MAHGIC: A Model Adapter for the Halo-Galaxy Inter-Connection

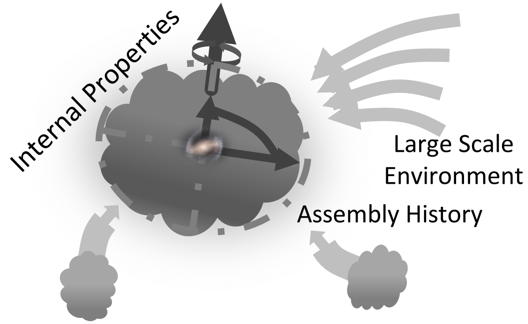
Yangyao Chen

yangyaochen.astro@foxmail.com

In collaboration with: Houjun Mo, Cheng Li, Kai Wang, Huiyuan Wang, Xiaohu Yang

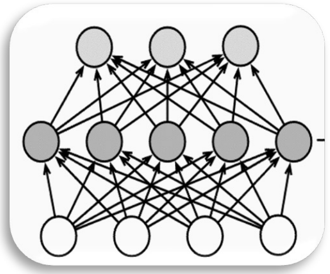
Toward Precise Halo-based Galaxy Model

1



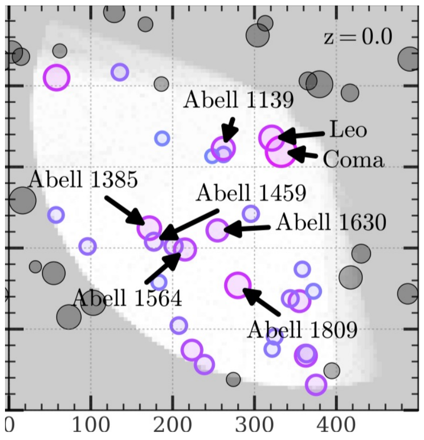
Dark matter halos & merger trees

2



Halo to galaxy mappings with tunable parameters

3



Galaxies and mock catalogs

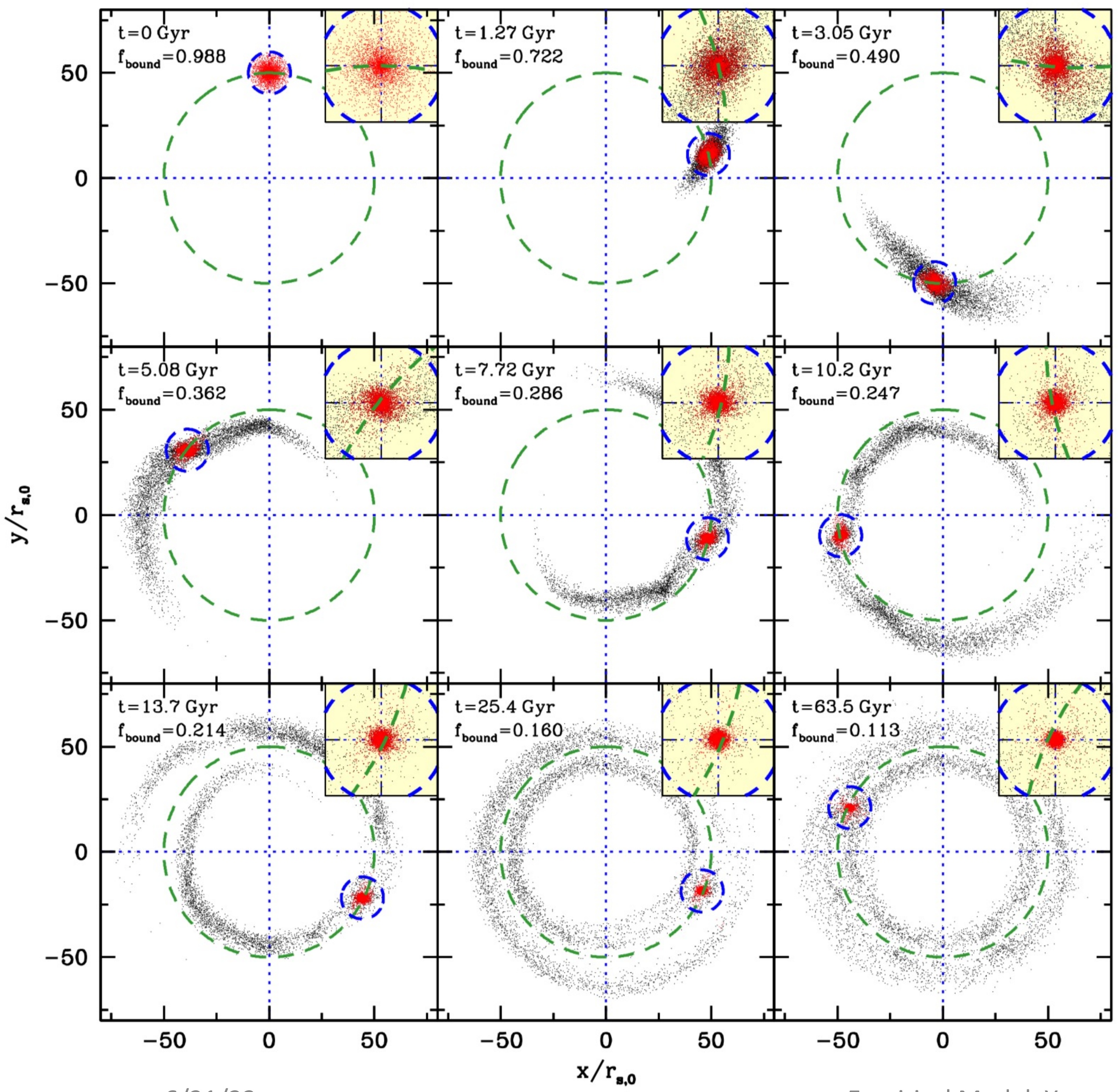
CAMTree

A Conditional Abundance Matching Method of Extending Subhalo Merger Trees

MAHGIC

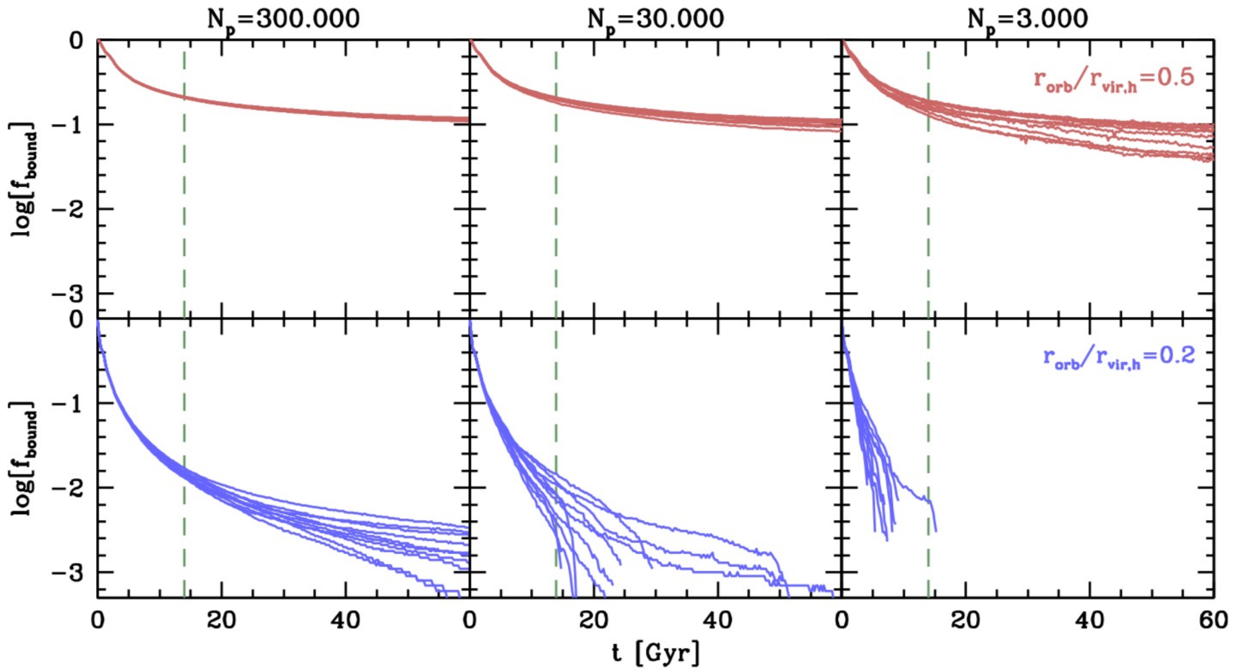
A Model Adapter for the Halo-Galaxy Inter-Connection

Massive Dark Matter Halos at High Redshift and Their Implications for Observations in the JWST Era



Incompleteness of large-volume N-body simulations

- Most of the satellite disruptions are artificial.
- Numerical defeats depend on subhalo mass ratio, time duration since infall, mass resolution, gravitational softening, etc.

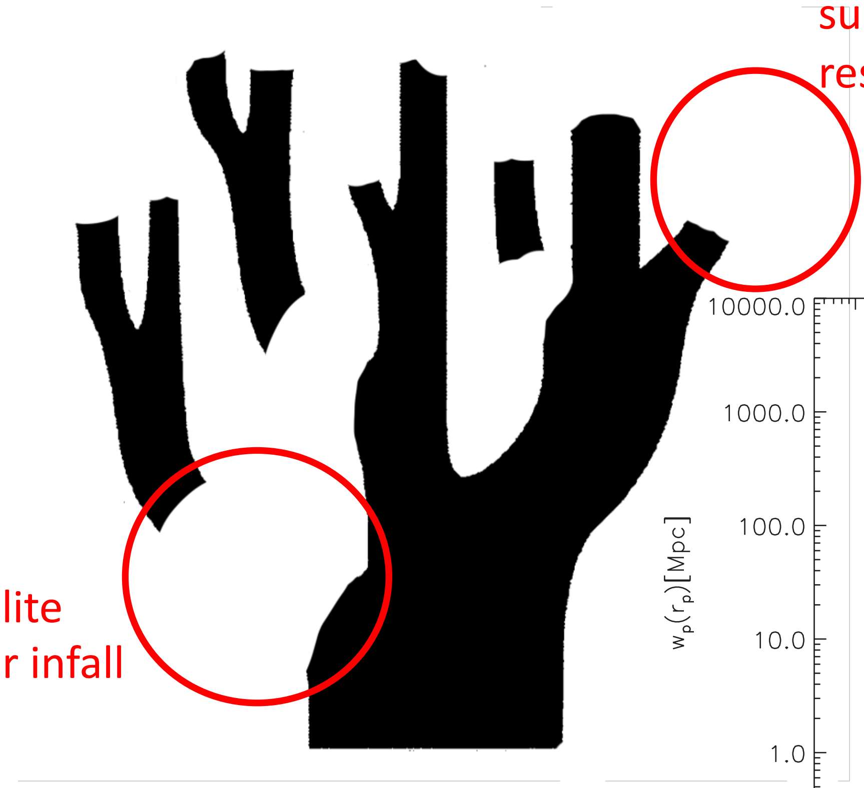


van den Bosch+ 2018

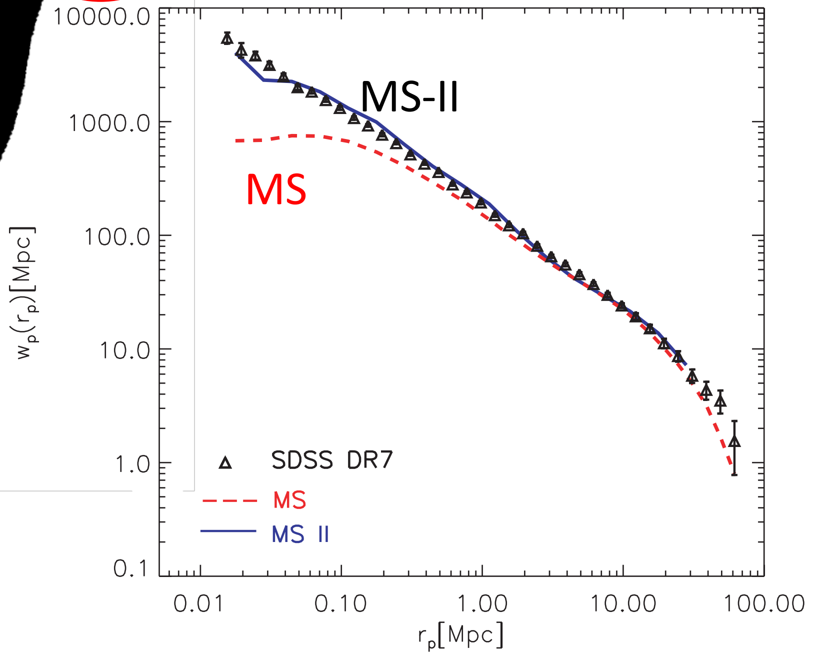
An Ideal Subhalo Merger Tree

A Trade-off between Resolution and Sample Size

Missed isolated (central) subhalo due to limited mass resolution



Missed satellite subhalo after infall



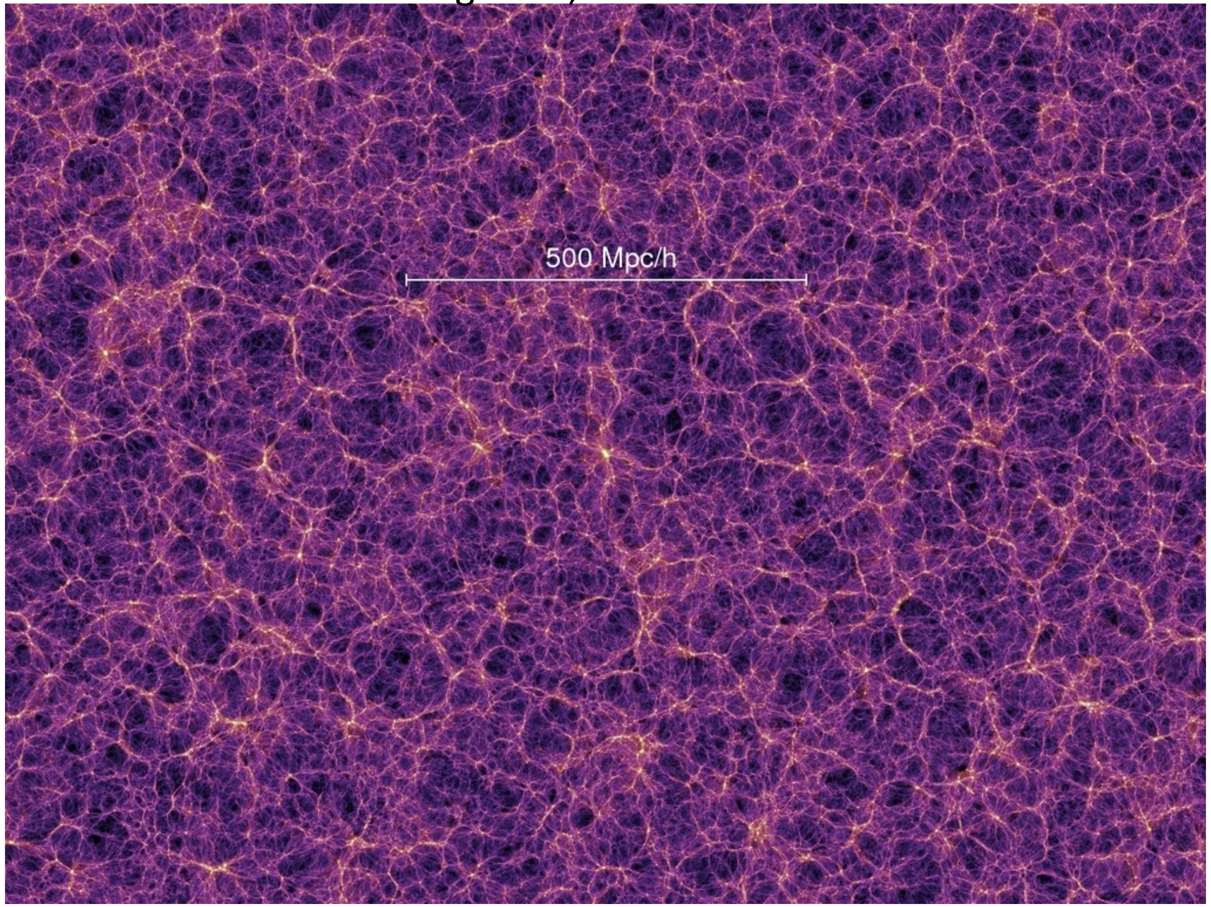
Guo+ 2010, Subhalo Abundance Matching

CAMTree: A Conditional Abundance Matching Method of Extending Subhalo Merger Trees

$$\mathrm{p}\left(X_{missed}, X_{conditioning}\right) = \underbrace{p\left(X_{conditioning}\right)}_{\text{Read from target simulation}} \underbrace{p\left(X_{missed} | X_{conditioning}\right)}_{\text{Learned from reference simulation}}$$

Target Simulation

Large box, low resolution



Missed variables:

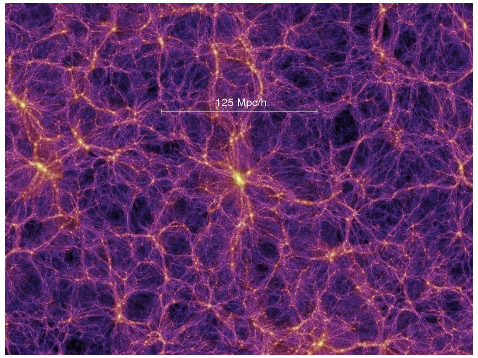
$t_{disruption}$, $\mathbf{x}$, $\mathbf{v}$, spin, $M_{subhalo}$, v_{max}, σ_v, $r_{half\ mass}$, …

Conditioning variables:

$M_{h,host}$, host halo shape, z_{form}, L_{orbit} , …

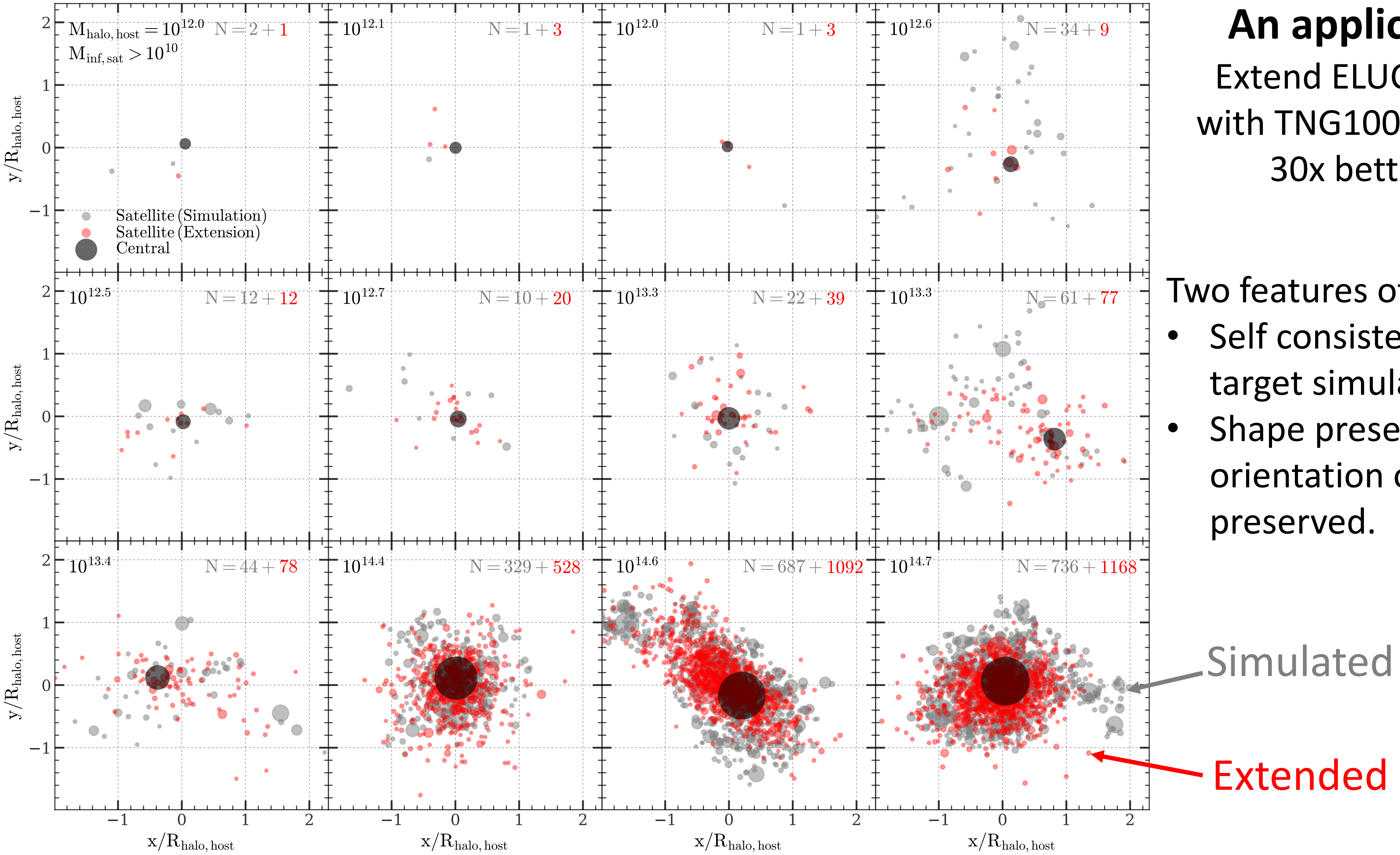
Reference Simulation

Small box, high resolution



An application of CAMTree

Extend ELUCID ($L = 500\ h^{-1}Mpc$)
with TNG100-Dark ($L = 75\ h^{-1}Mpc$,
30x better mass resolution)

