Application Case Study: HACC

ALCF: Simulation, Data, and Learning Workshop

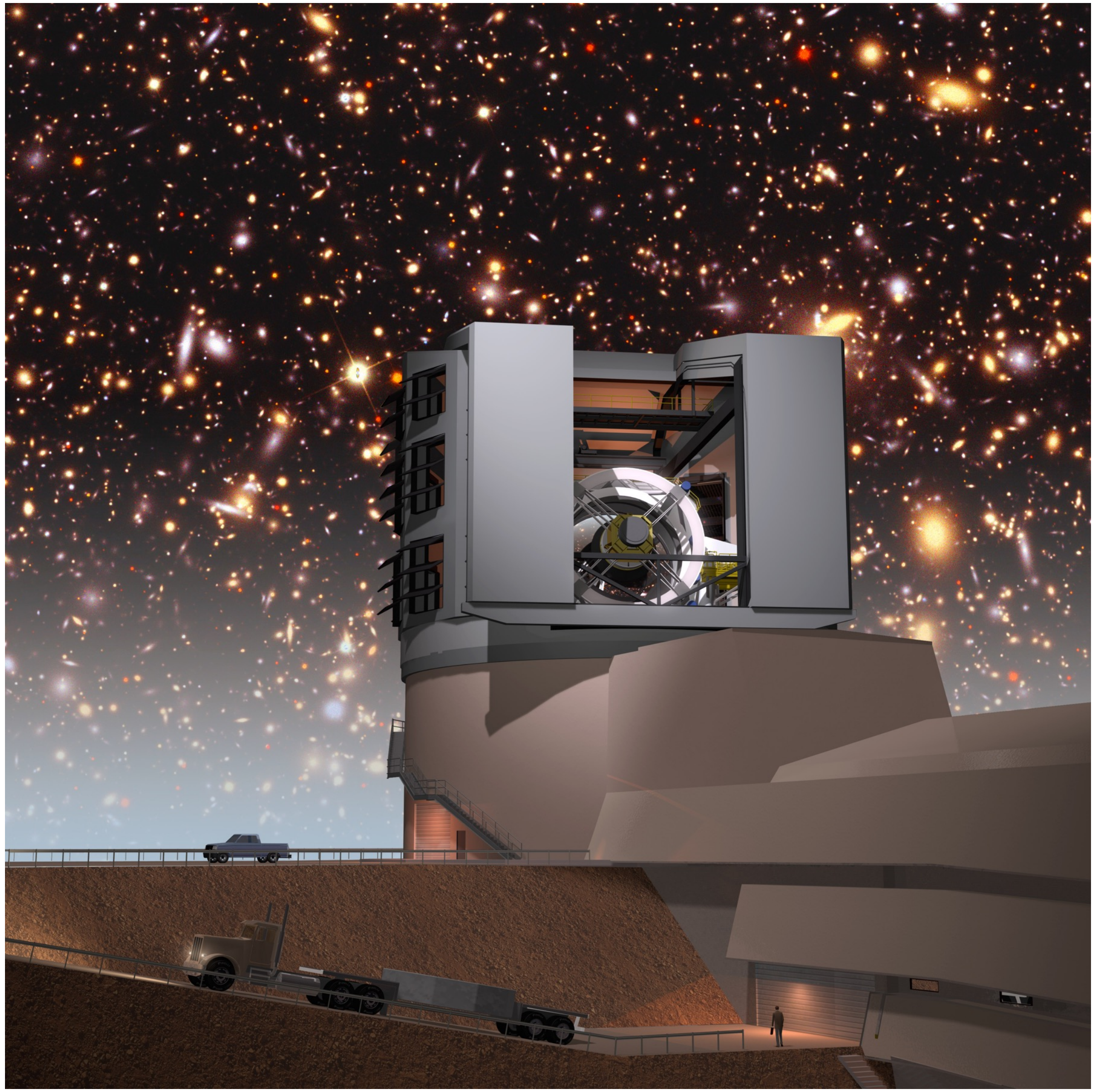
Steve Rangel, CPS
October 4, 2022

Overview

- HACC background
  - —Gravity-only, Hydro
  - —Data products and analysis
- Results from tuning subgrid model parameters (work done on Polaris)

