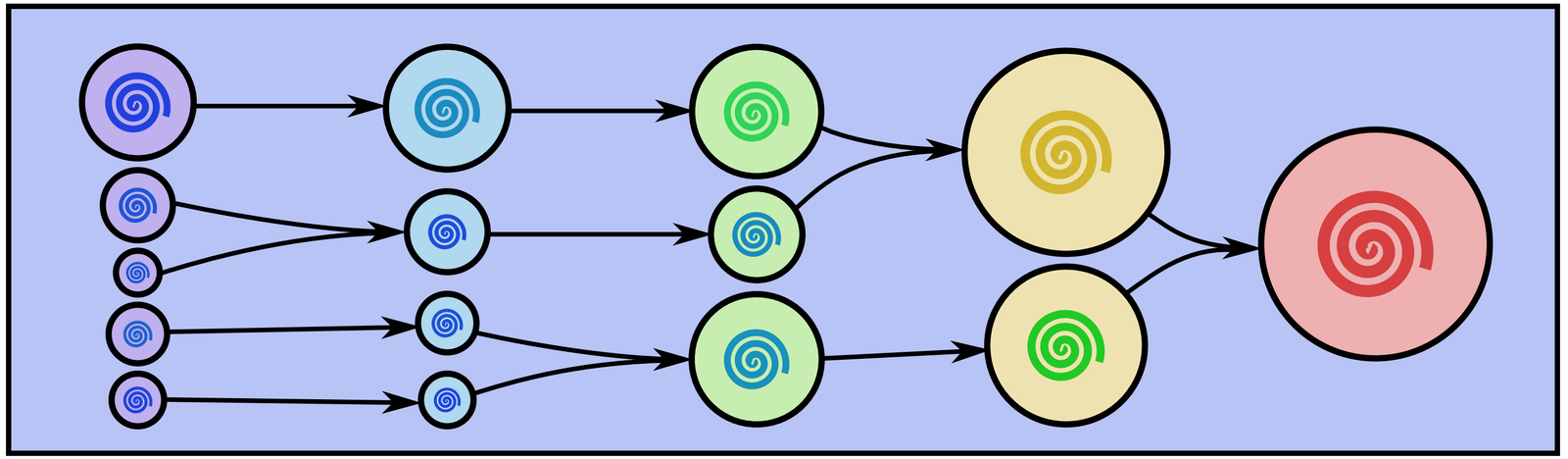


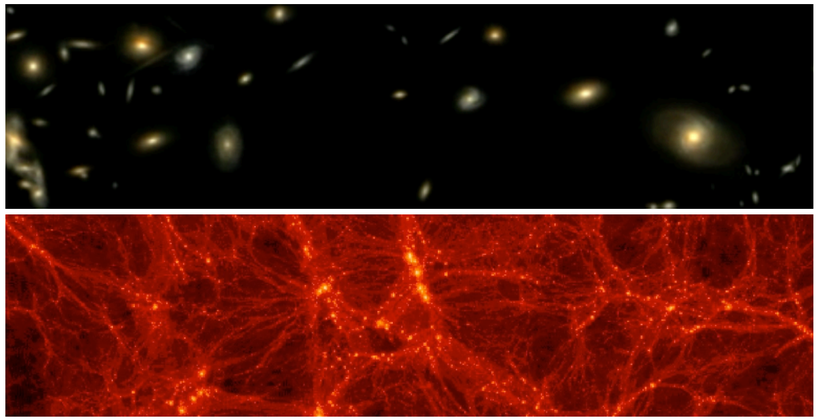


DM Simulation Klypin+ (2016)

$$\langle f_{\rm ISM} \rangle \approx \frac{Z_{\rm ISM} \cdot M_{\rm ISM}}{y M_*}$$