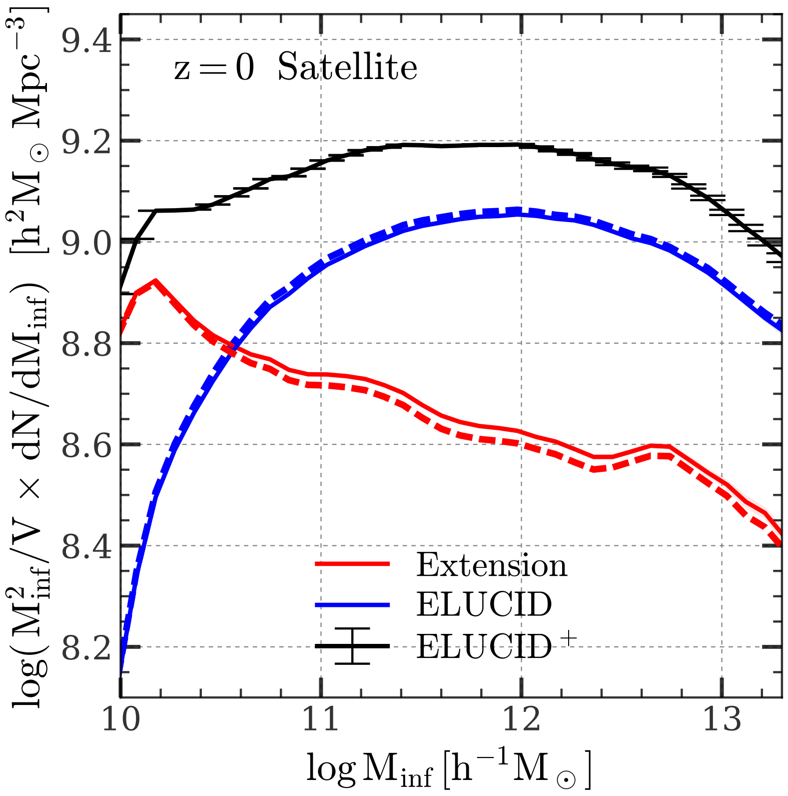


Two features of the extension algorithm:

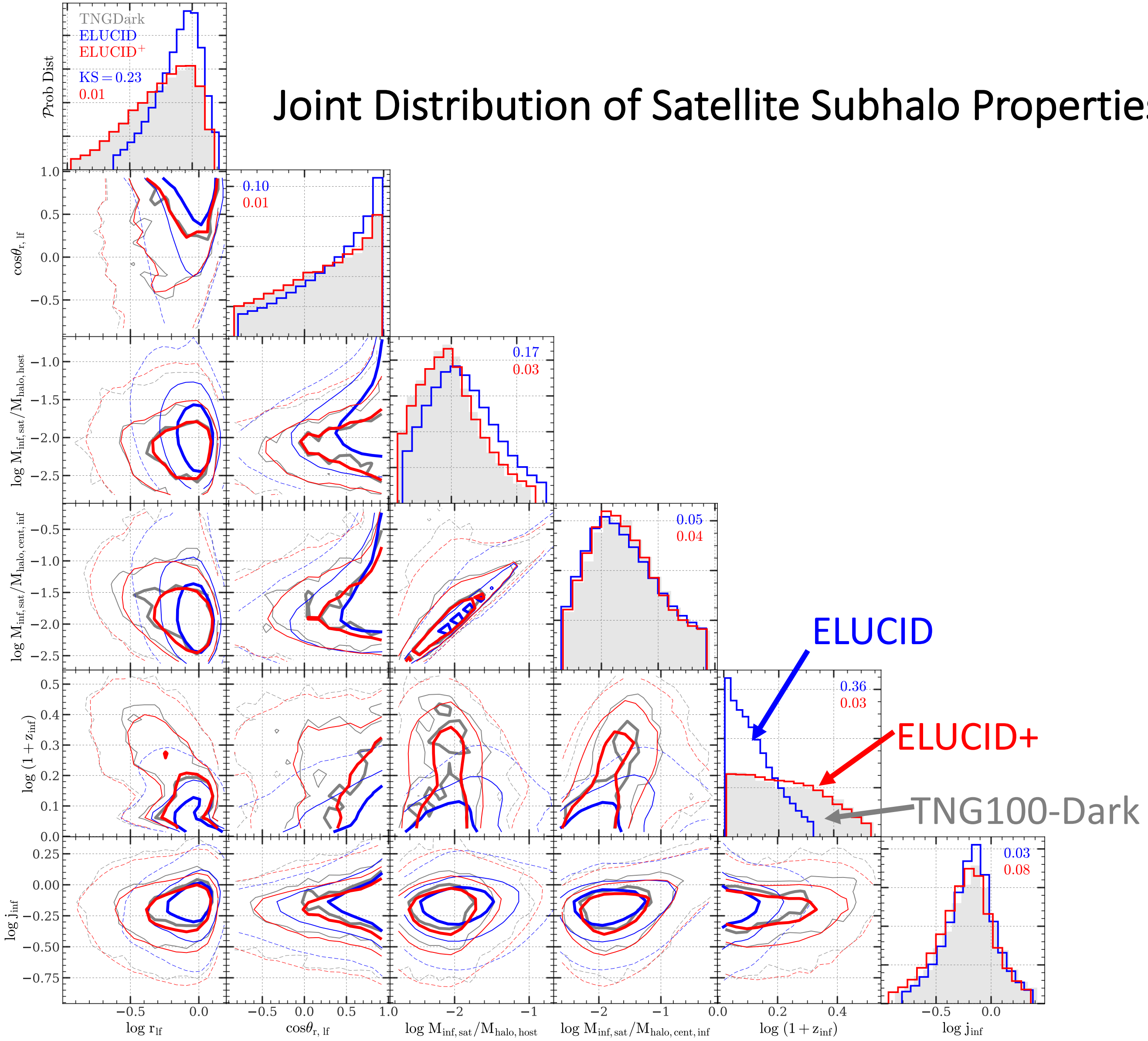
- Self consistency: subhalos resolved by target simulation are kept.
- Shape preservation: shape and orientation of the host halo is preserved.

CAMTree: Recovery of Key Subhalo Statistics

Satellite Subhalo Mass Function

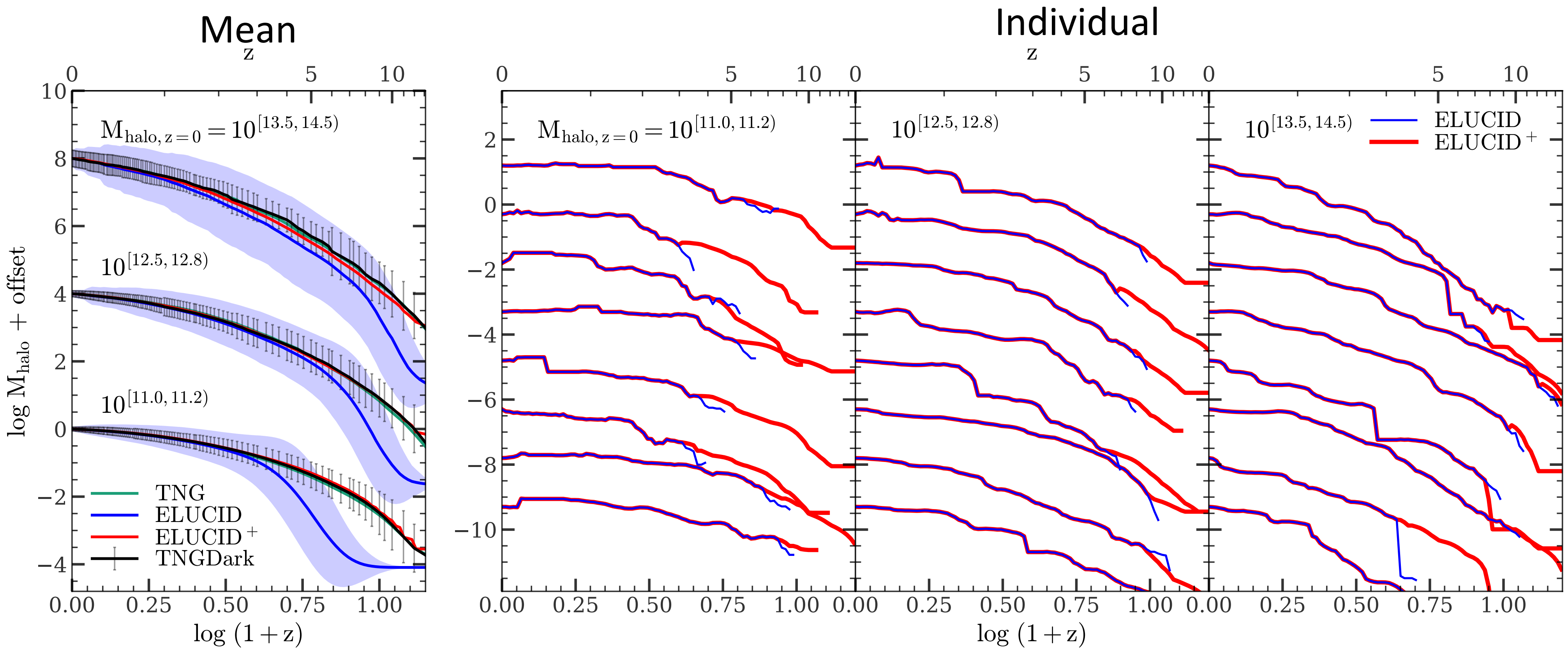


Joint Distribution of Satellite Subhalo Properties



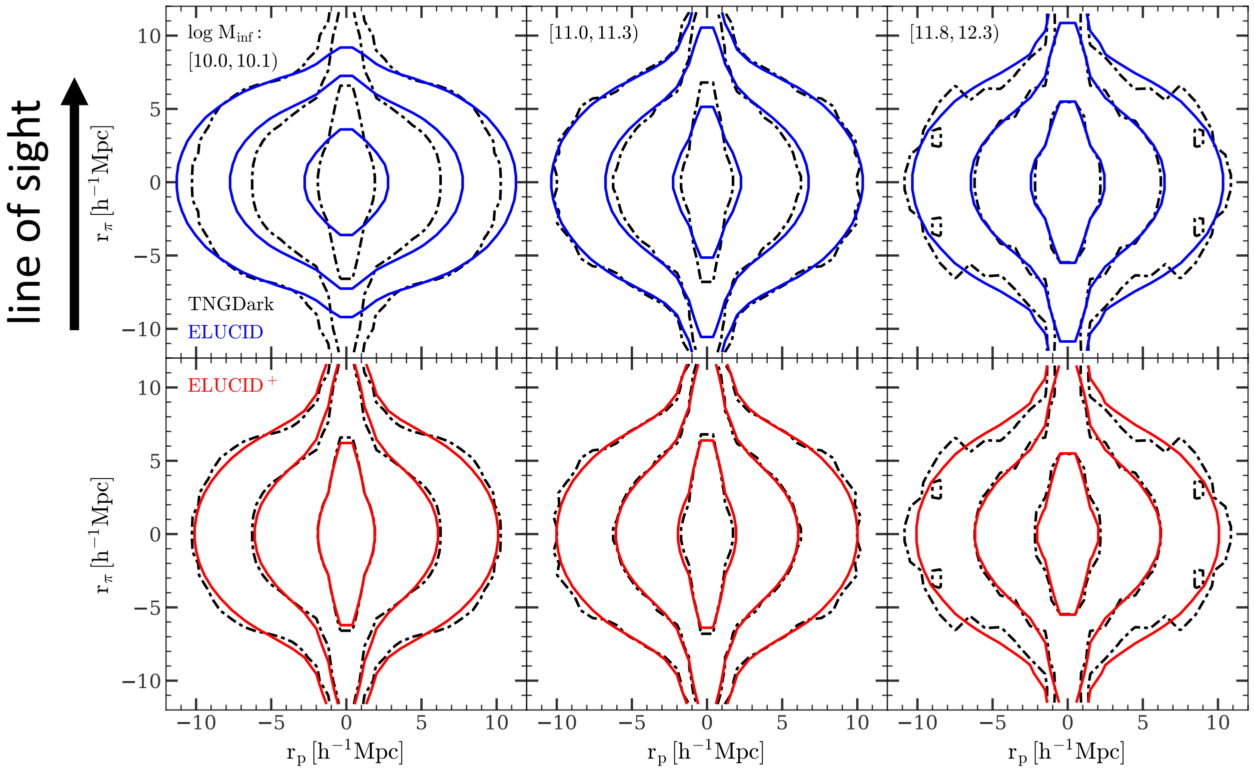
CAMTree: Recovery of Key Subhalo Statistics

Extended Central Subhalo Assembly History

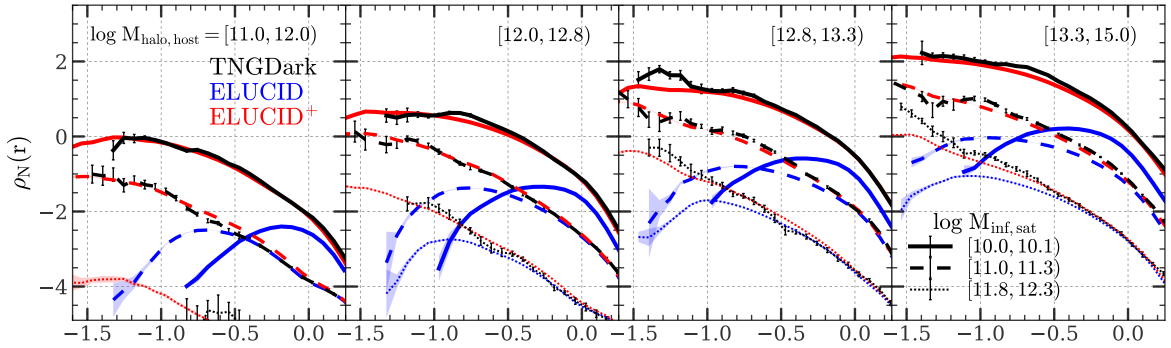


CAMTree: Recovery of Key Subhalo Statistics

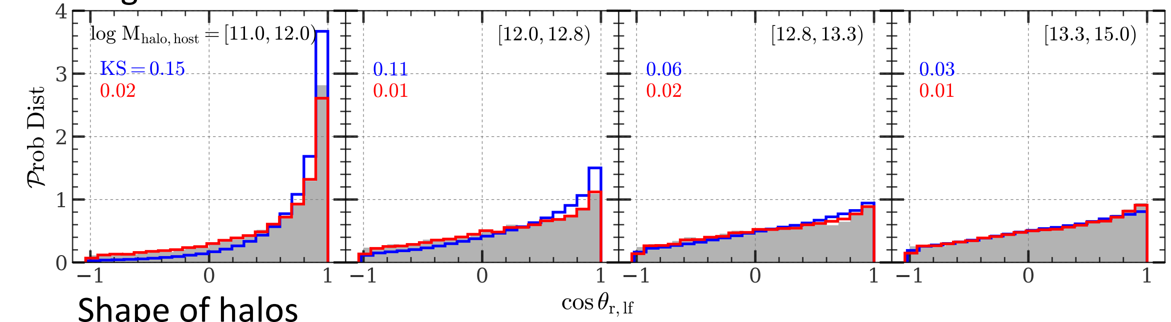
Redshift-space distortion pattern



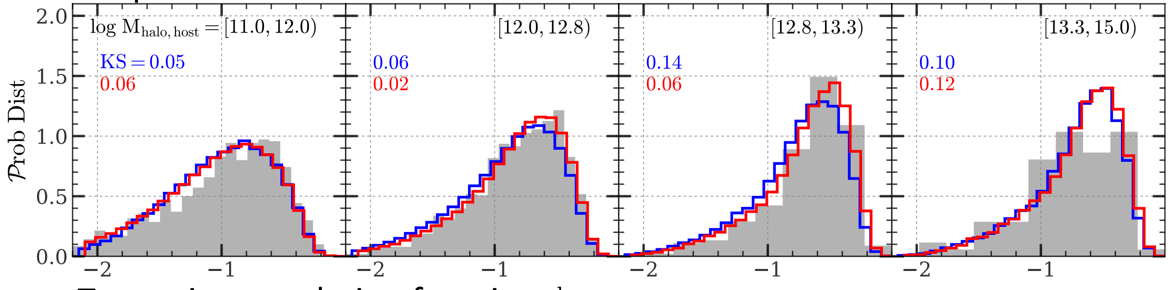
Number density profile



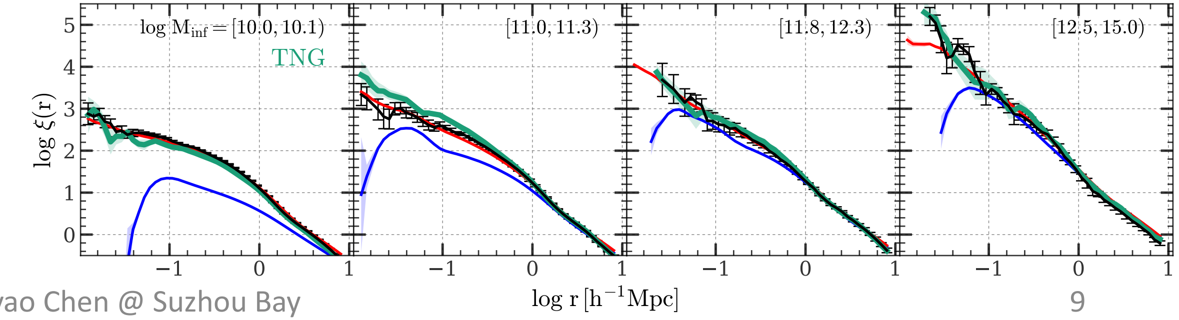
Angular distribution in halos



Shape of halos

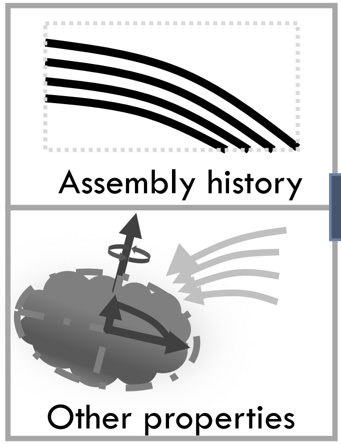


Two point correlation function



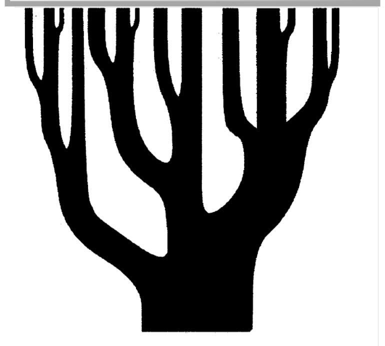
MAHGIC: a Model Adapter for the Halo-Galaxy Inter-Connection

Dark Matter Halos



Extension Algorithm

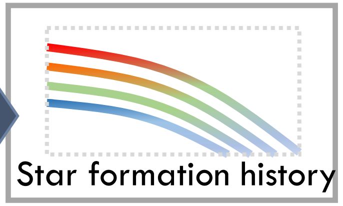
Extended Merger Tree



Halo Properties in Reduced Dimensions

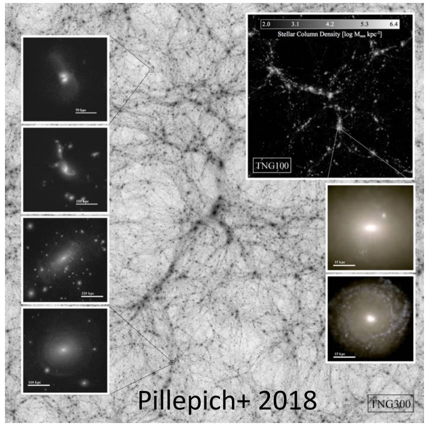
Galaxy Properties in Reduced Dimensions

Galaxies



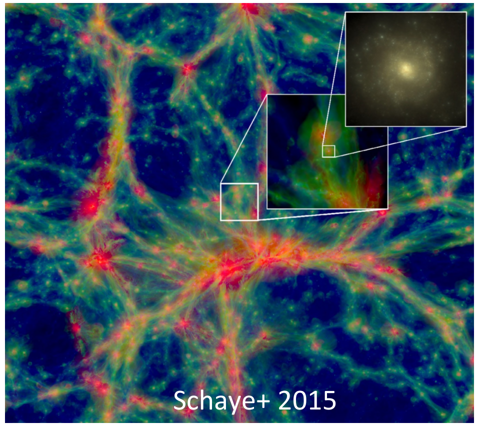
Learn representations and basis functions from a hydrodynamic simulation,
adapt them to constraining data,
and implement the model into large-volume simulation.