Cosmological N-Body Simulations

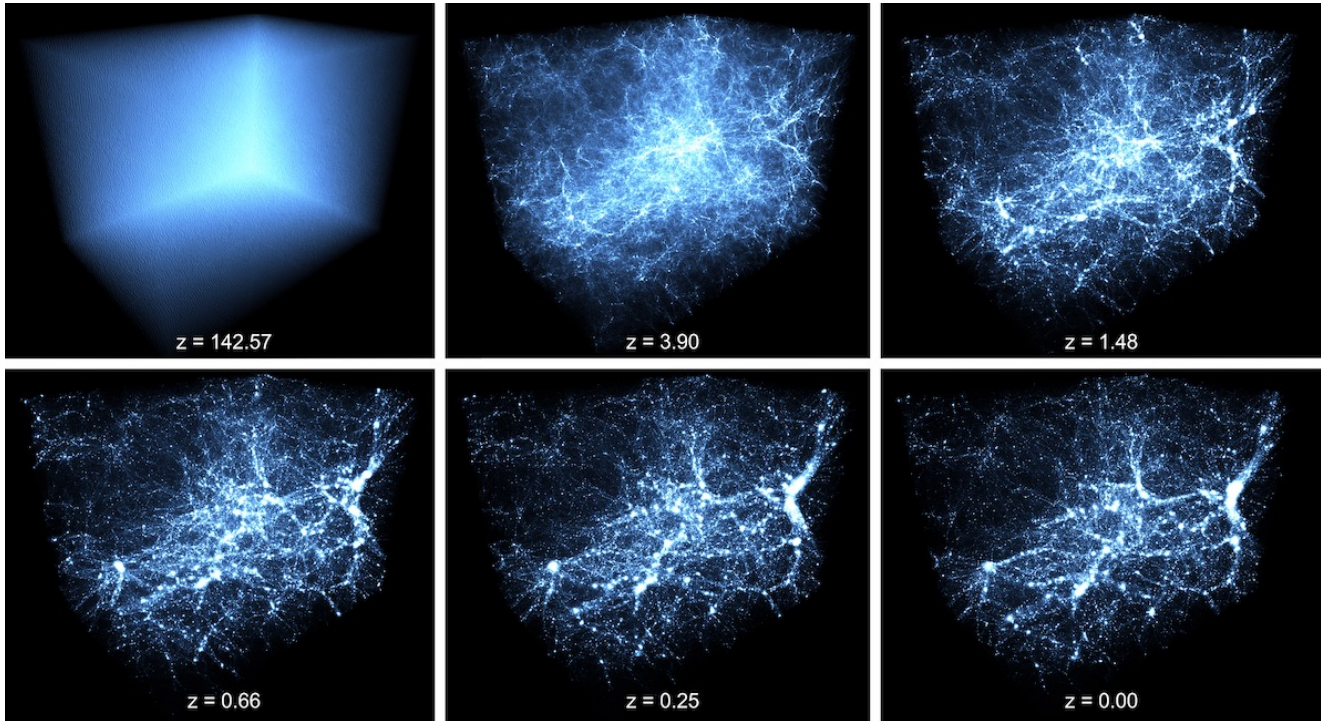
- Why do we run HACC (Hardware/Hybrid Accelerated Cosmology Code)?
- Theoretical research
  - Understand how Large-Scale Structures (LSS) form and evolve over cosmic time
  - Look for signatures of new/interesting physics
- Comparison with observations
  - Grand astronomical surveys
    - Rubin-LSST >$0.5B (NSF + DOE + …)
  - Create mock Universes for survey design
  - Provide theoretical models of summary statistics for data analysis (eg. emulators)
  - Understand data covariance for parameter estimation
  - Single observed Universe means forward-modeling