Modeling Stages



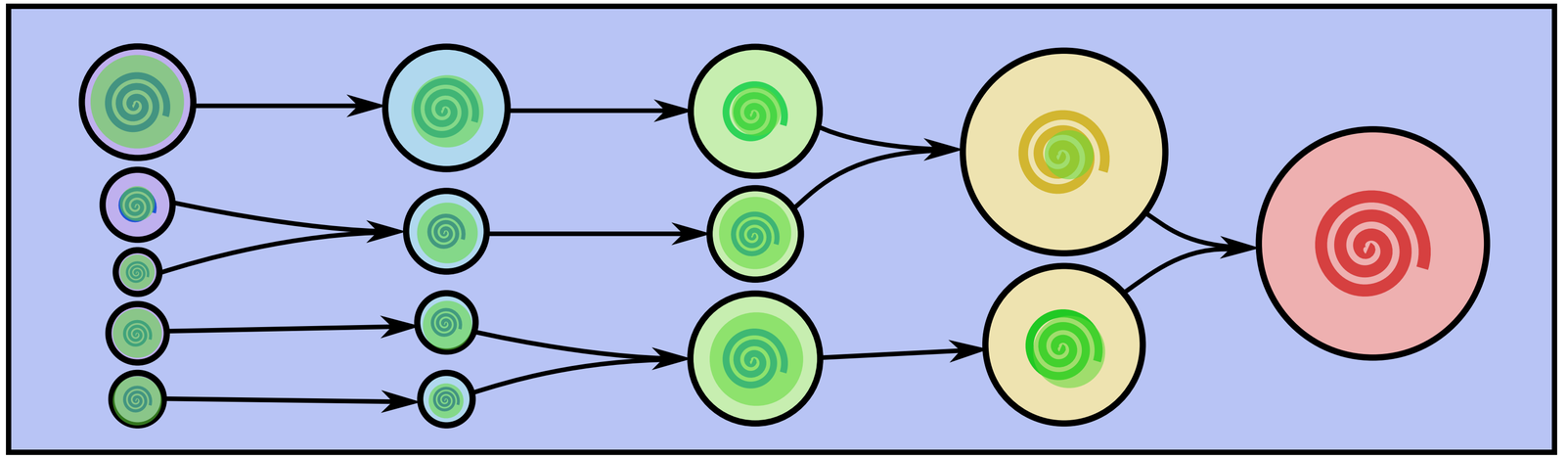


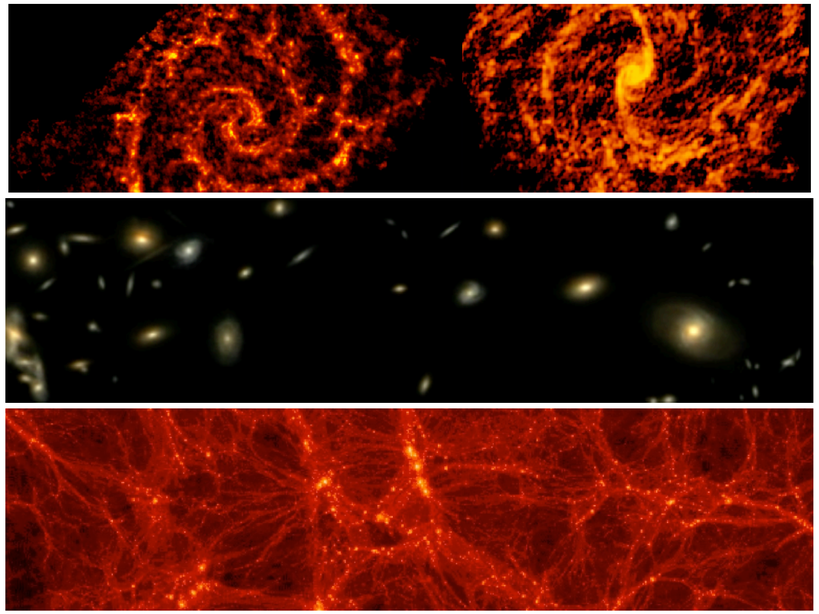
Galaxy Masses — UniverseMachine (PB+ 2019)

DM Simulation — Klypin+ (2016)

$$\langle f_{\rm ISM} \rangle \approx \frac{Z_{\rm ISM} \cdot M_{\rm ISM}}{y M_*}$$

Modeling Stages



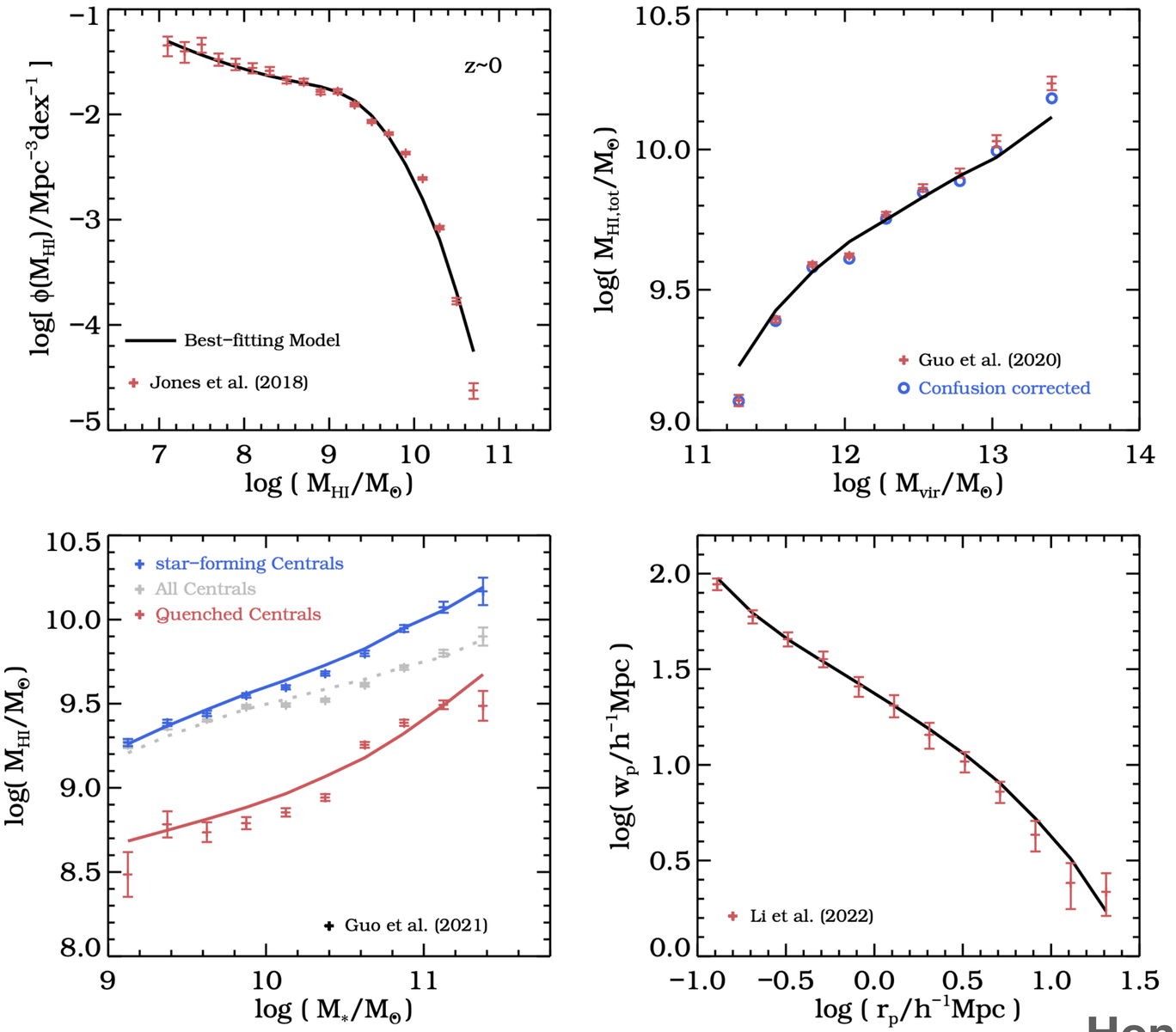


HI, H_2 Masses — NeutralUniverseMachine (Guo et al. 2023)