TNG



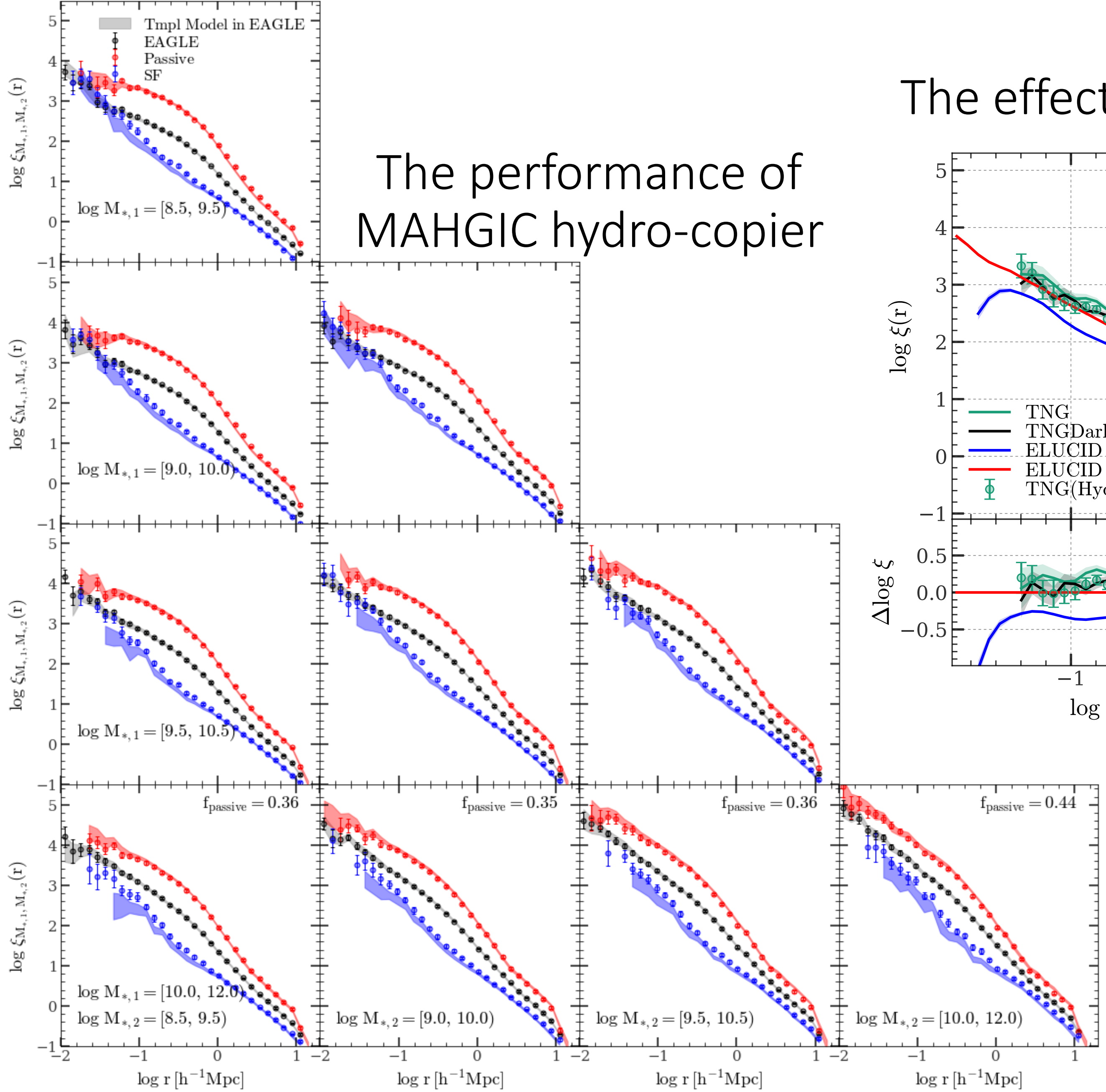
Pillepich+ 2018

EAGLE

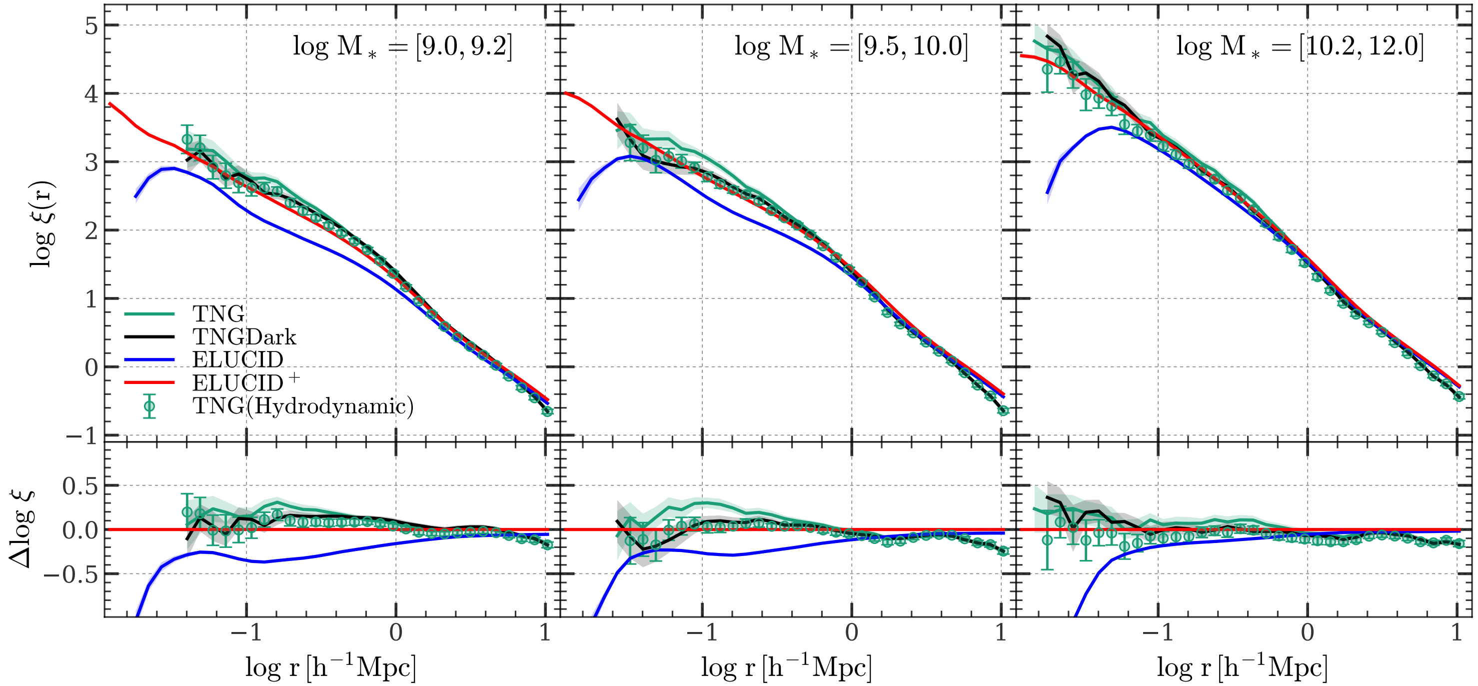


Schaye+ 2015

The performance of MAHGIC hydro-copier

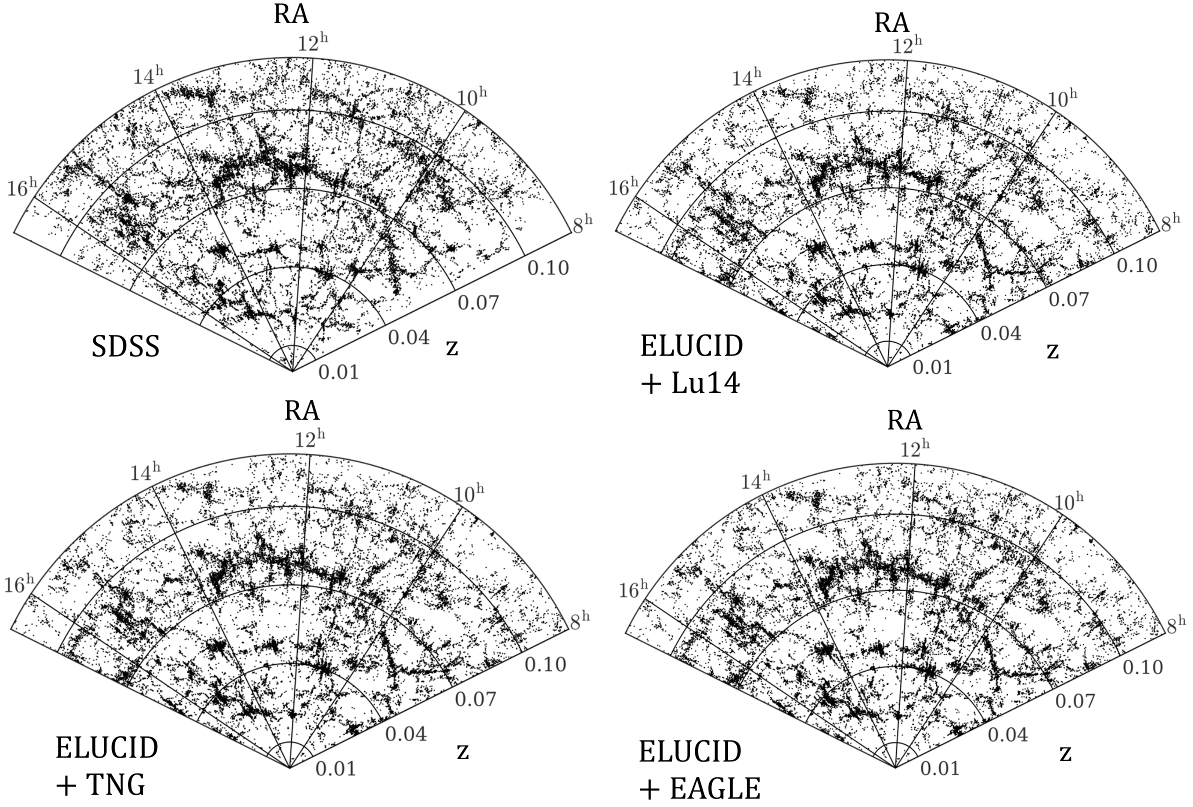


The effects of merger tree extension to halo-based model

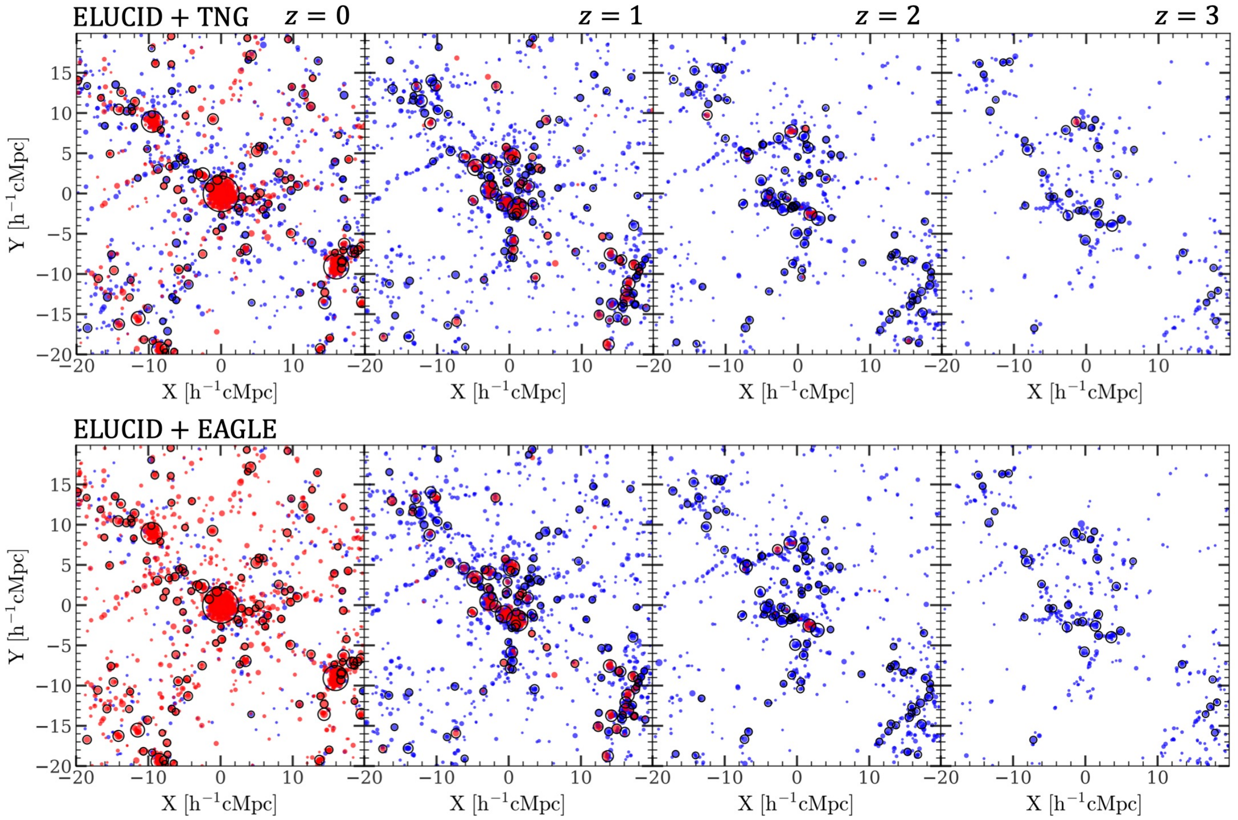


Map the distribution and assembly of galaxies with MAHGIC “copies” from different hydrodynamic simulations (TNG, EAGLE) into the reconstructed density field (ELUCID)

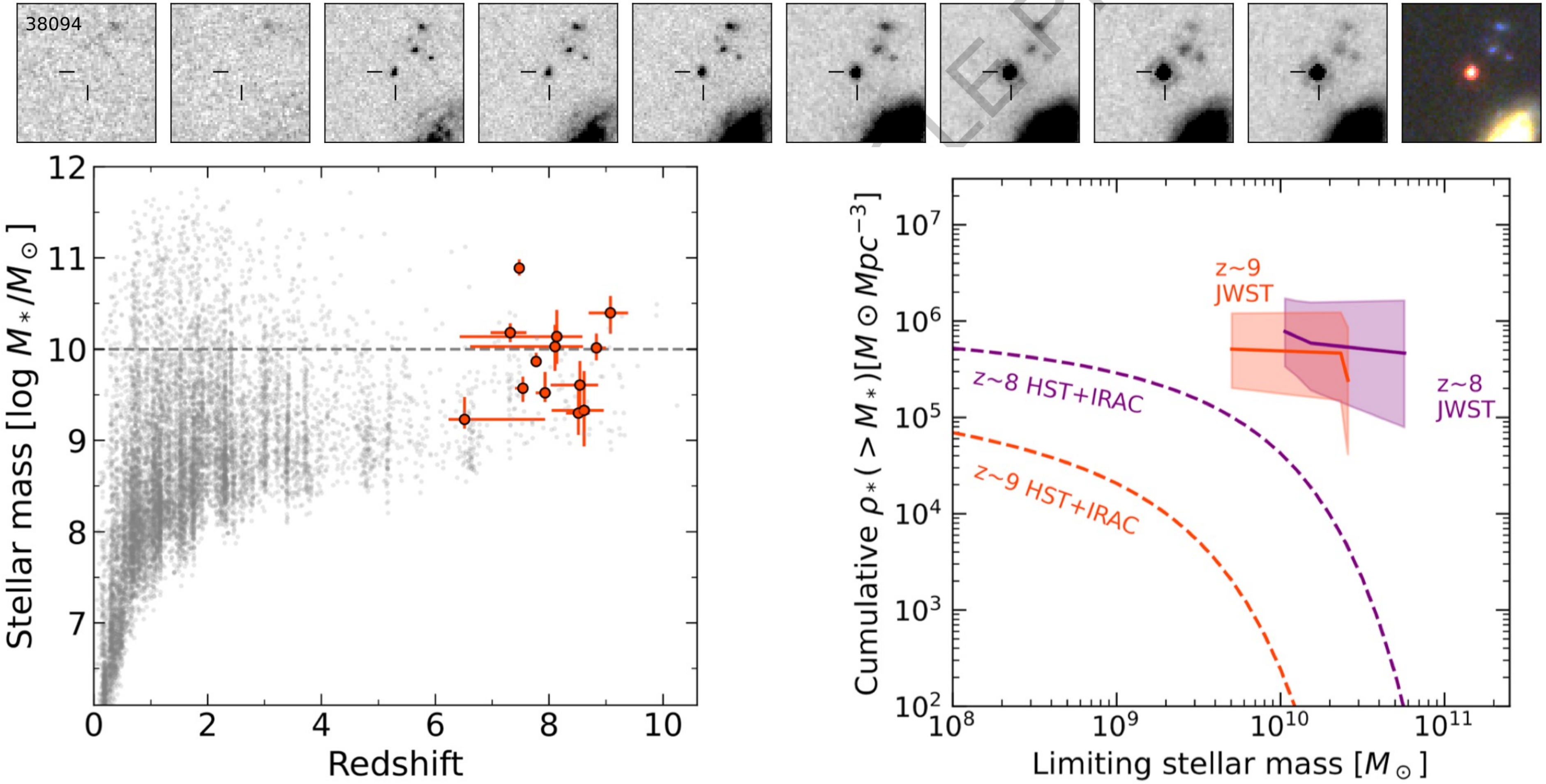
For large-scale structures



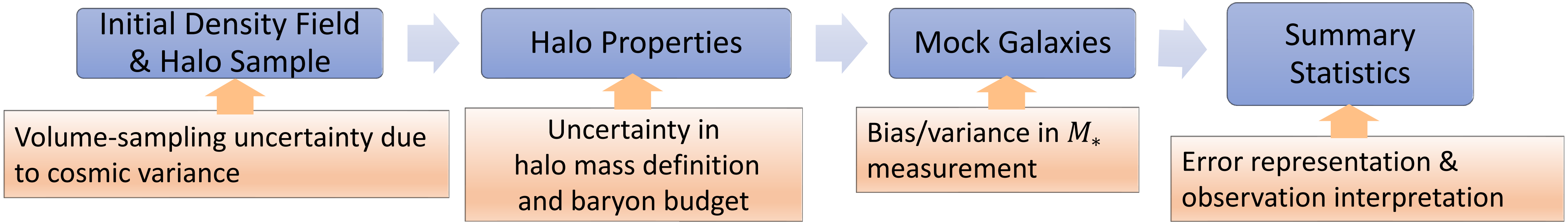
For individual clusters (take Coma as an example)