Rubin LSST: https://www.lsst.org/

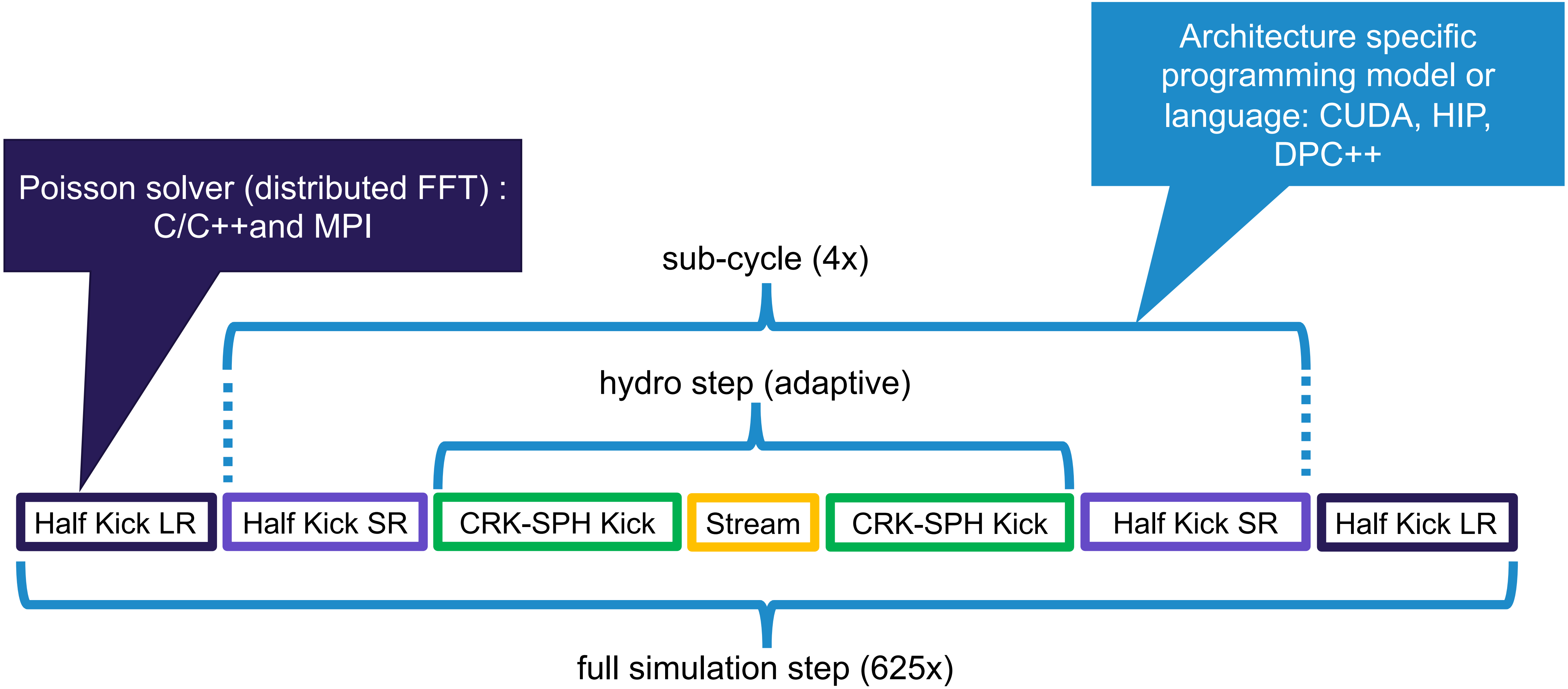
HACC N-Body: Matter Distribution

- Gravity
  - Infinite extent, unshielded, dominates on largest scales
  - As the Universe expands, structure condenses from very smooth initial conditions
  - Dark matter is dominant mass component and is modeled as effectively collision-less
- Hydrodynamics
  - Physics for (sub-dominant) baryonic matter component
  - Adiabatic: gas, temperature/pressure, shocks
  - Sub-grid: star formation, feedback from supernovae and Active Galactic Nuclei