Galaxy Masses — UniverseMachine (PB+ 2019)

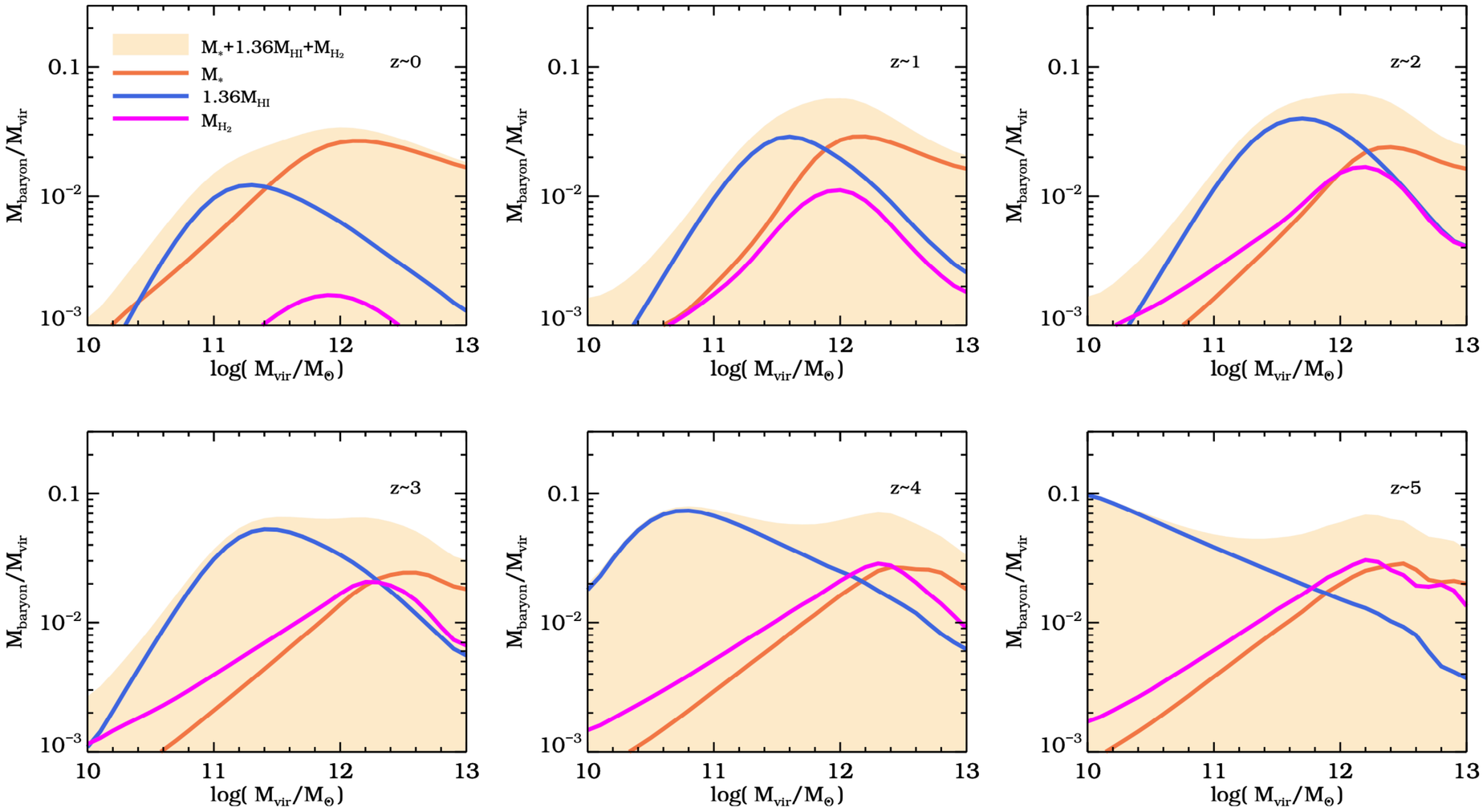
DM Simulation — Klypin+ (2016)

Neutral UniverseMachine



Hong Guo, et al (2023)