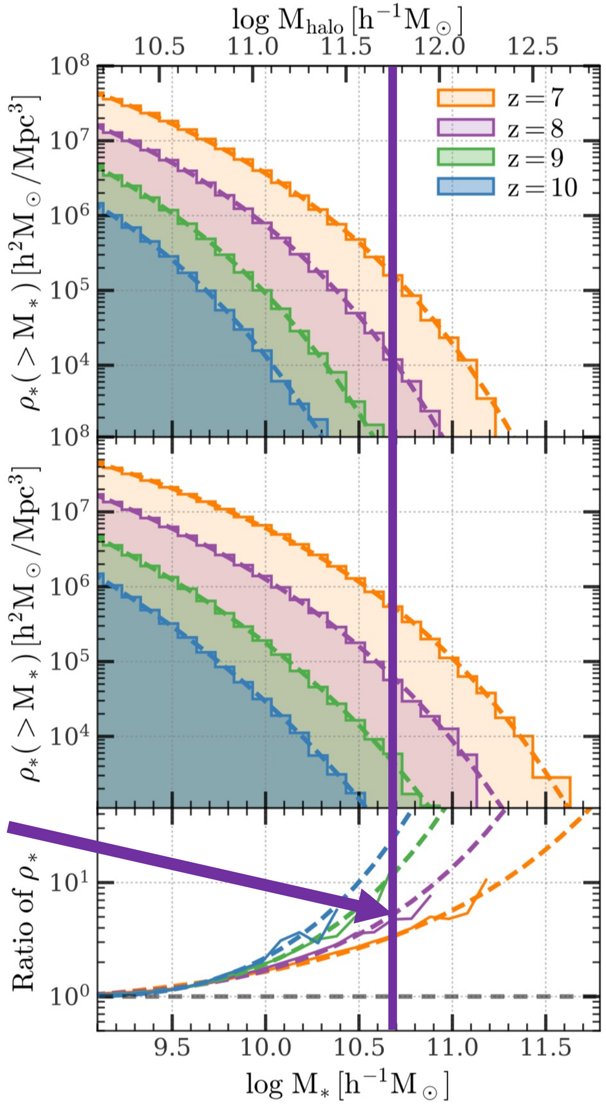
Massive Dark Matter Halos at High z
and Their Implications for Observations in the JWST Era

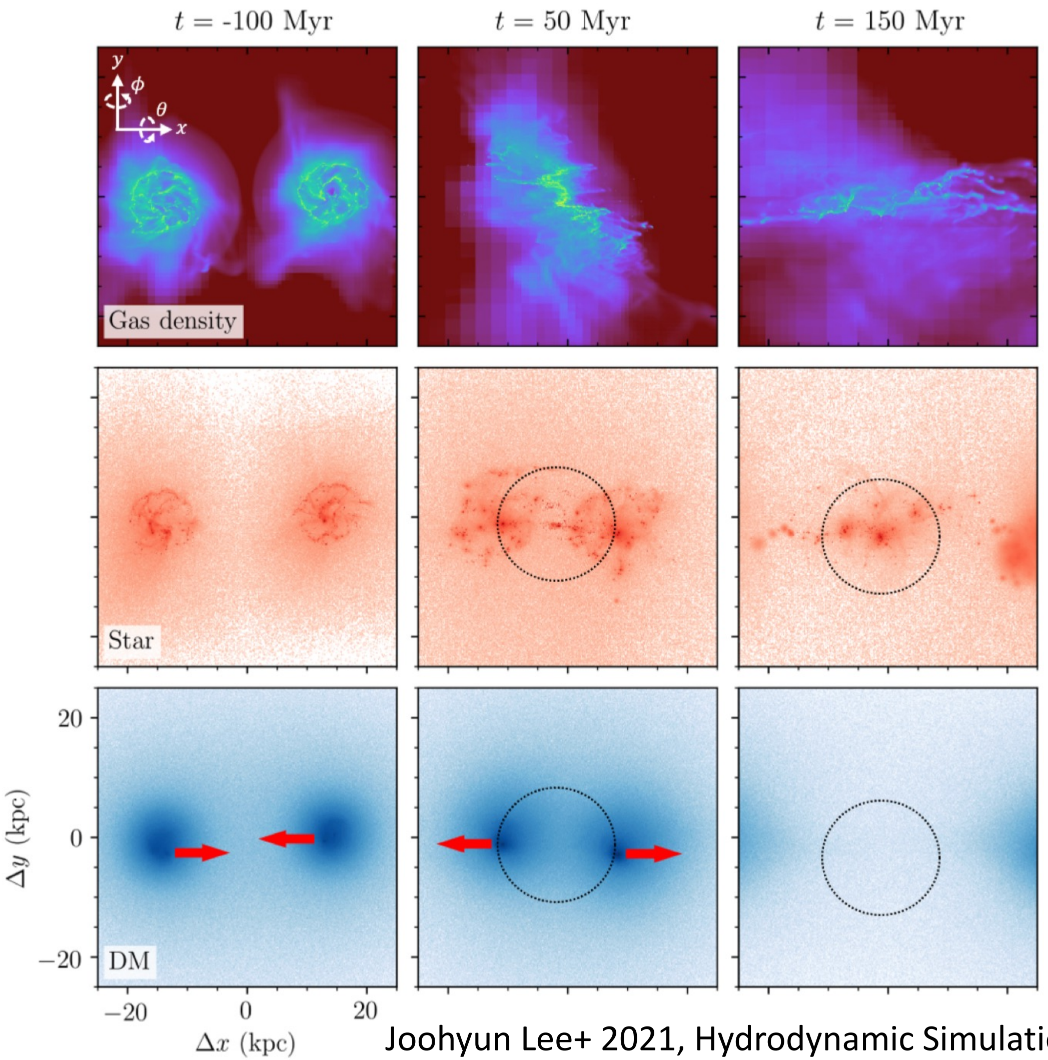


Labbe+ 2023, JWST/CEERS images and SED-fitted stellar masses



Most massive JWST/CEERS galaxy at z ~ 8

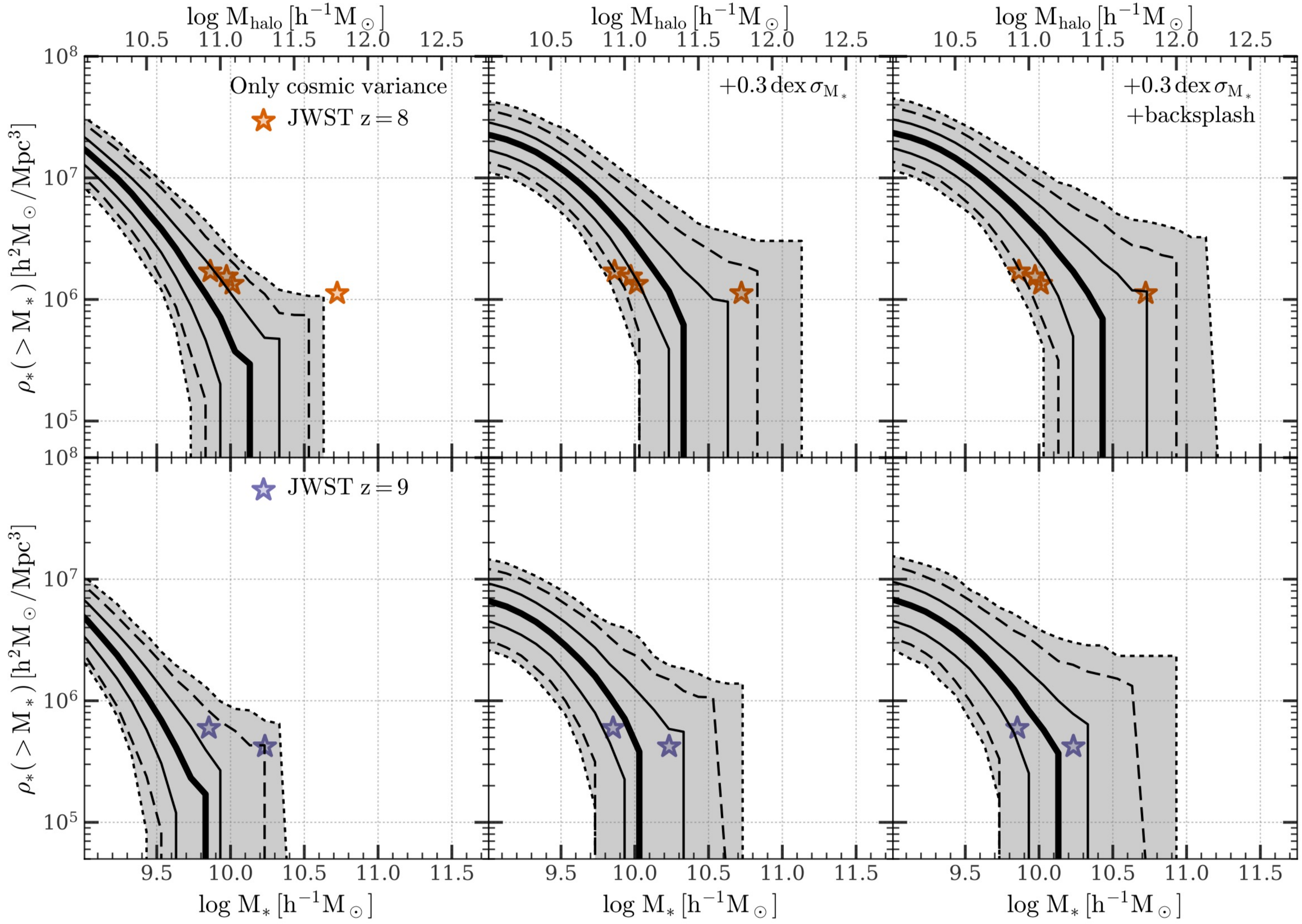




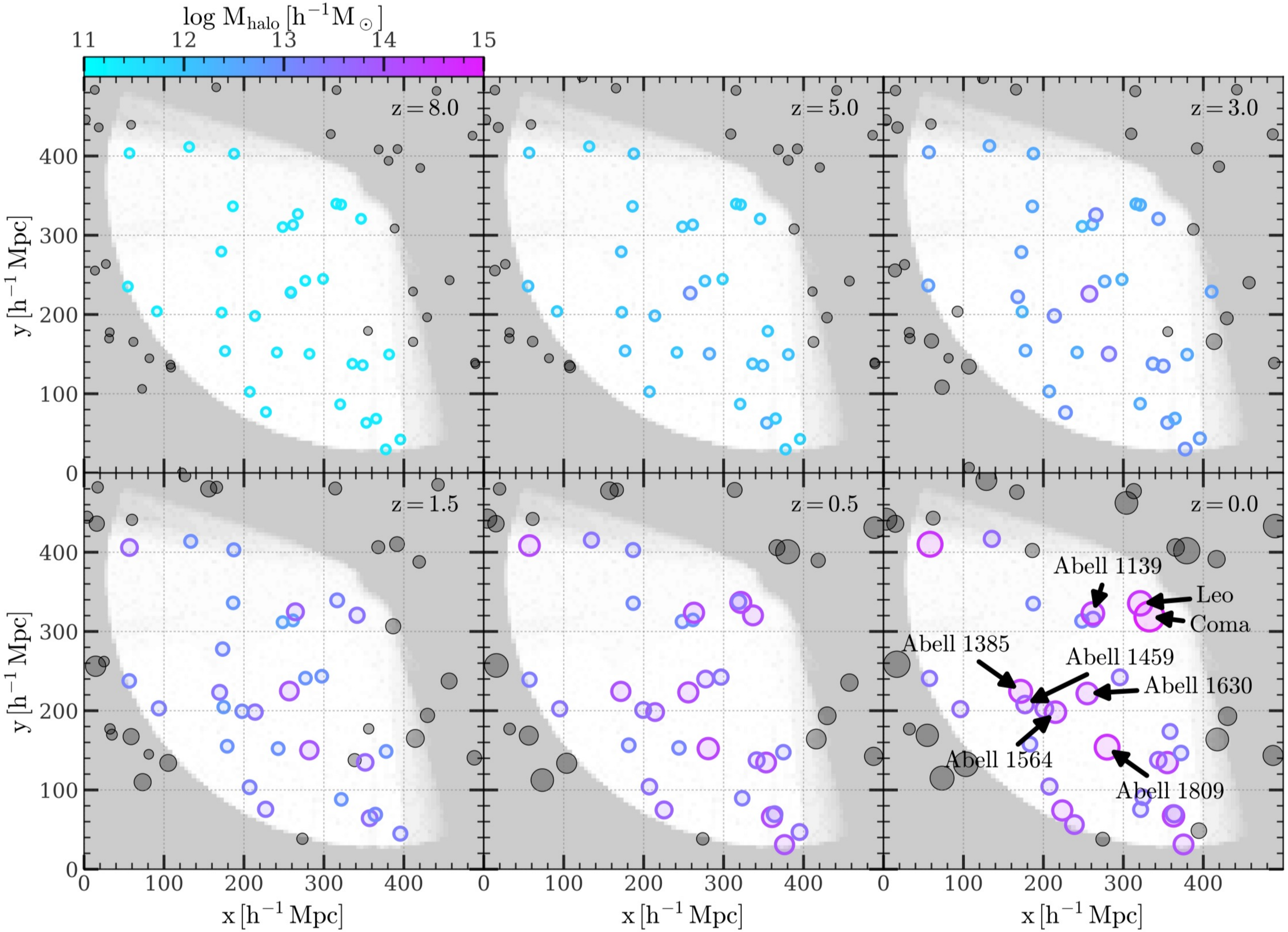
Joohyun Lee+ 2021, Hydrodynamic Simulations

A forward model is the key to completely incoporate sources of errors and fairly interpret the observations

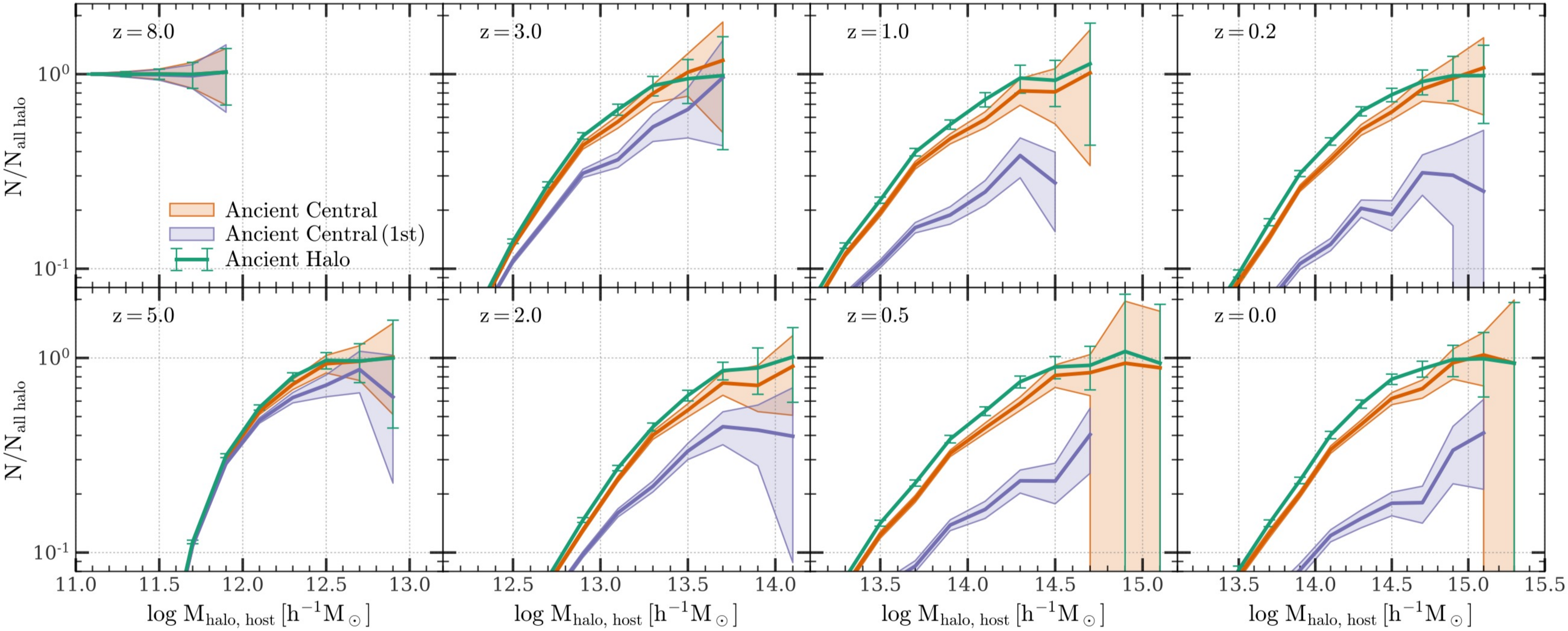
The simplest model: a constant star formation efficiency $\varepsilon_* = 50\%$.