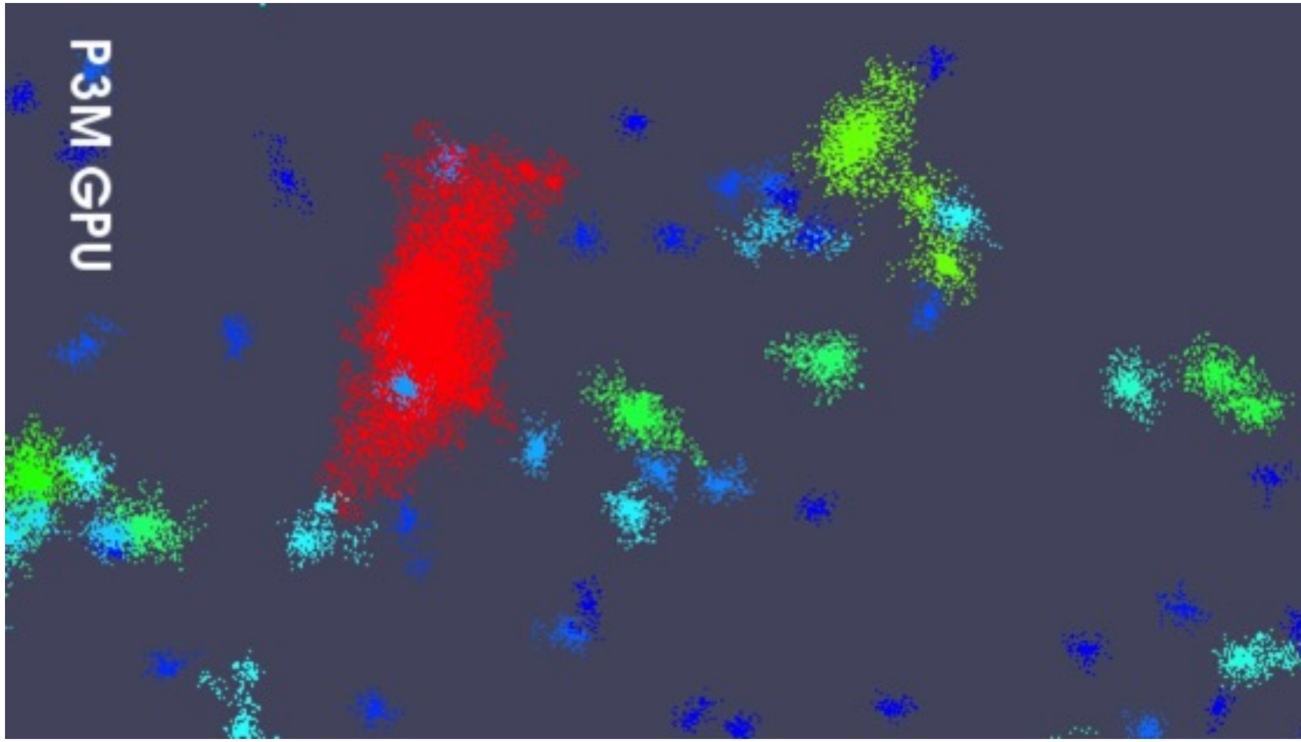
Evolution of matter distribution over cosmic time for a sub-volume of a HACC simulation. FarPoint: https://arxiv.org/abs/2109.01956

HACC: Execution Overview



HACC Analysis: Halos

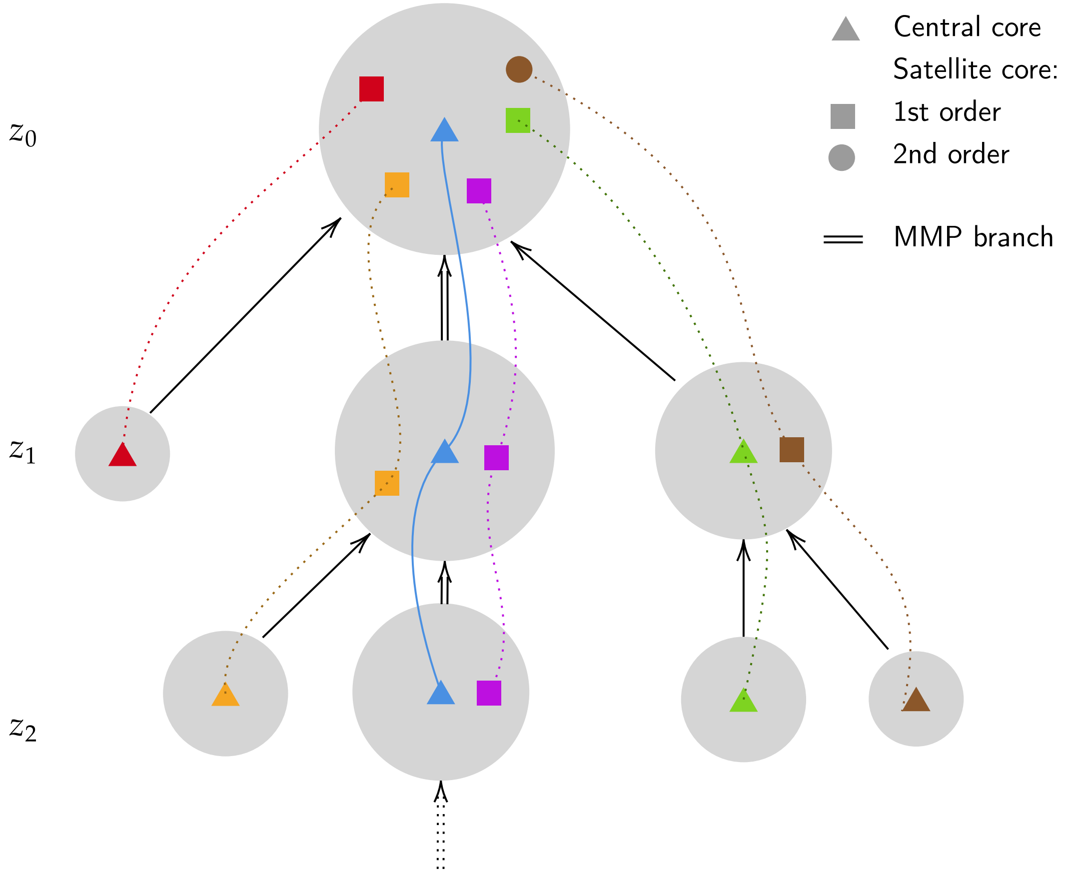
- Dark matter collects into “halos”
- Halos provide deep gravitational potential wells where baryonic matter can collect and eventually cool and condense to form stars and galaxies
- Roughly half of the mass in the Universe ends up in halos by our current epoch
- Halos are identified in simulations by looking for coherent structures with densities >100x of the background density