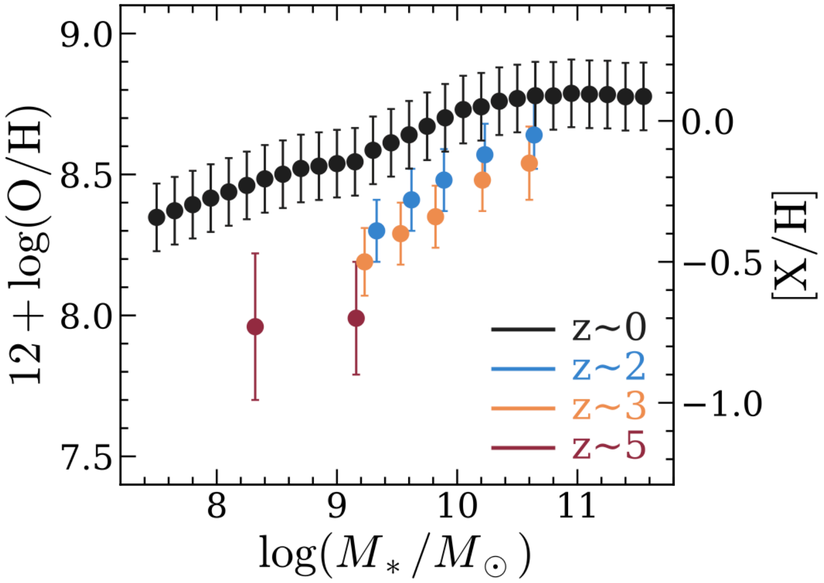
Neutral UniverseMachine



Hong Guo, et al (2023)

$$\langle f_{\rm ISM} \rangle \approx \frac{Z_{\rm ISM} \cdot M_{\rm ISM}}{y M_*}$$

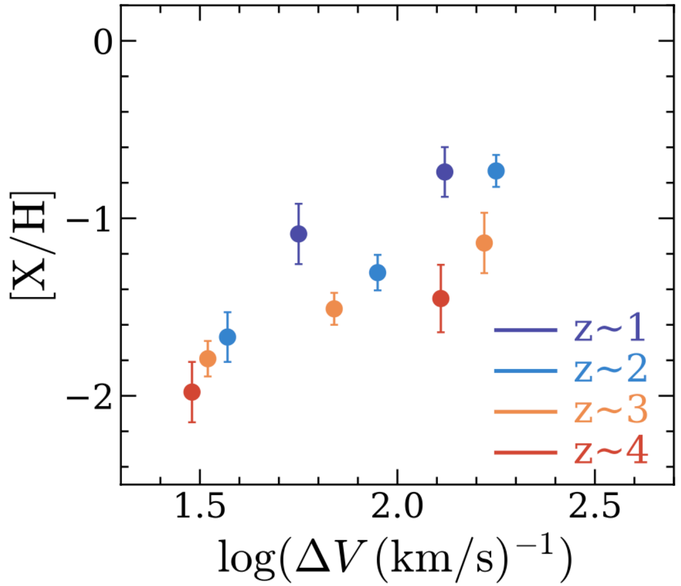
ISM metallicity (HII)



Emission lines

Curti+20; Sanders+21; Nakajima+23

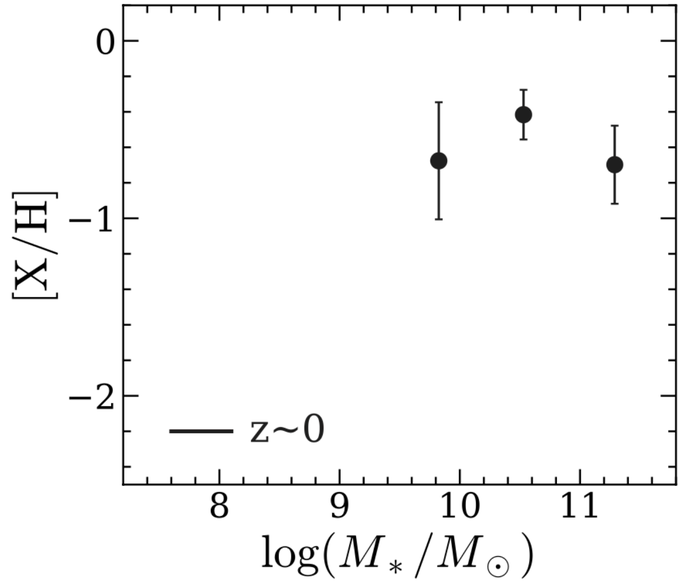
ISM metallicity (HI)