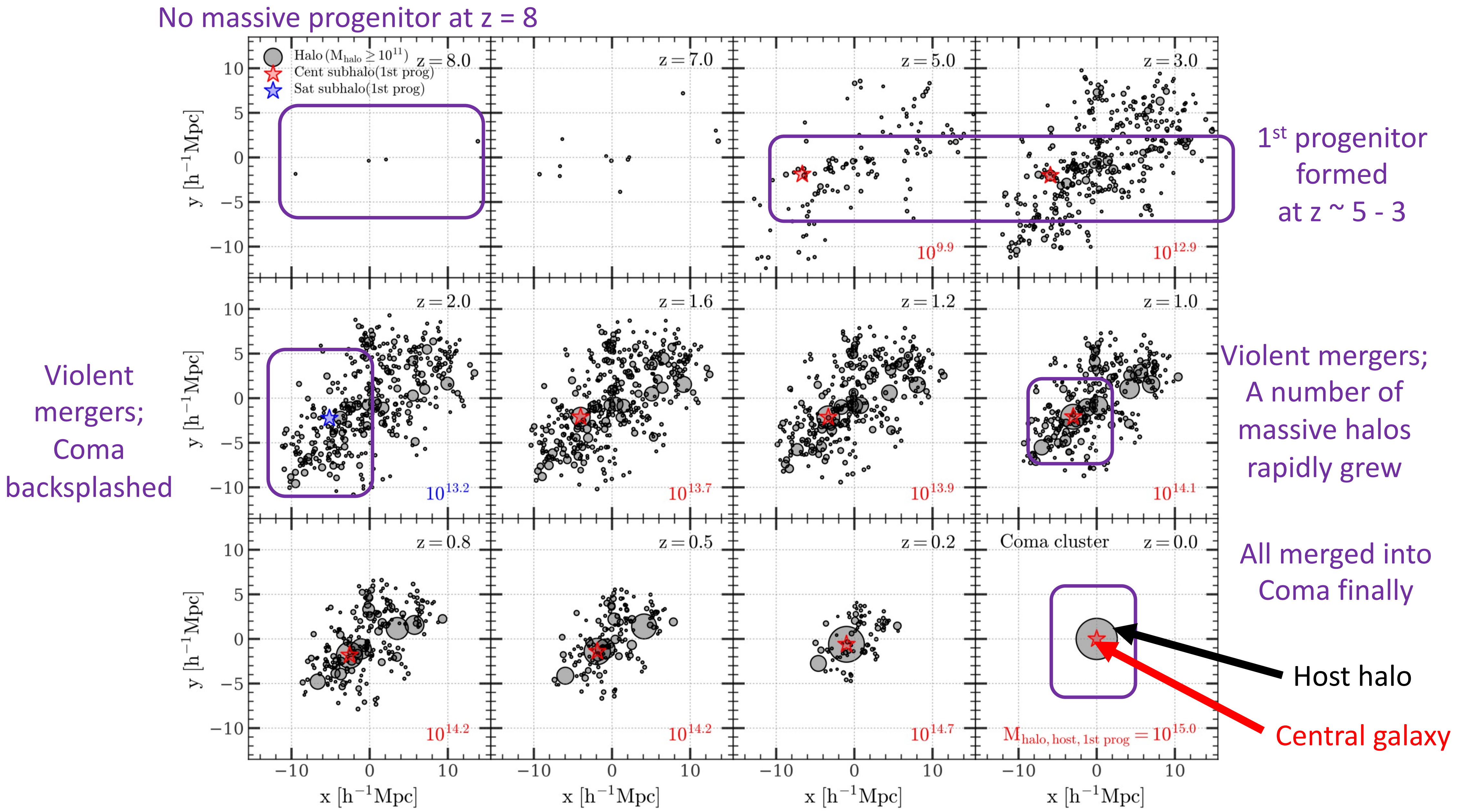
Find low-z descendants of JWST high-z counterparts using constrained simulation



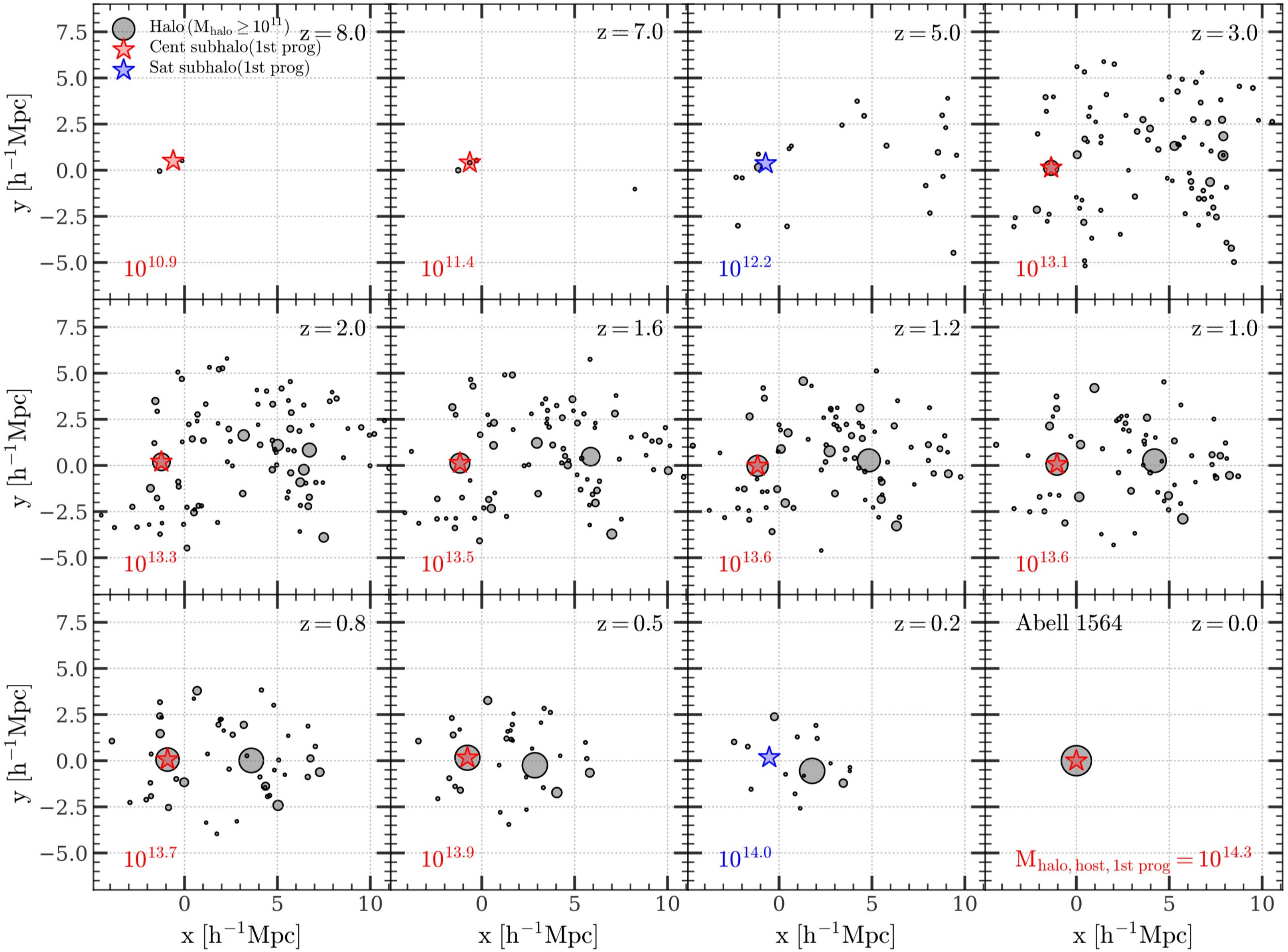
- >90% z=0 central galaxies in $\mathrm{M_h} \geq 10^{15} M_\odot$ clusters are descendants of massive z=8 JWST galaxies.
- ~30% are the main-descendants.



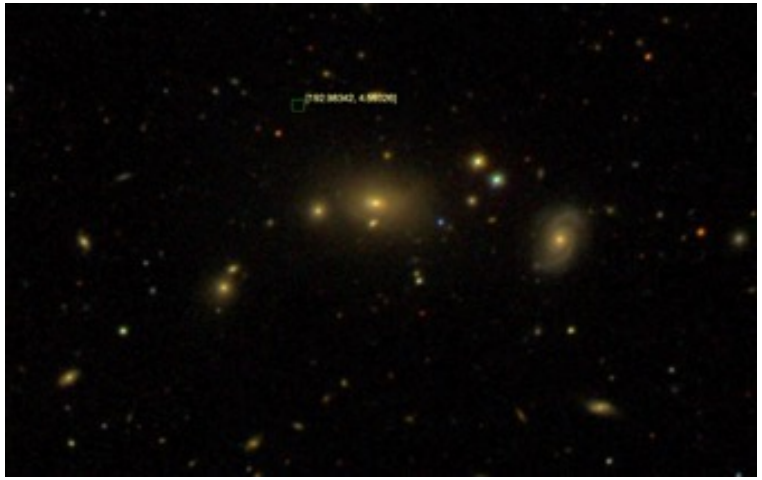
Coma's Assembly History

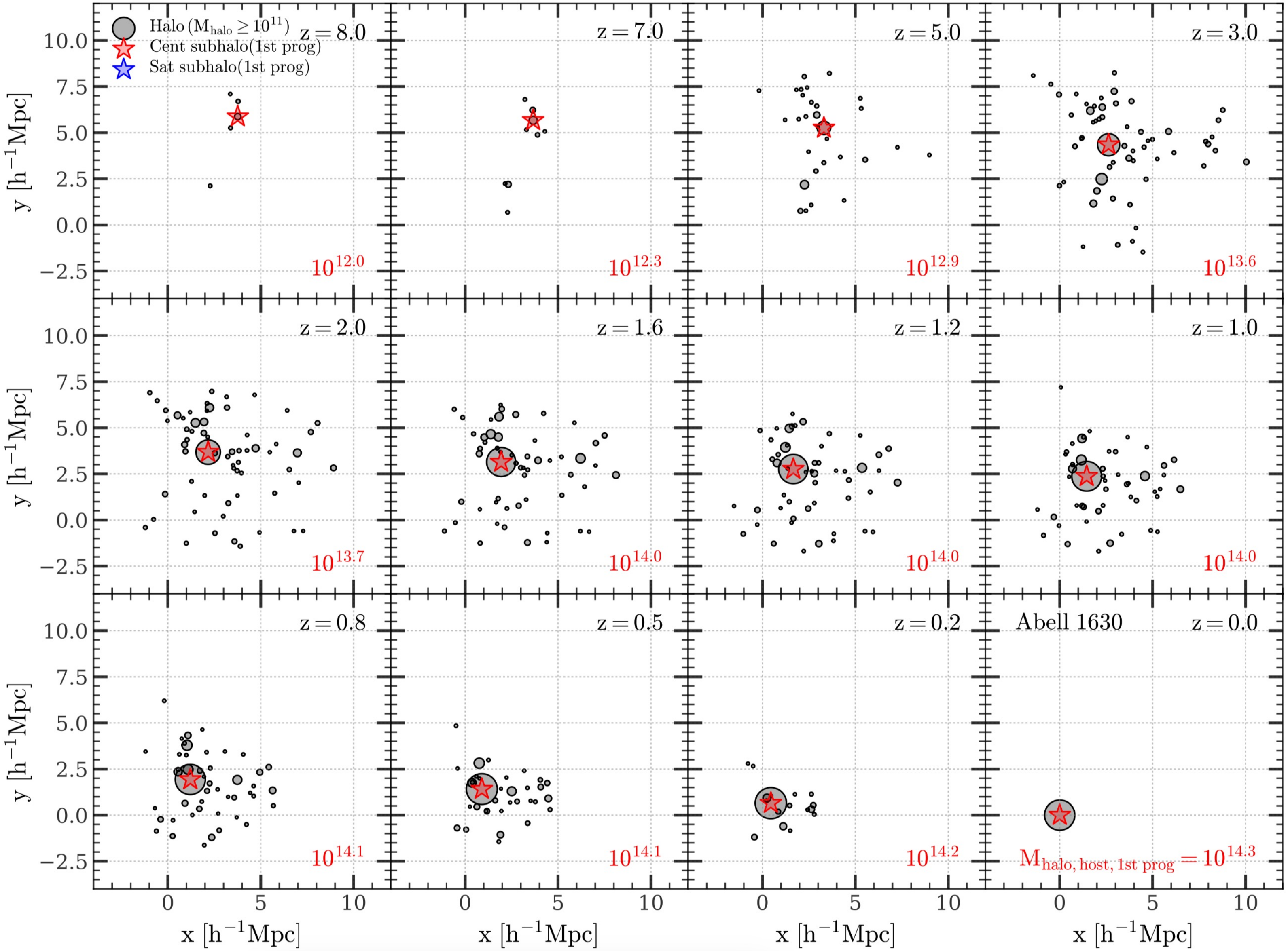


Abell 1564 - an equal-mass merger remnant



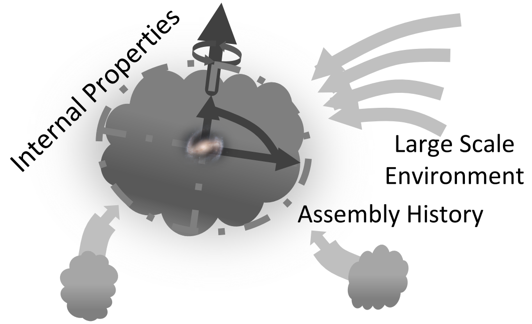
Abell 1630 - a lonely giant



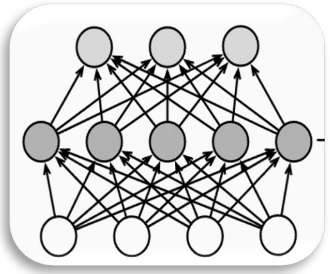


Toward Precise Halo-based Galaxy Model

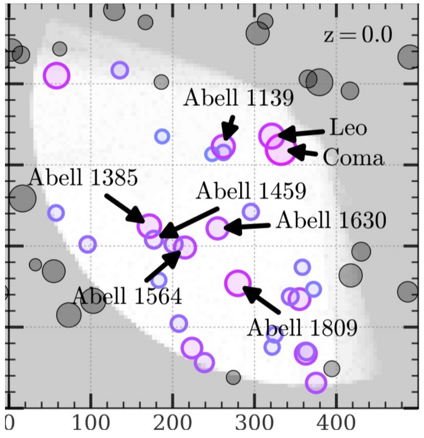
1



2



3



Dark matter halos & merger trees → Halo to galaxy mappings with tunable parameters → Galaxies and mock catalogs

CAMTree

A Conditional Abundance Matching Method of Extending Subhalo Merger Trees

MAHGIC

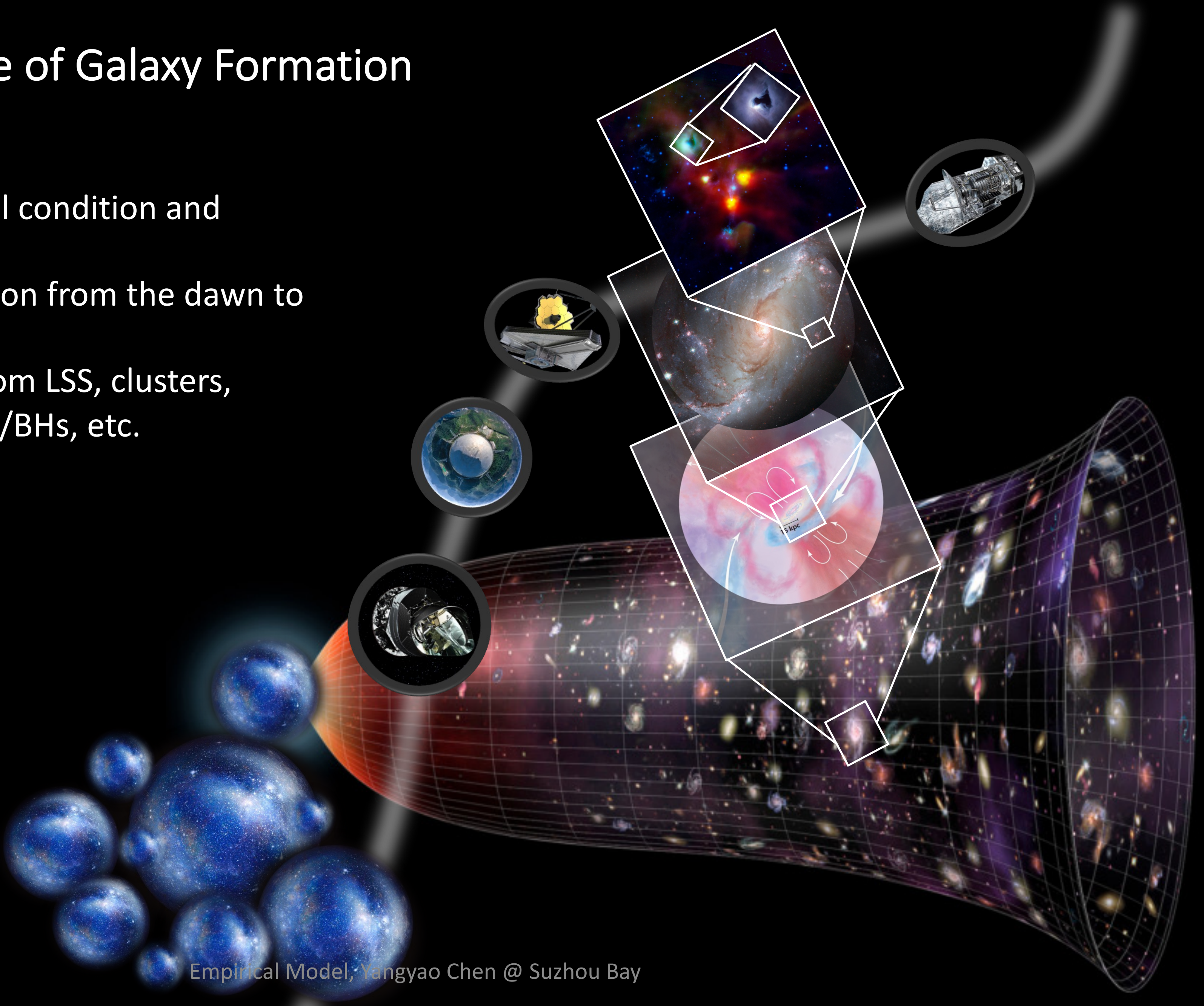
A Model Adapter for the Halo-Galaxy Inter-Connection

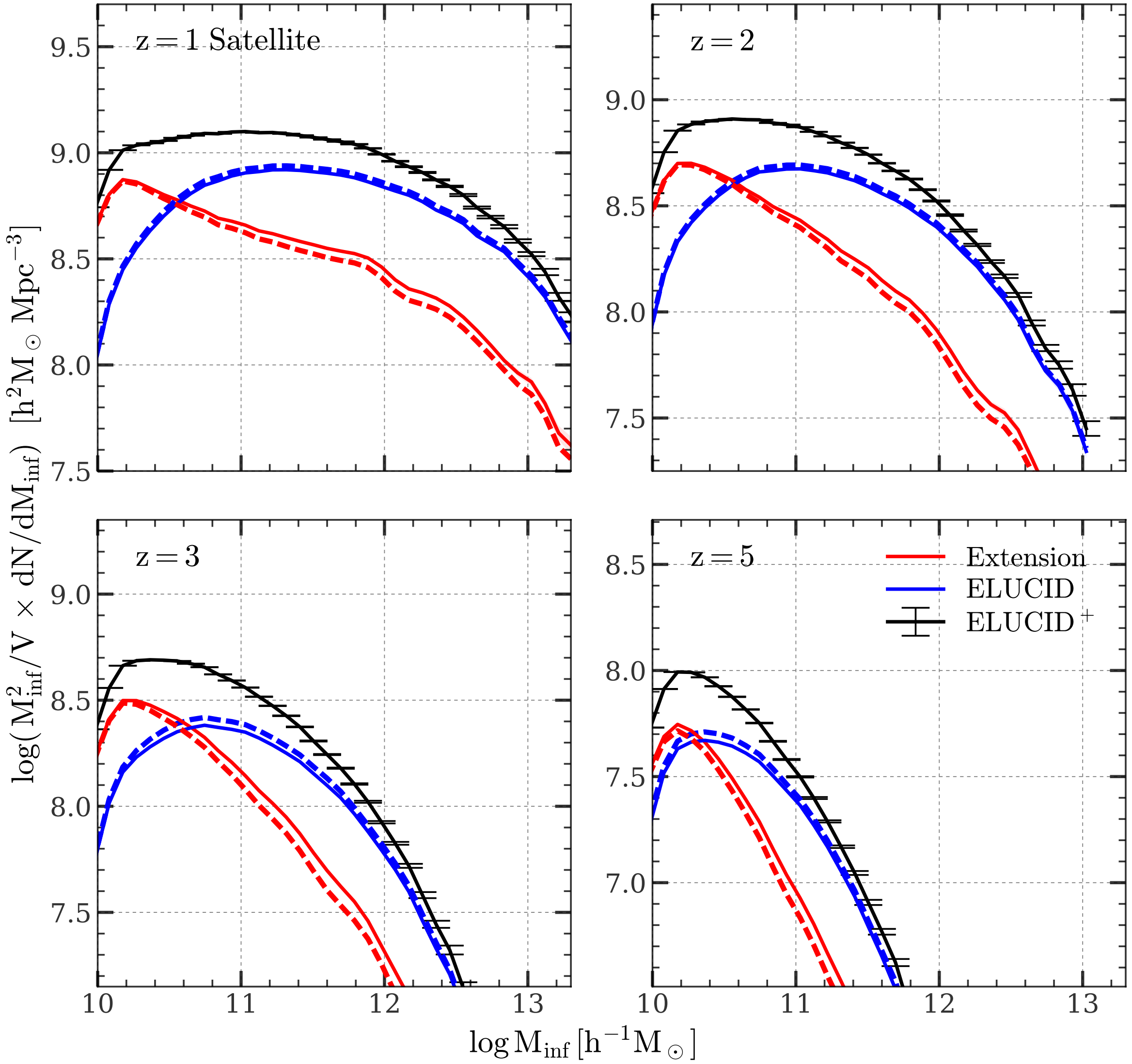
Massive Dark Matter Halos at High Redshift and Their Implications for Observations in the JWST Era

Credit: Planck, FAST, JWST, Herschel, TNG teams.