DLAs

Moller+13

CGM metallicity



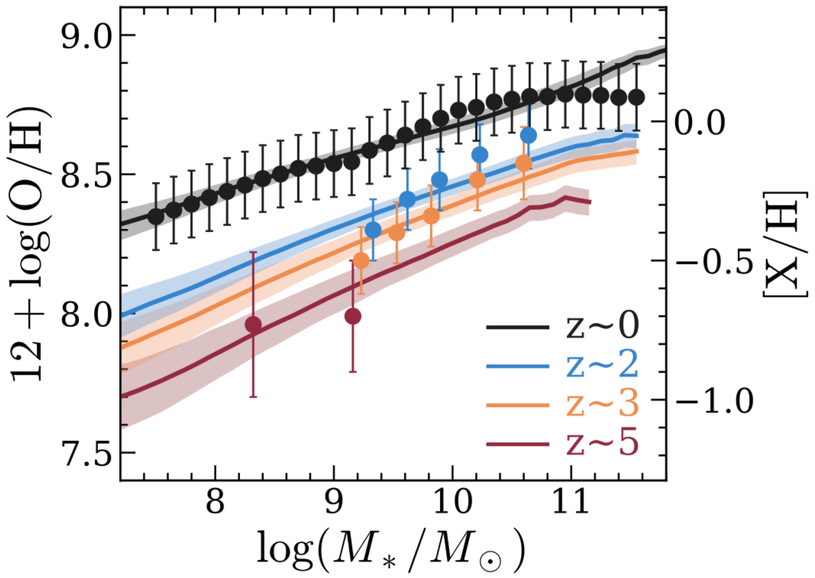
CGM absorption

Prochaska+17

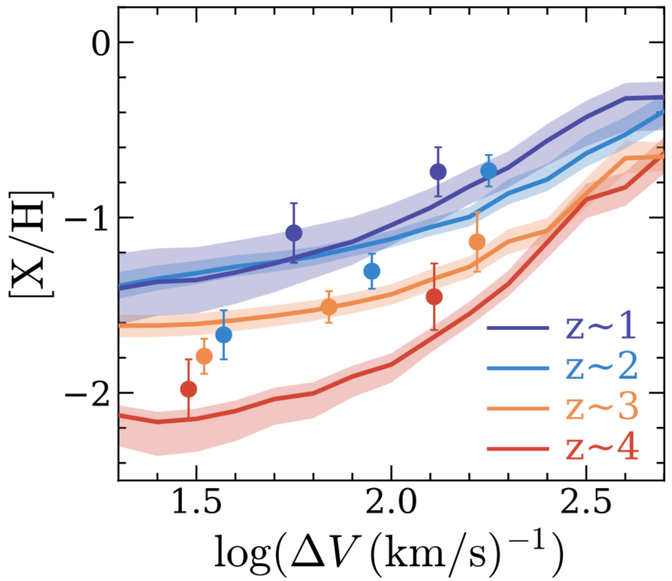
Decoupled ISM Metallicity

Metallicity of HI different from Metallicity of H_2

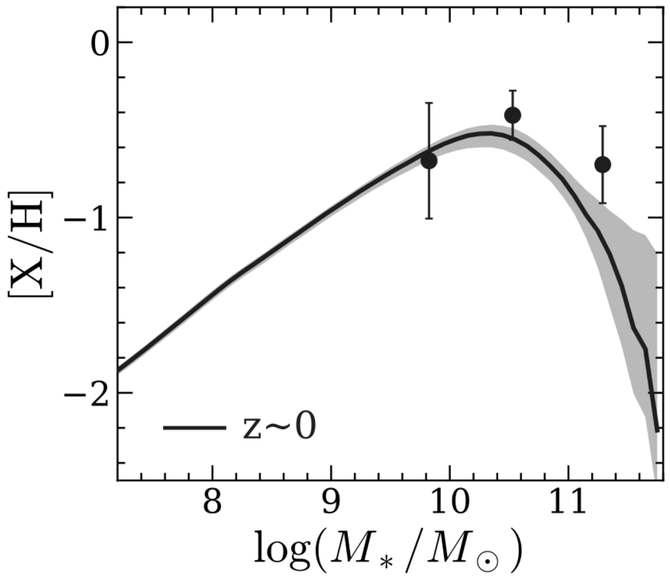
ISM metallicity (H2)



ISM metallicity (HI)



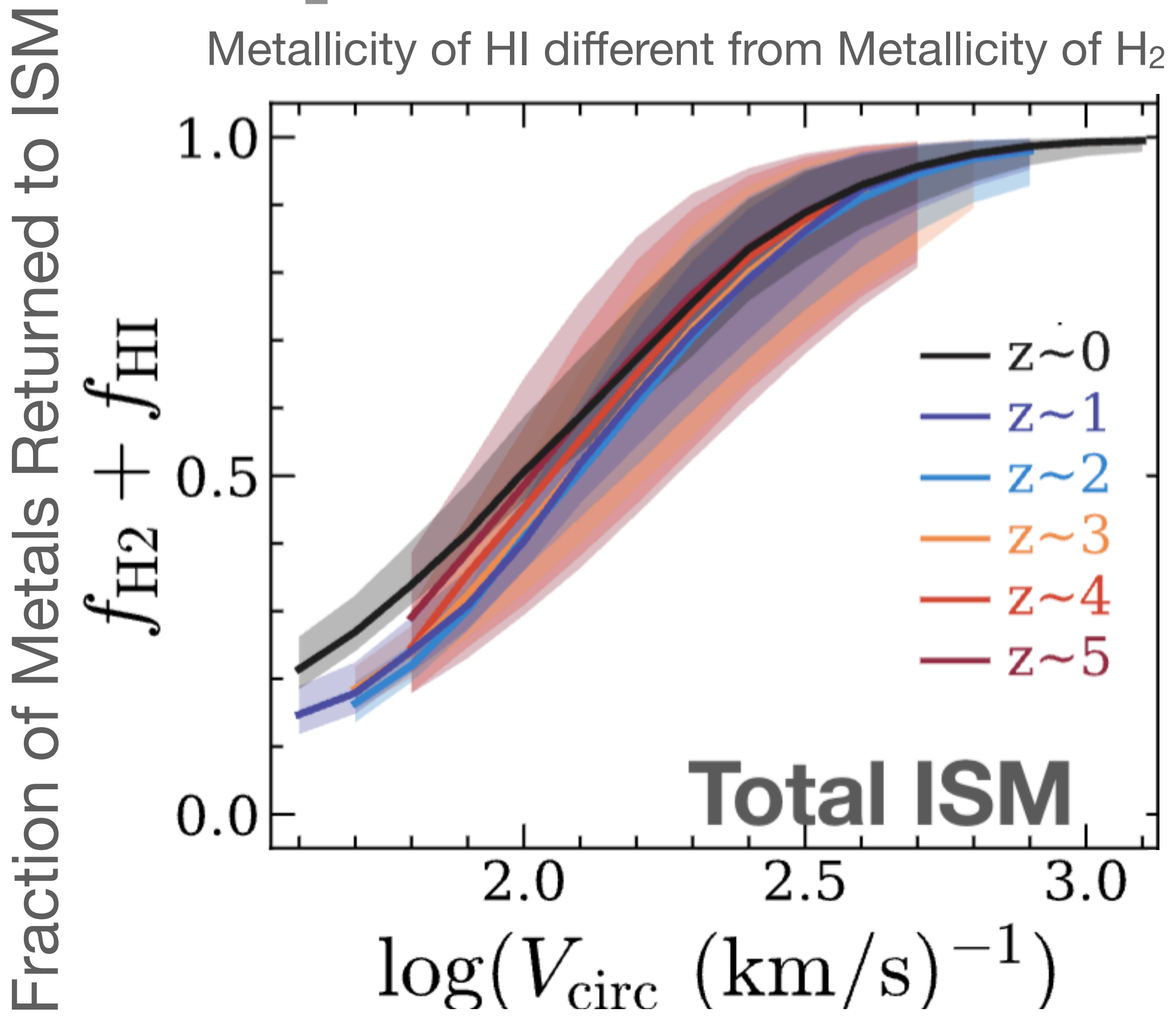
CGM metallicity



Nishigaki et al., 2025

Decoupled ISM Metallicity

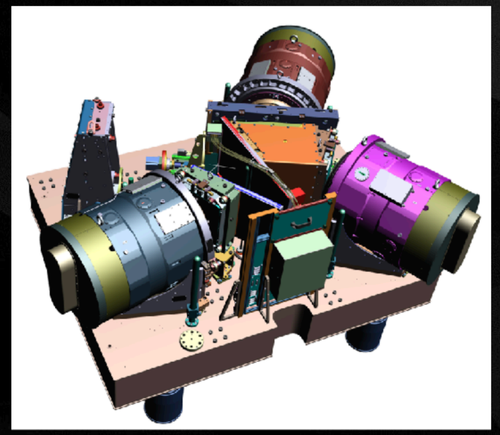
Metallicity of HI different from Metallicity of H_2