The Complex Picture of Galaxy Formation

- Cosmology as the initial condition and background.
- Long time-scale evolution from the dawn to present.
- Multi-scale coupling from LSS, clusters, galaxies, gas cloud/star/BHs, etc.




z = 1 Satellite
z = 2
z = 3
z = 5
Extension
ELUCID
ELUCID+
$\log(M_{inf}^2/V \times dN/dM_{inf})$ $[h^2 M_\odot \mathrm{Mpc}^{-3}]$
$\log M_{inf} [h^{-1} M_\odot]$
9.5
9.0
8.5
8.0
7.5
7.0
10
11
12
13

Missing-Subhalo Problem for Halo-based Models

In a real application of galaxy modeling

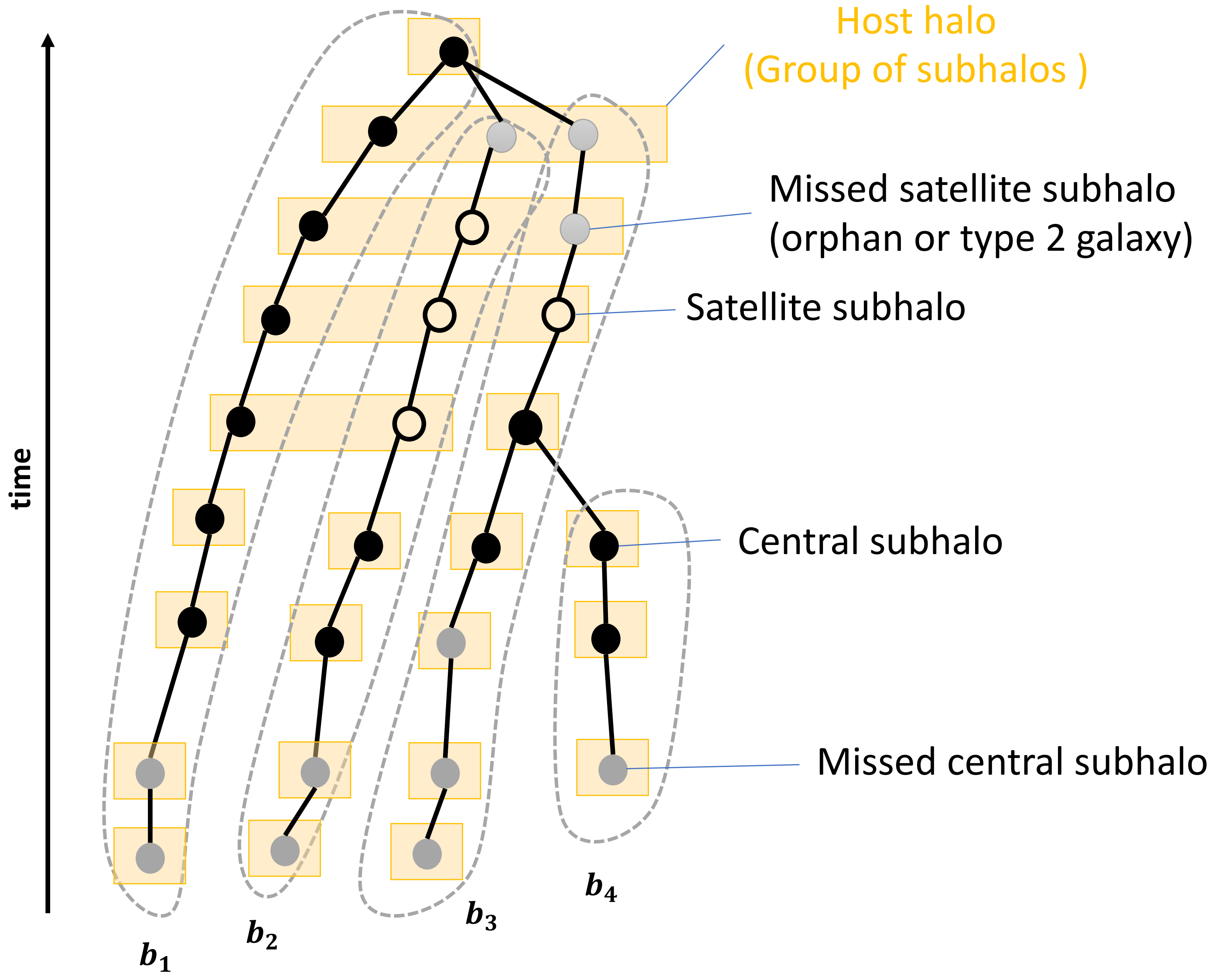
- The lower limit of sample size (simulation volume) is determined by the statistical target.
- The upper limit of CPU hours are detemined by the fundings at hand.
- Resolution upper limit = max CPU hours / min sample size.

Something that is missed with limited resolution power

- The assembly history of a central subhalo at high-z is missed, when its halo mass is below the resolution limit.
- The dynamic evolution of a satellite subhalo is missed, after it is disrupted numerically.

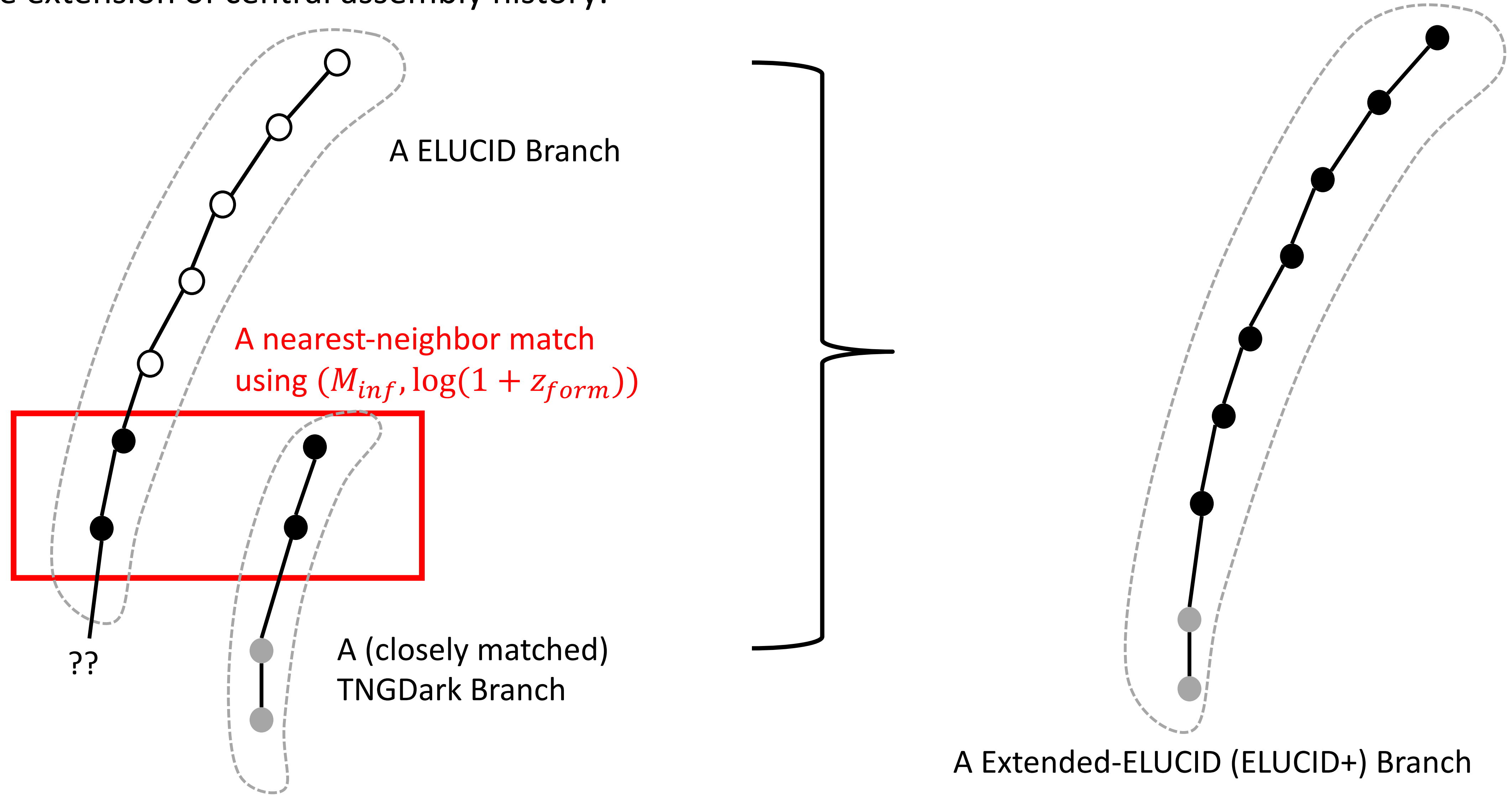
For ELUCID, $\mathrm{M_{halo,min}} = 10^{10} h^{-1} M_{\odot}$
For TNG100-1-Dark, $\mathrm{M_{halo,min}} = 2\times10^{8} h^{-1} M_{\odot}$

A example subhalo merger tree



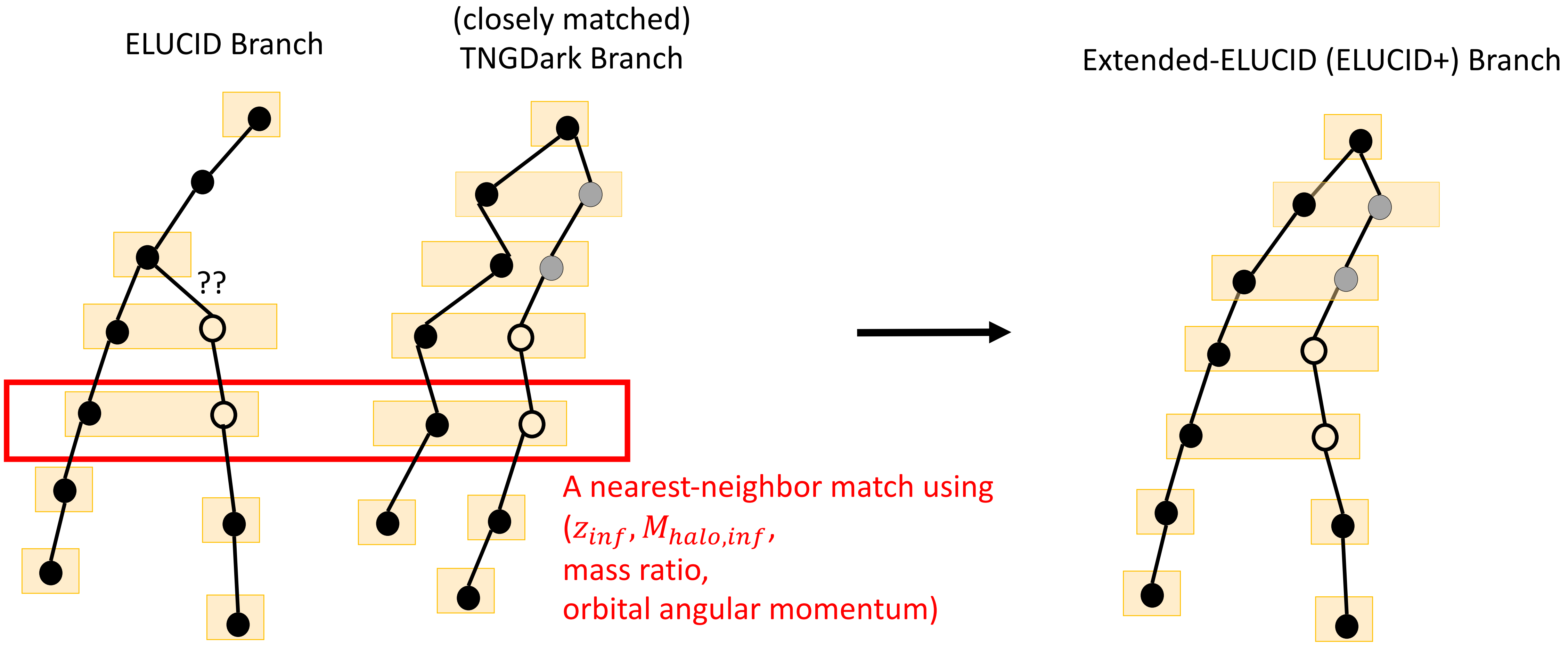
The Method: Learn From a High-resolution Simulation to Extend a Low-resolution Simulation

The extension of central assembly history:



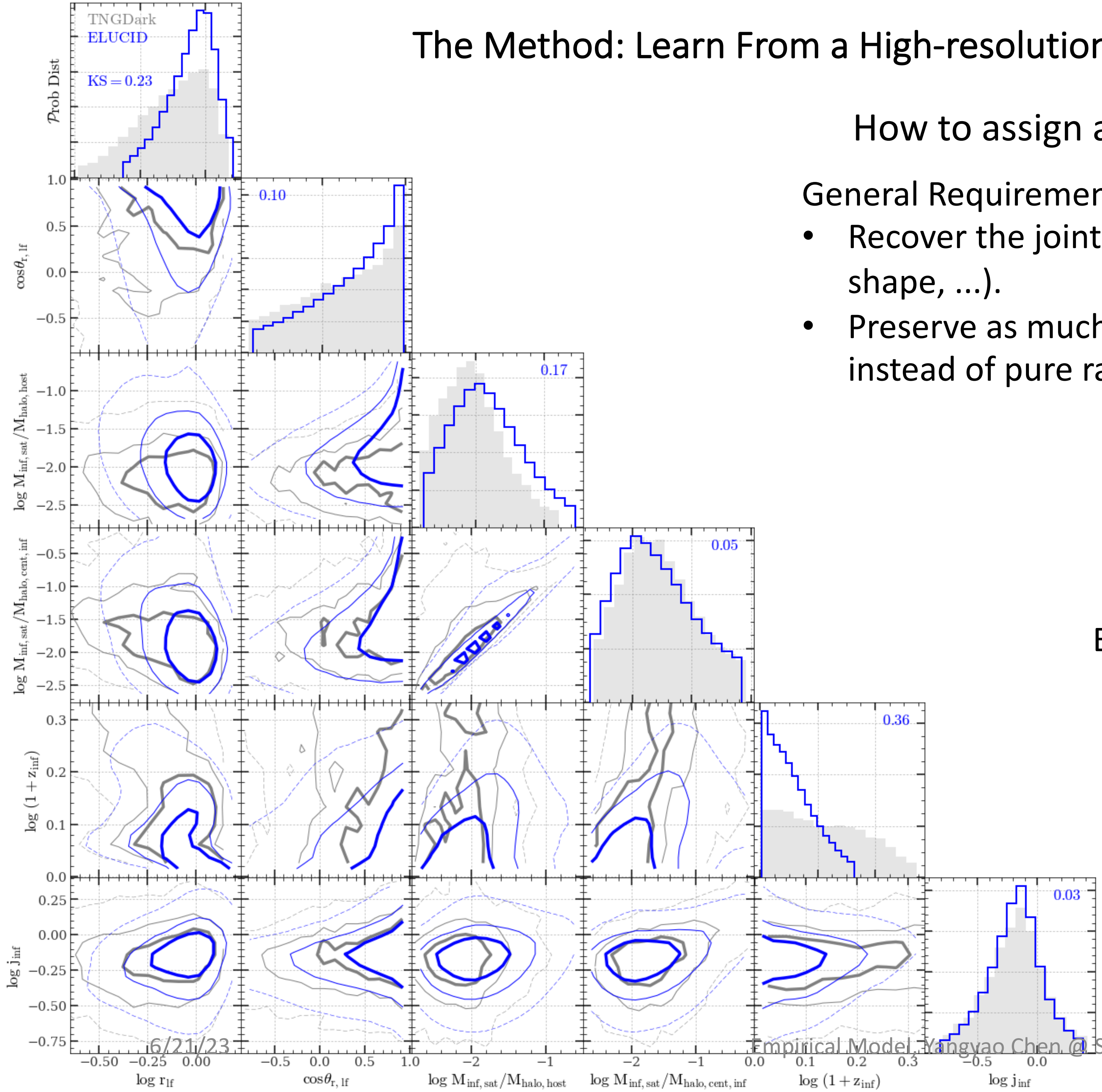
The Method: Learn From a High-resolution Simulation to Extend a Low-resolution Simulation

The extension of satellite dynamic evolution



A nearest-neighbor match using
(z_{inf}, $M_{halo,inf}$,
mass ratio,
orbital angular momentum)

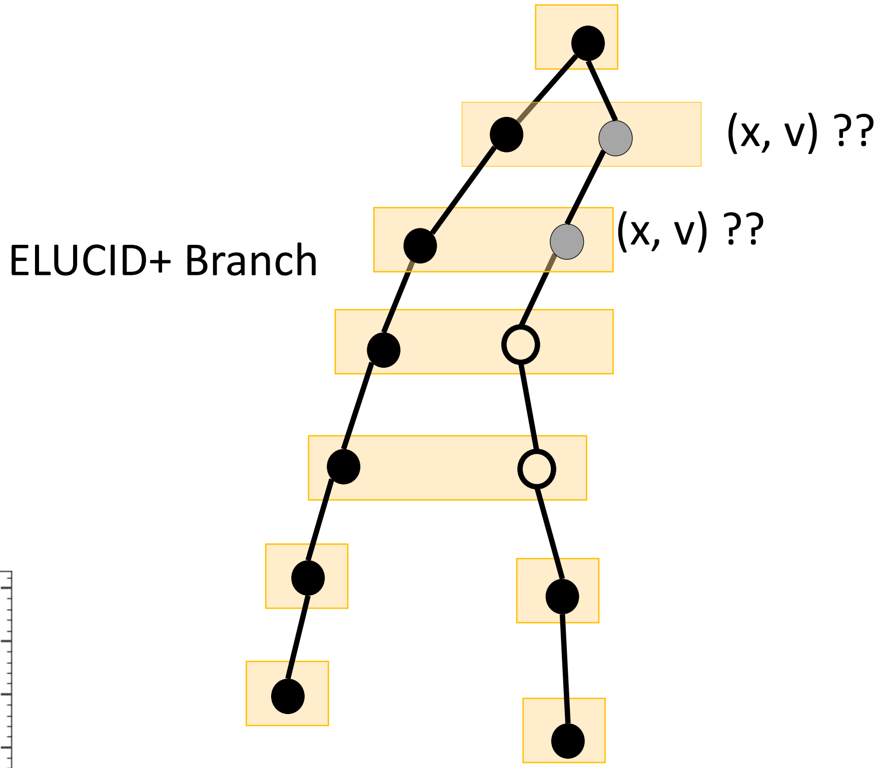
The Method: Learn From a High-resolution Simulation to Extend a Low-resolution Simulation