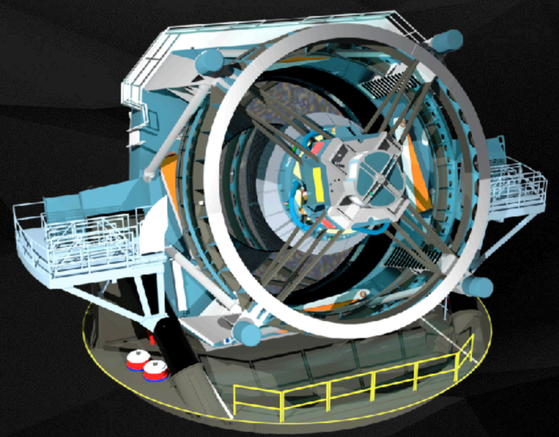
Potential Well Depth

Nishigaki et al., 2025

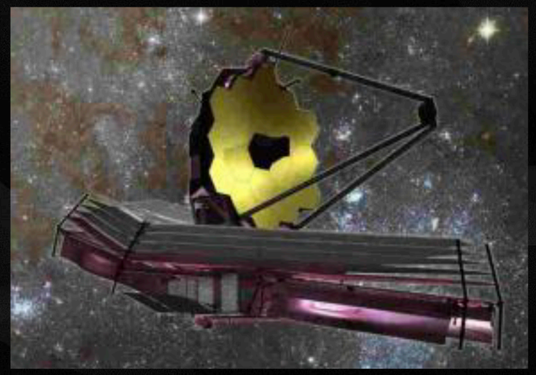
What the Future Brings



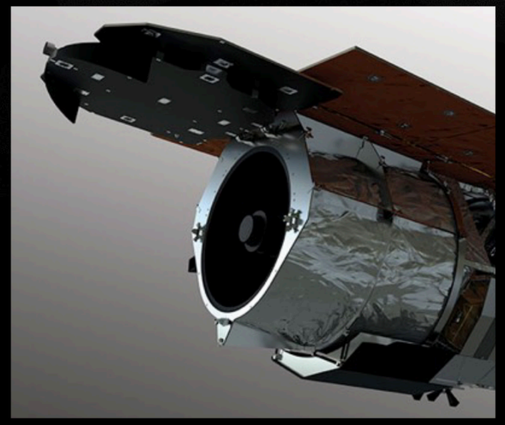
Subaru PFS



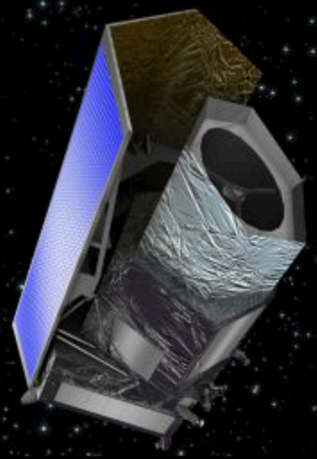
LSST



JWST



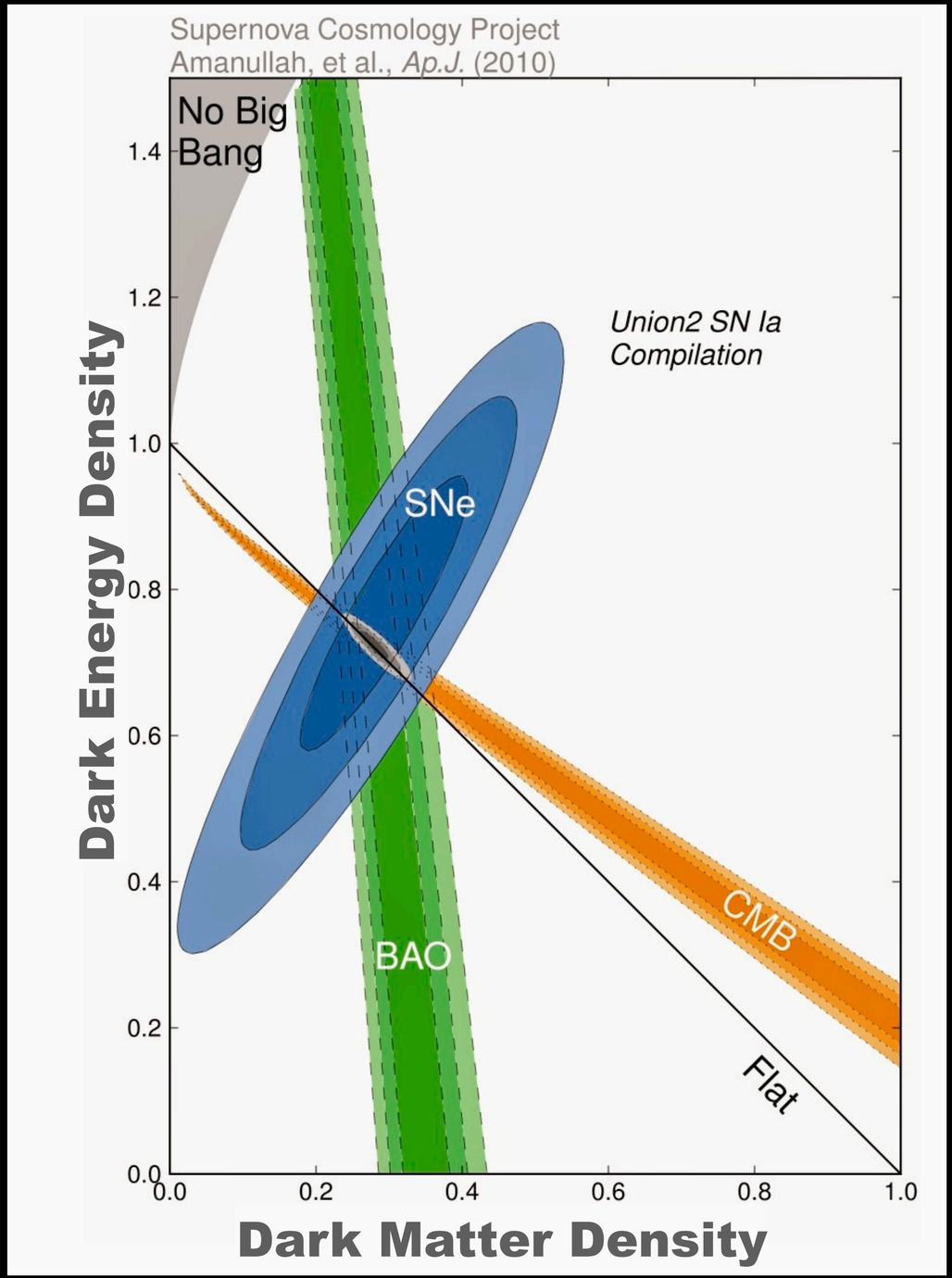
WFIRST



Euclid

XRISM

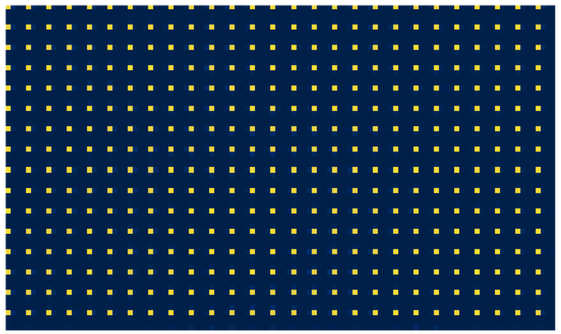
Credits: Subaru PFS Team, LSST Project Office, NASA, ESA, JAXA


Supernova Cosmology Project
Amanullah, et al., Ap.J. (2010)
No Big Bang
Union2 SN Ia Compilation
SNe
BAO
CMB
Flat
Dark Energy Density
Dark Matter Density
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1.4

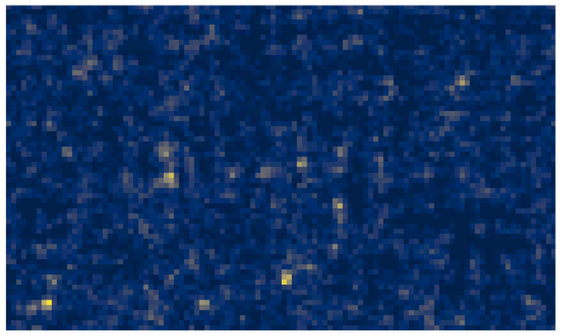
Summary:

Methods that scale better to large datasets will reveal more science(!)

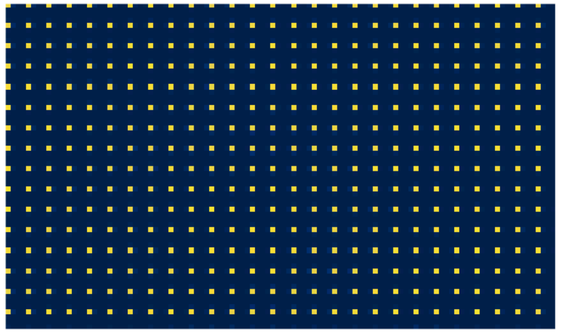
Enormous number of discoveries will be made with current and future datasets!