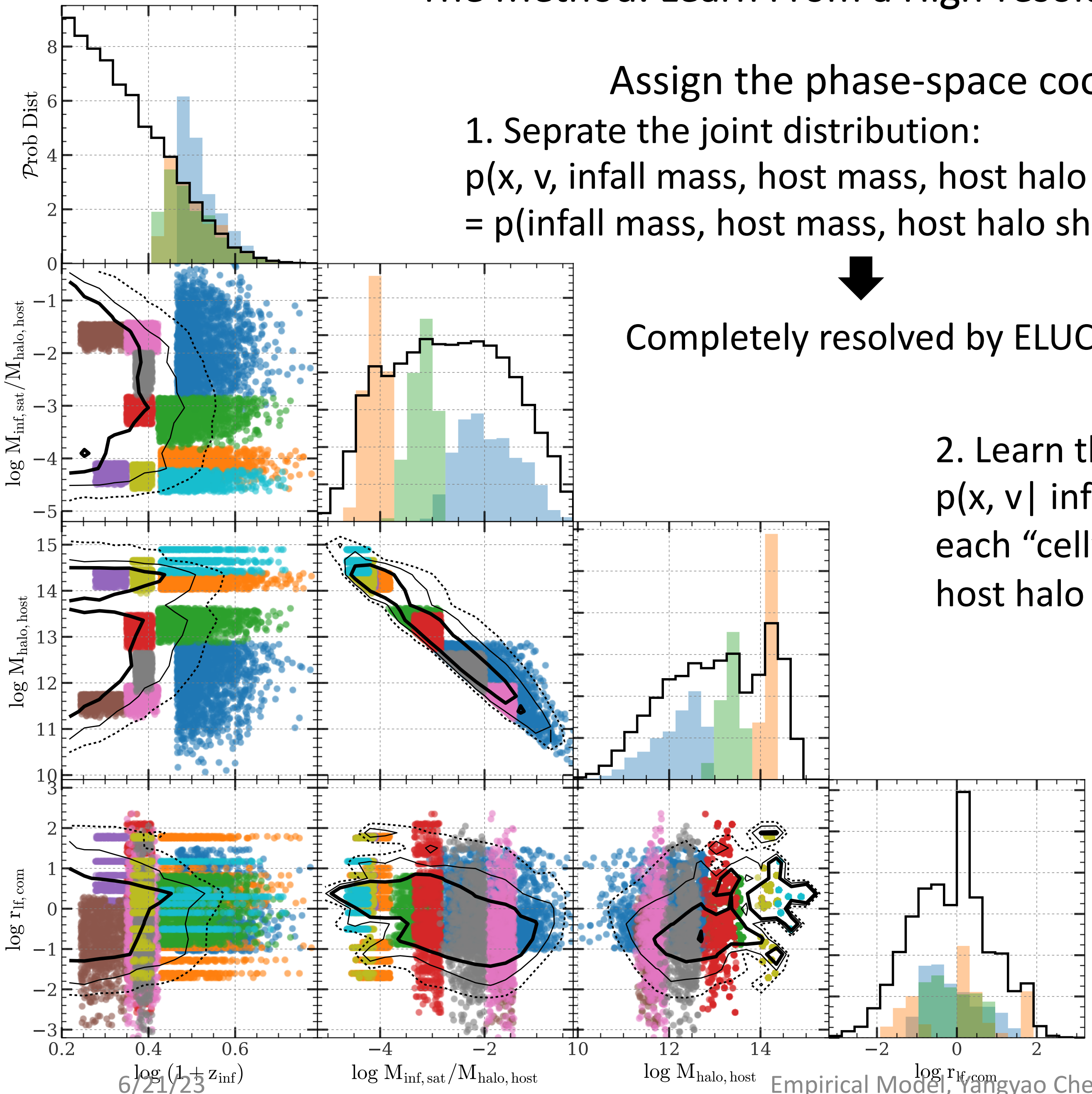
How to assign a phase-space coordinate to satellite?

General Requirements:

- Recover the joint distribution p(x, v, infall mass, host mass, host halo shape, ...).
- Preserve as much information as possible from original simulation, instead of pure random sampling from the joint distribution.



The Method: Learn From a High-resolution Simulation to Extend a Low-resolution Simulation



ELUCID satellites in cells

Assign the phase-space coordinates with conditional abundance matching

1. Seprate the joint distribution:
p(x, v, infall mass, host mass, host halo shape, ...)
= p(infall mass, host mass, host halo shape, ...) p(x, v | infall mass, host mass, host halo shape, ...)

Completely resolved by ELUCID

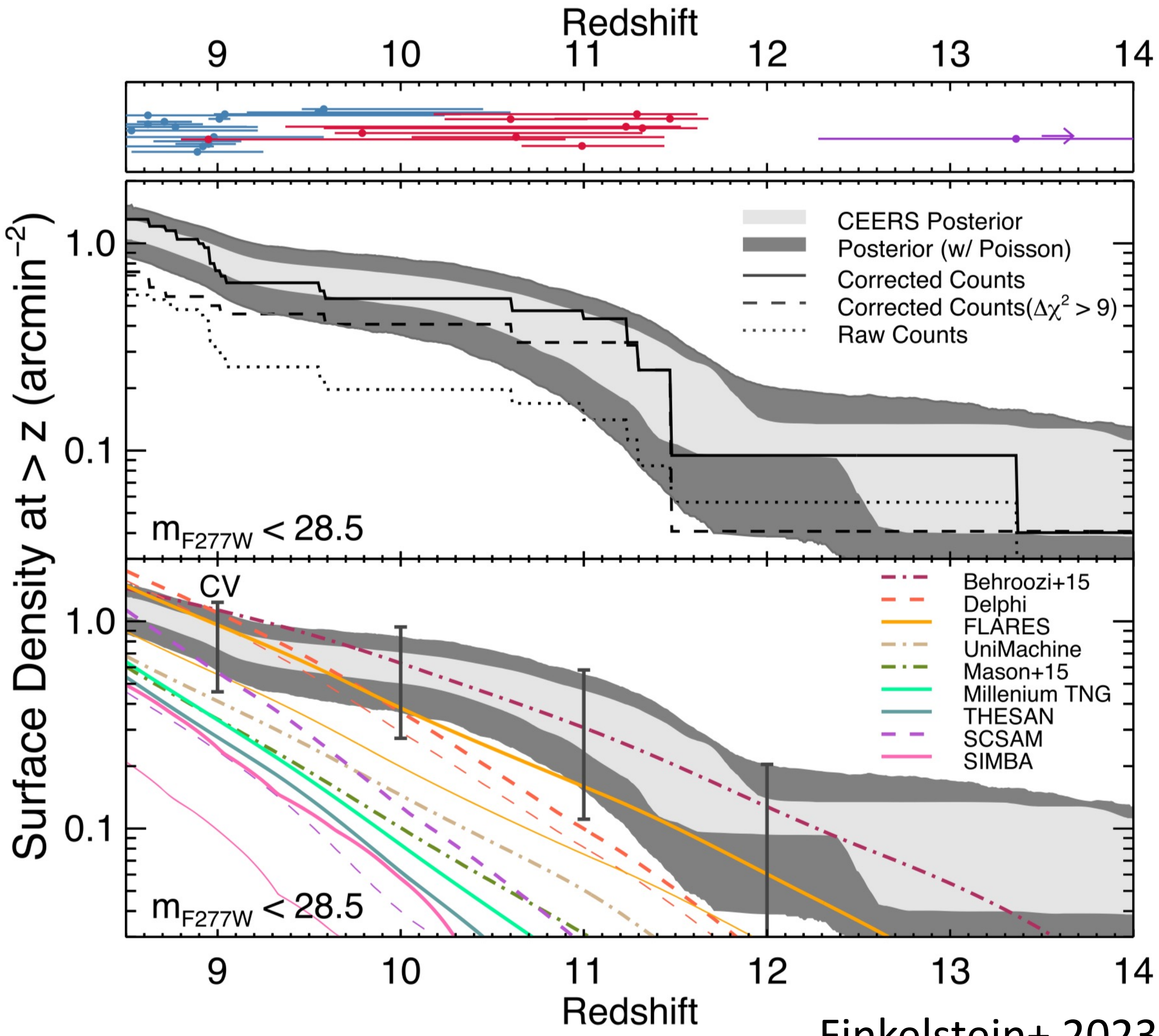
Partly missed by ELUCID

2. Learn the missed part from TNGDark:
p(x, v| infall mass, host mass, host halo shape, ...) is estimated in each "cell" of the conditioning variable (infall mass, host mass, host halo shape, ...).

3. In each cell, match each ELUCID-resolved satellite to a TNGDark one (in some predefined order), and remove them from the cell.

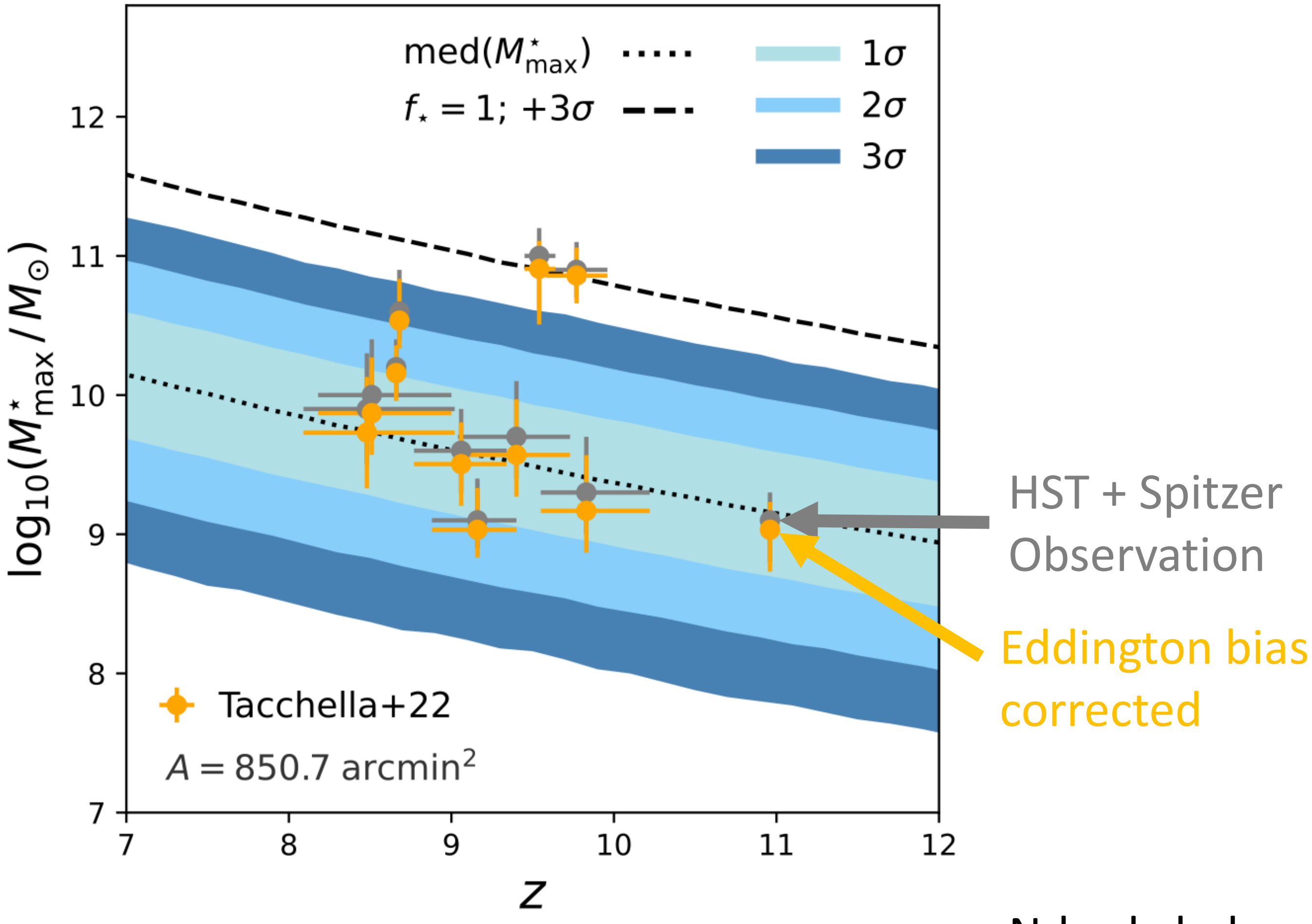
4. Randomly match ELUCID-extended satellites to the remaining ones of TNGDark.

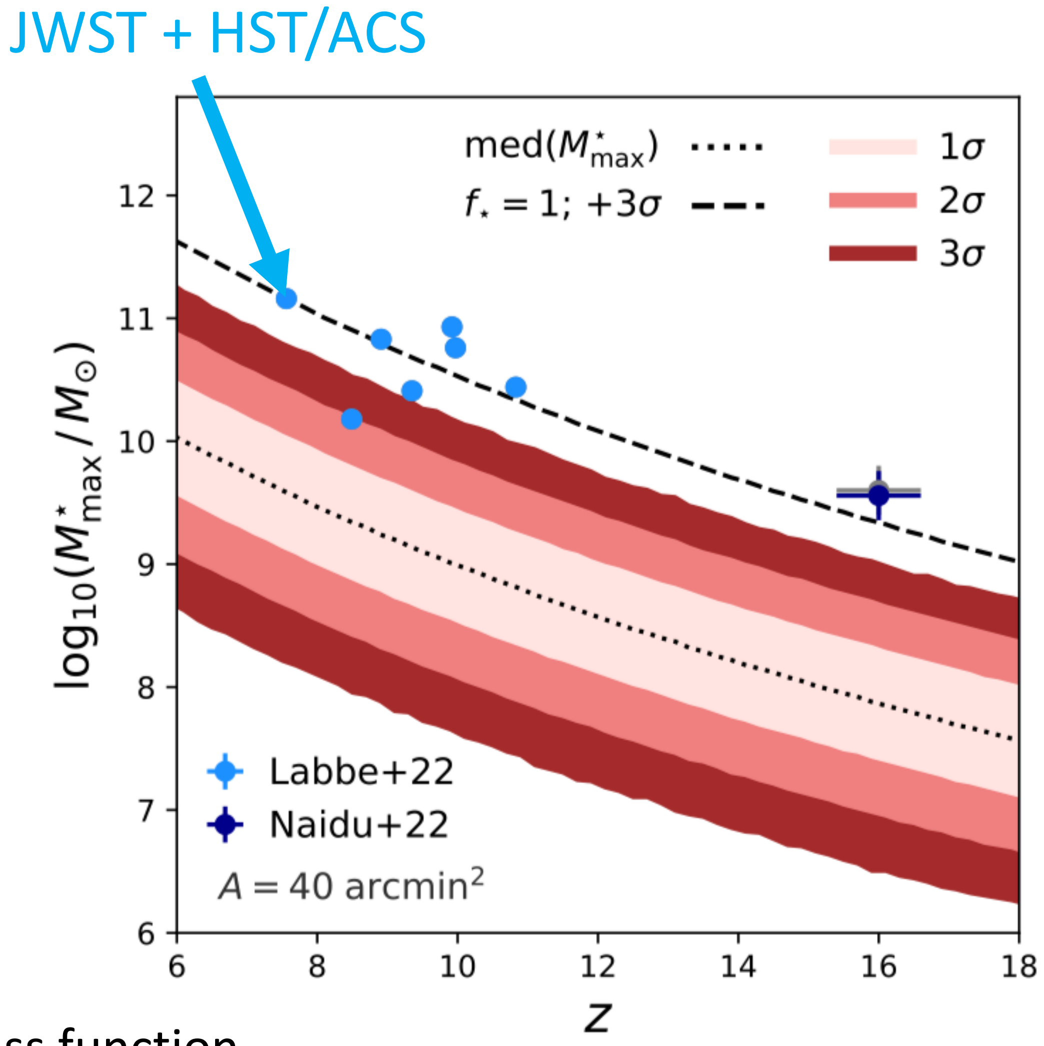
z = 8 – 14 Cosmic UV Luminosity Densities from JWST/CEERS Samples



Finkelstein+ 2023, CEERS Key Paper

Tensions with ΛCDM



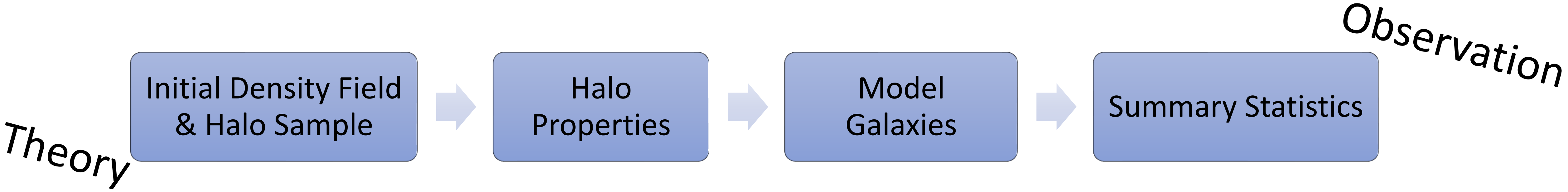


Lovell+ 2022

- N-body halo mass function.
- Maximal star formation efficiency $\epsilon_* = 1$.
- Extreme value statistics.

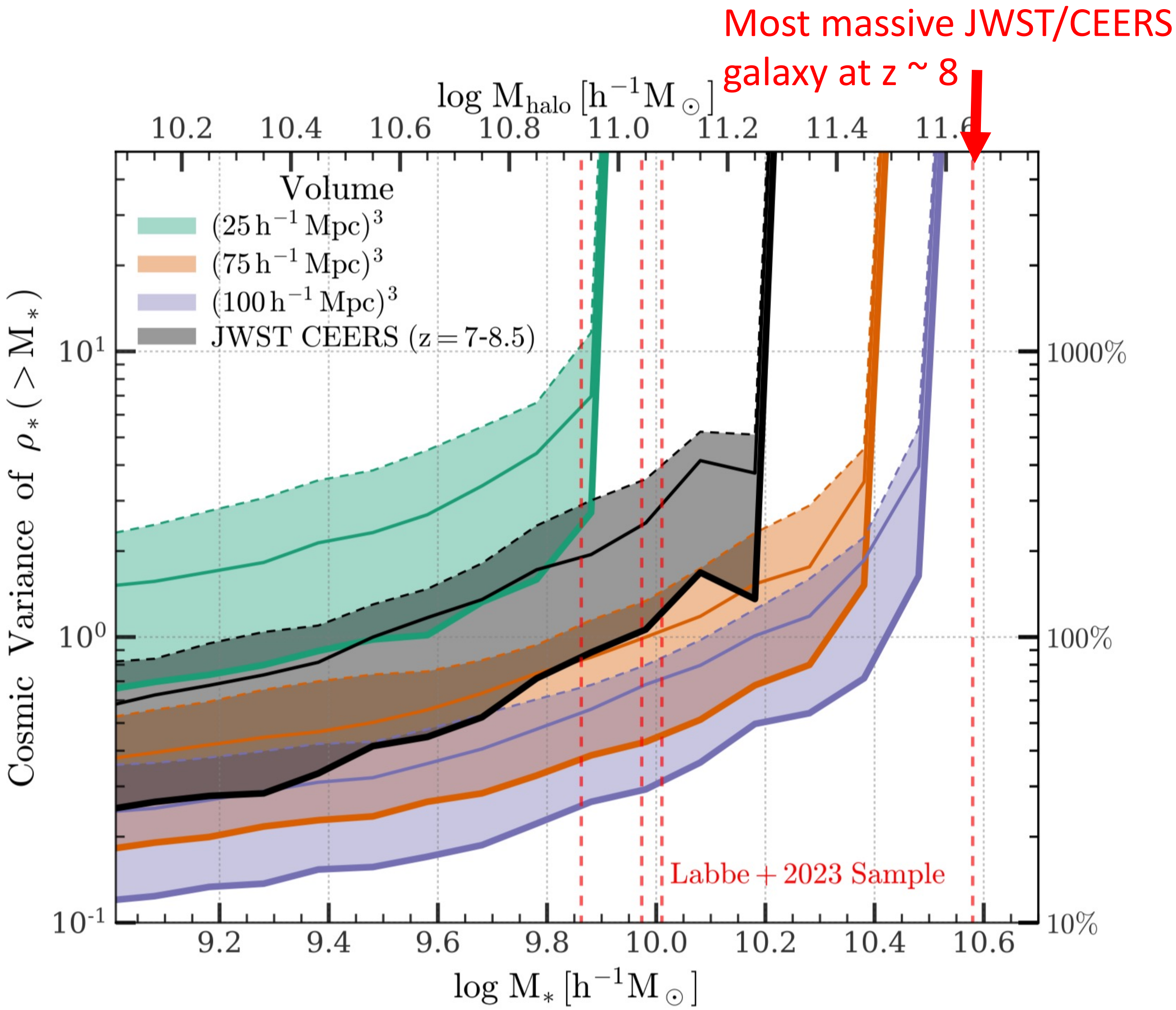
How to reproduce JWST observations within ΛCDM paradigm?

- Split theory-to-observation mocking into steps.
- Inject ΛCDM-compatible uncertainties into each step.
- Propagate uncertainties to the final observable forwardly and make comparison with JWST results.



Method: repeated sampling from a large N-body simulation with given volume and survey geometry.

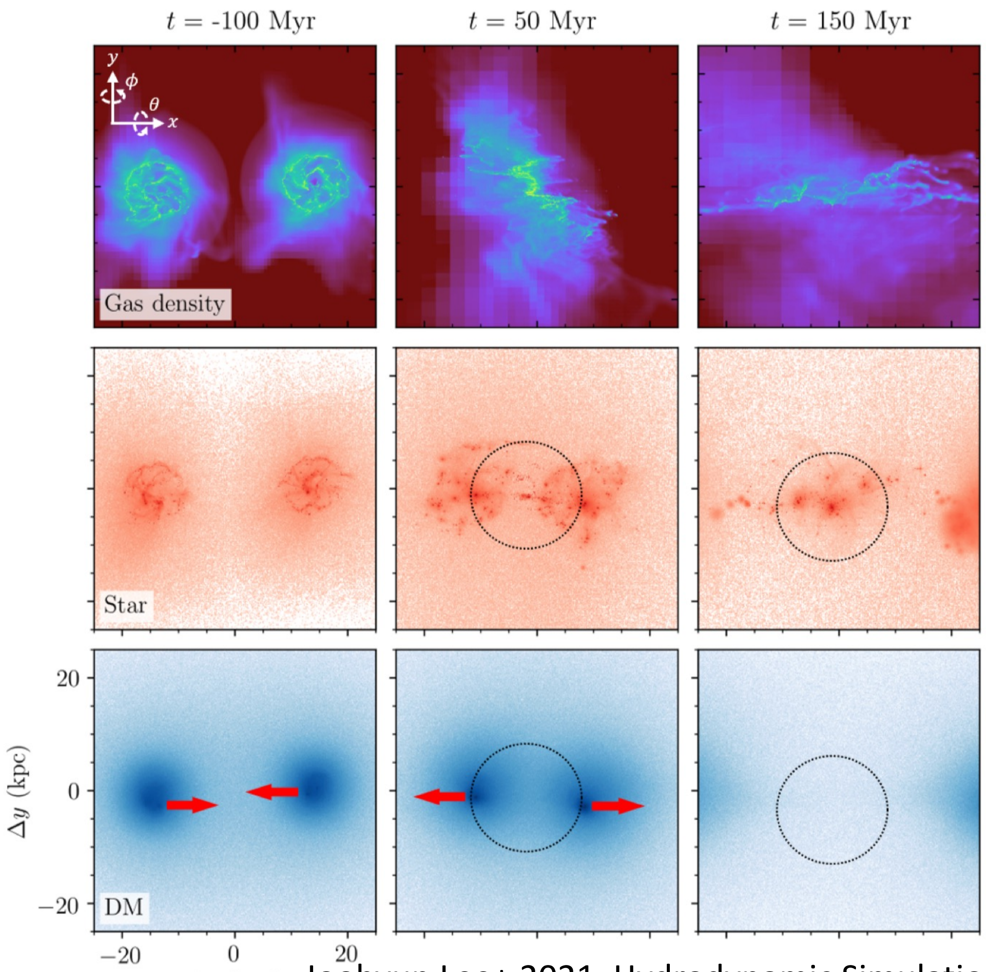
Result: 1σ cosmic variance at $z\sim8$
- 30% at low mass.
- 100% for the 3 JWST/CEERS galaxies.
- Divergent for the most massive JWST/CEERS galaxy.



Solid, dashed and dotted: 1, 2 and 3 sigma

Initial Density Field & Halo Sample → Halo Properties → Model Galaxies → Summary Statistics

Uncertainty in Halo Mass Definition

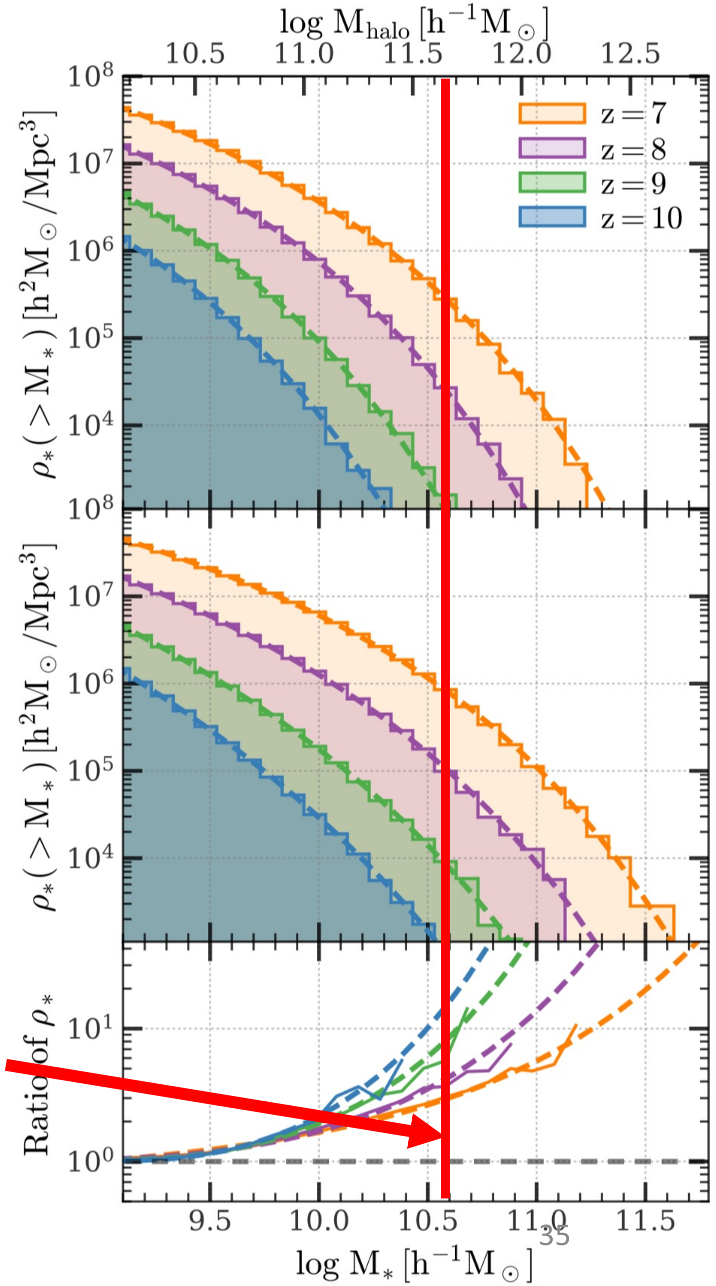


Joohyun Lee+ 2021, Hydrodynamic Simulations

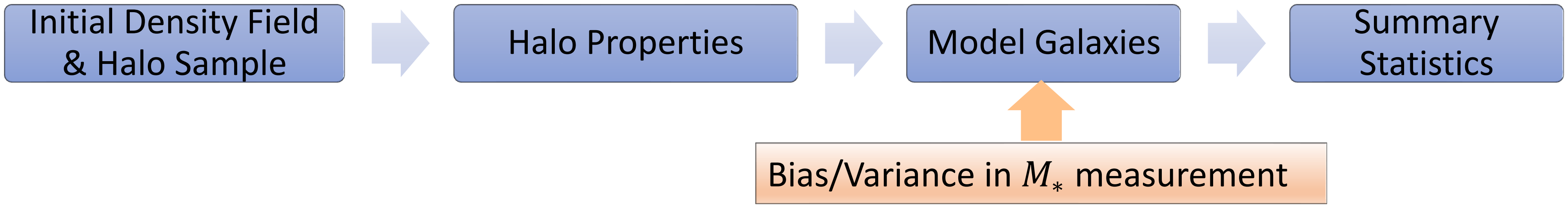
Method: add "backsplashed" mass to subhalo mass.

Result: M_* cosmic density enhancement at $z > 7$

- No effect at low mass.
- 100% at $M_* \sim 10^{10} M_\odot$.
- 1000% at the extreme masses.



Most massive JWST/CEERS galaxy at z ~ 8

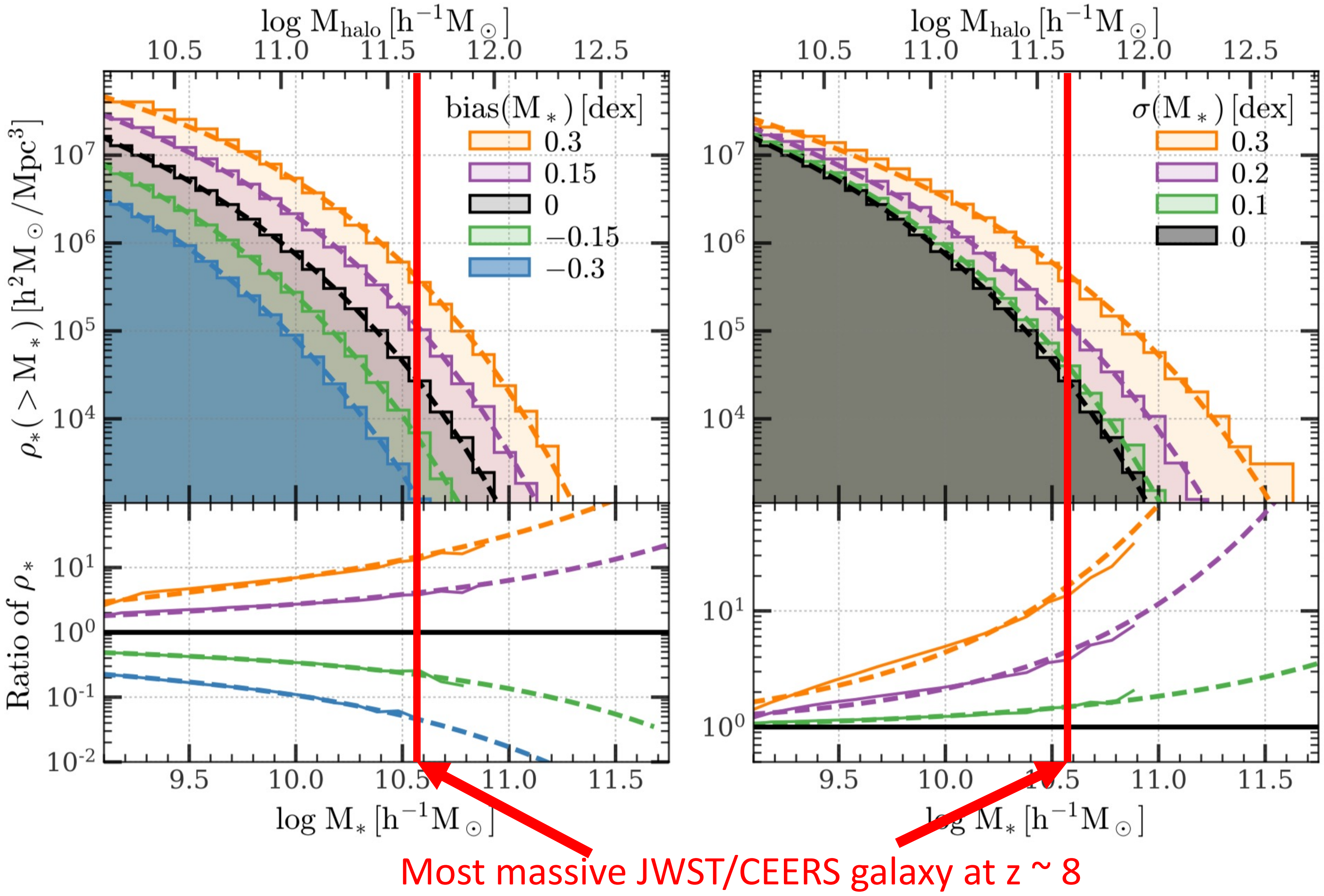


Method:

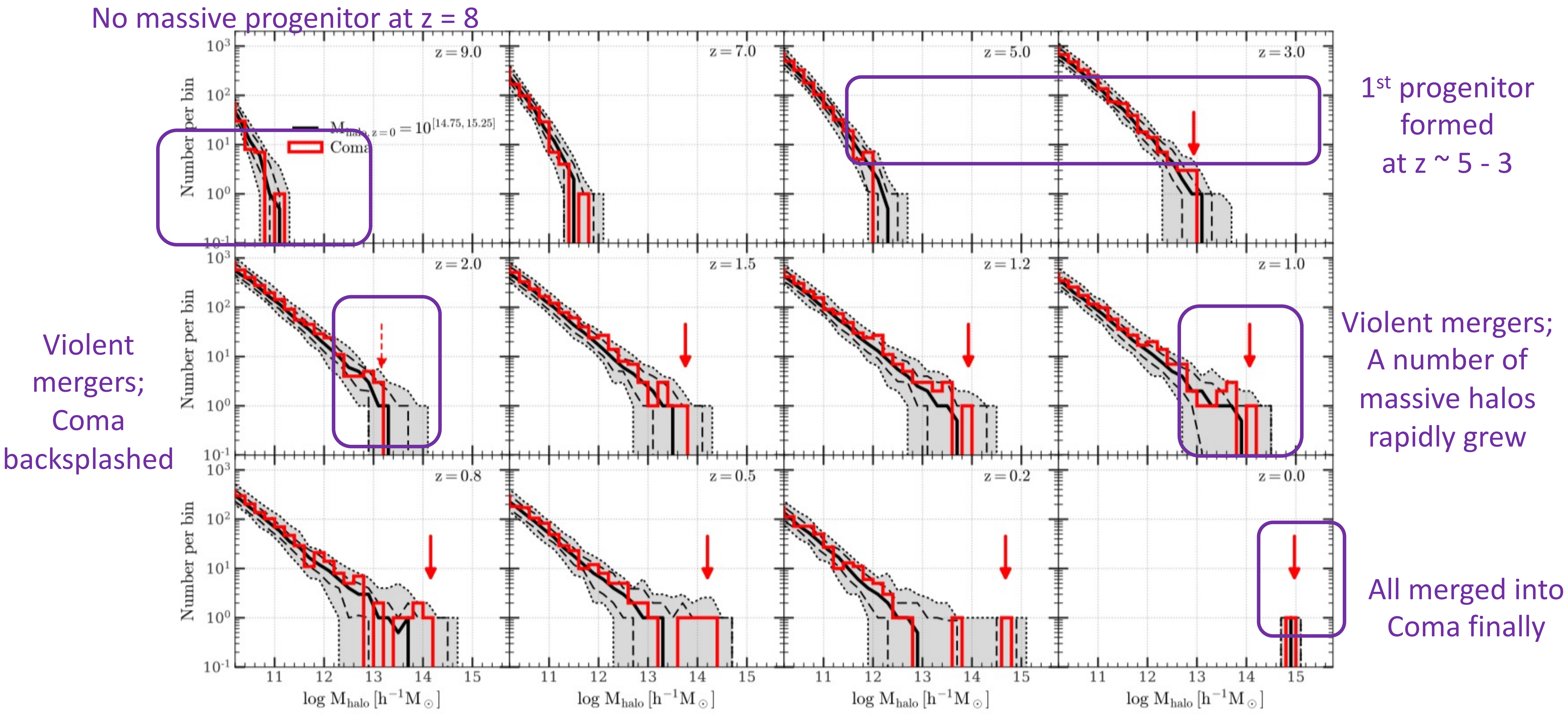
- Assume a constant star formation efficiency $\varepsilon_* = 50\%$.
- Add systematic and random errors into M_* to mimic observations.

Result: M_* changes at $z \sim 8$.

- Moderate at low mass.
- +0.3 dex bias/scatter in M_* produces 1000% enhancement on cosmic M_* density at extreme masses.

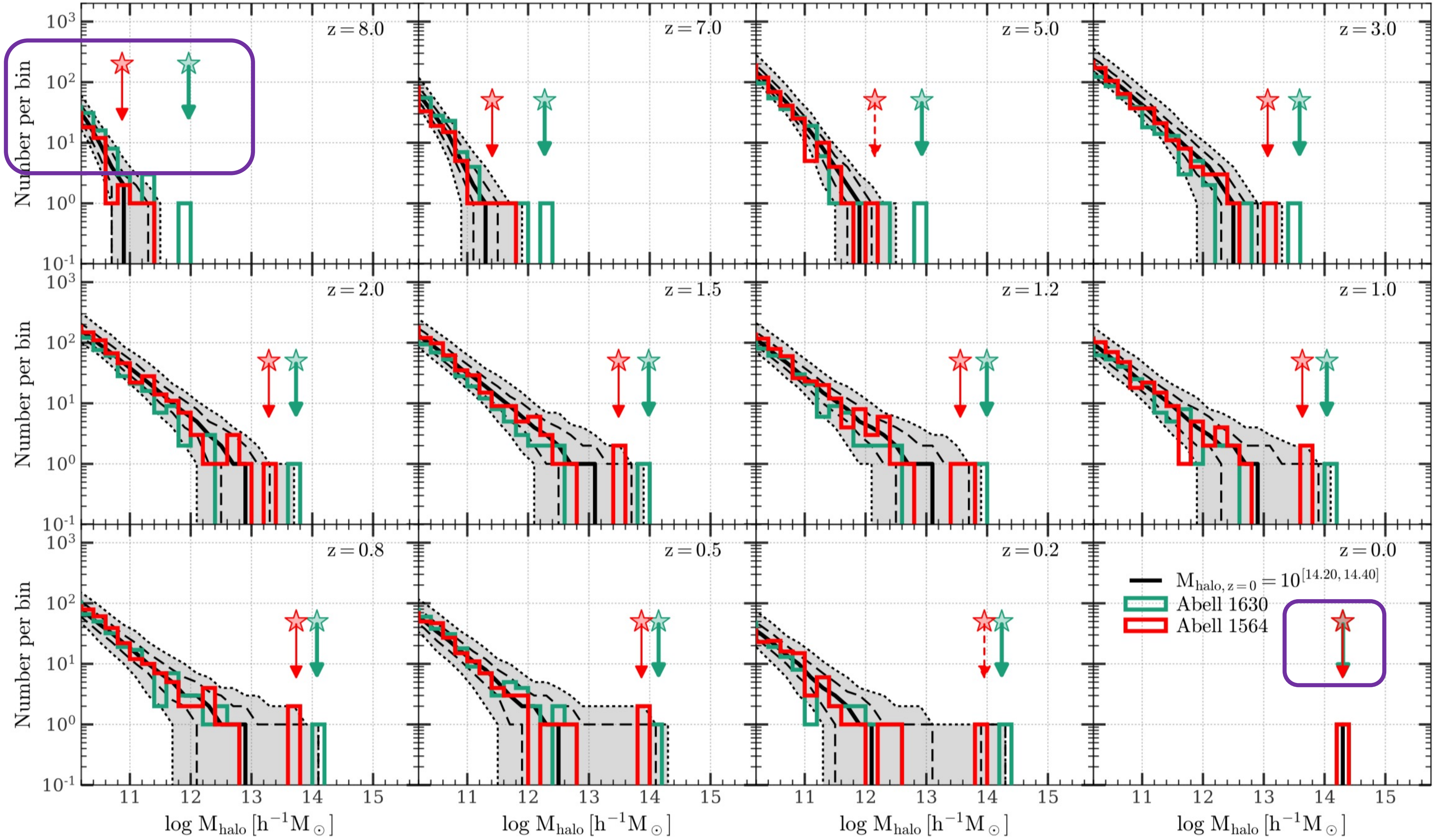


Assembly History of Massive Clusters: Coma



Assembly History of Massive Clusters: Abell 1630 and 1564

More massive than Coma at z = 8



less than 1/5 mass of Coma at z = 0