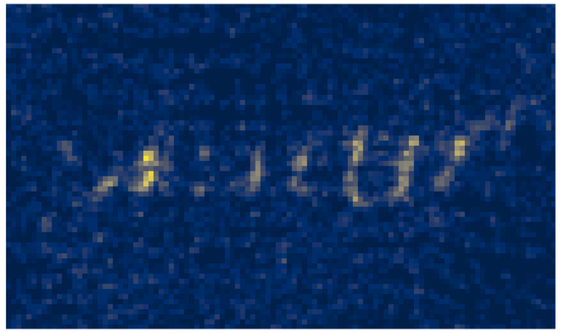
(a) initial conditions at $a = 1/64$



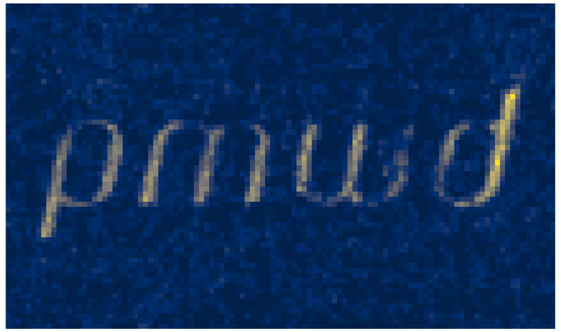
(b) final snapshot at $a = 1$



(c) reverse evolution back to $a = 1/64$



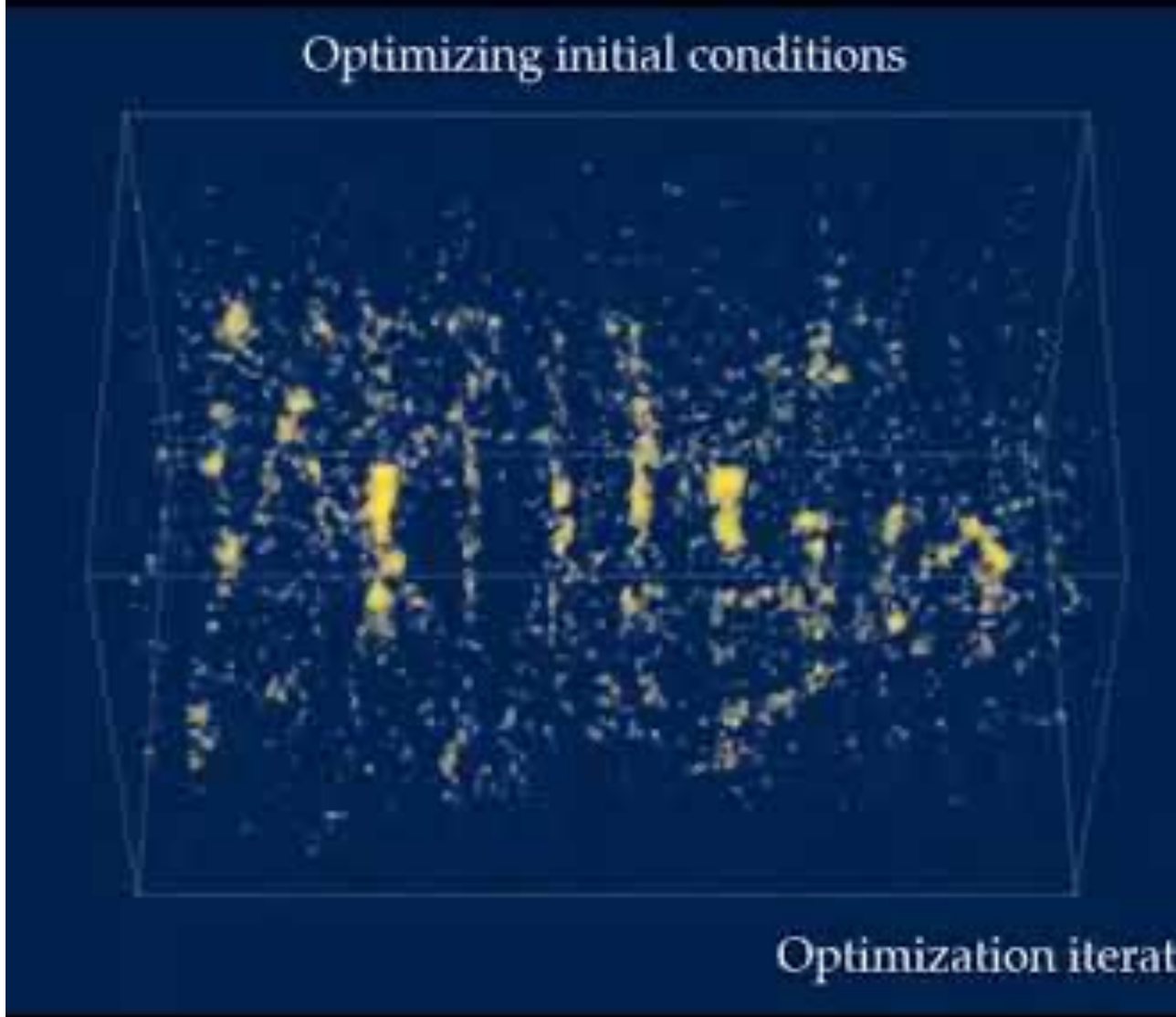
(d) optimization for 10 iterations at $a = 1$



(e) optimization for 100 iterations at $a = 1$

(f) optimization for 1000 iterations at $a = 1$

Figure 4. A toy problem where we optimize the initial conditions by gradient descent to make some interesting pattern after projection. The particles originally fill a $16 \times 27 \times 16$ grid, and then evolve from $a = 1/64$ to $a = 1$ for 63 time steps with *single precision* and a



Normalizing Flow

Normalizing Flow

Diffusion Model

Cosmo-logical Simulations → Physical Galaxy Properties → Integrated Photometry, Spectra → Projected Images, IFU Cubes