Particles in a small volume of a HACC simulation colored by halo membership.
HACC: https://arxiv.org/abs/1410.2805

HACC Analysis: Halo Merger Trees

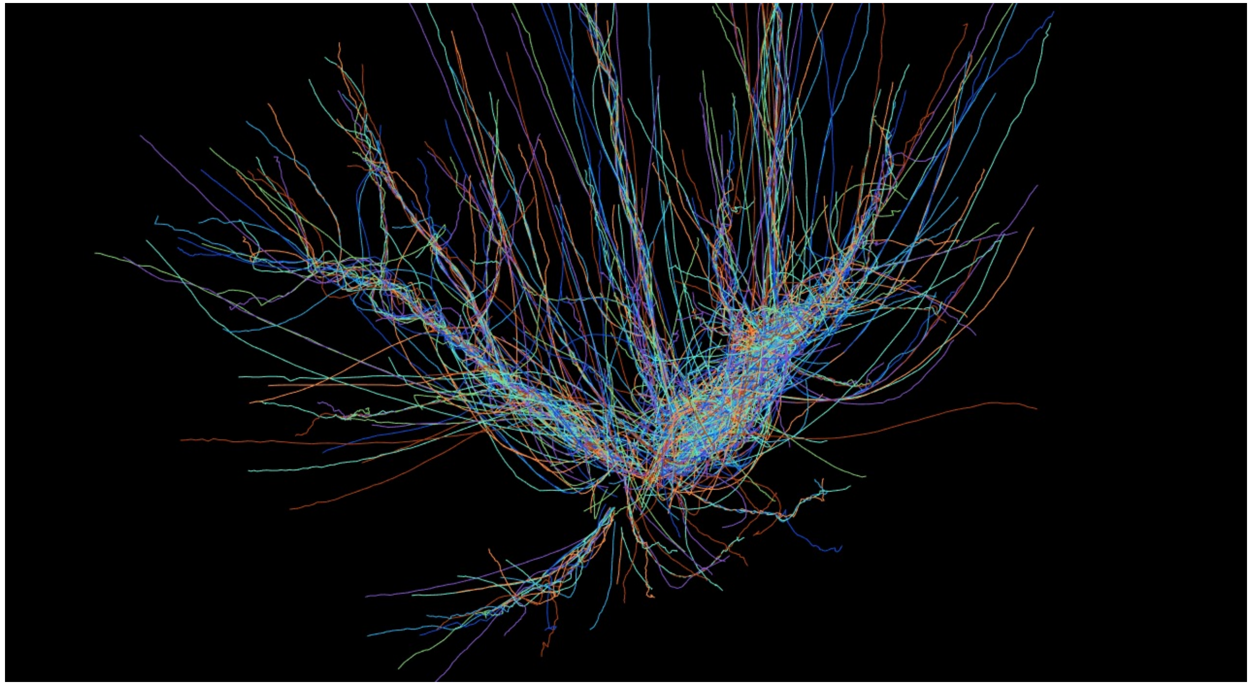
- Halos interact with each other as the Universe evolves, colliding and merging
- The interaction history of halos is important because interactions between galaxies within halos can trigger epochs of star formation, and the total history of star formation in a galaxy determines its luminosity/color
- The interaction histories of halos is summarized in a data structure called a merger tree



Logical merger tree.
SMACC: https://arxiv.org/abs/2012.09262

HACC Analysis: Halo Core Tracking

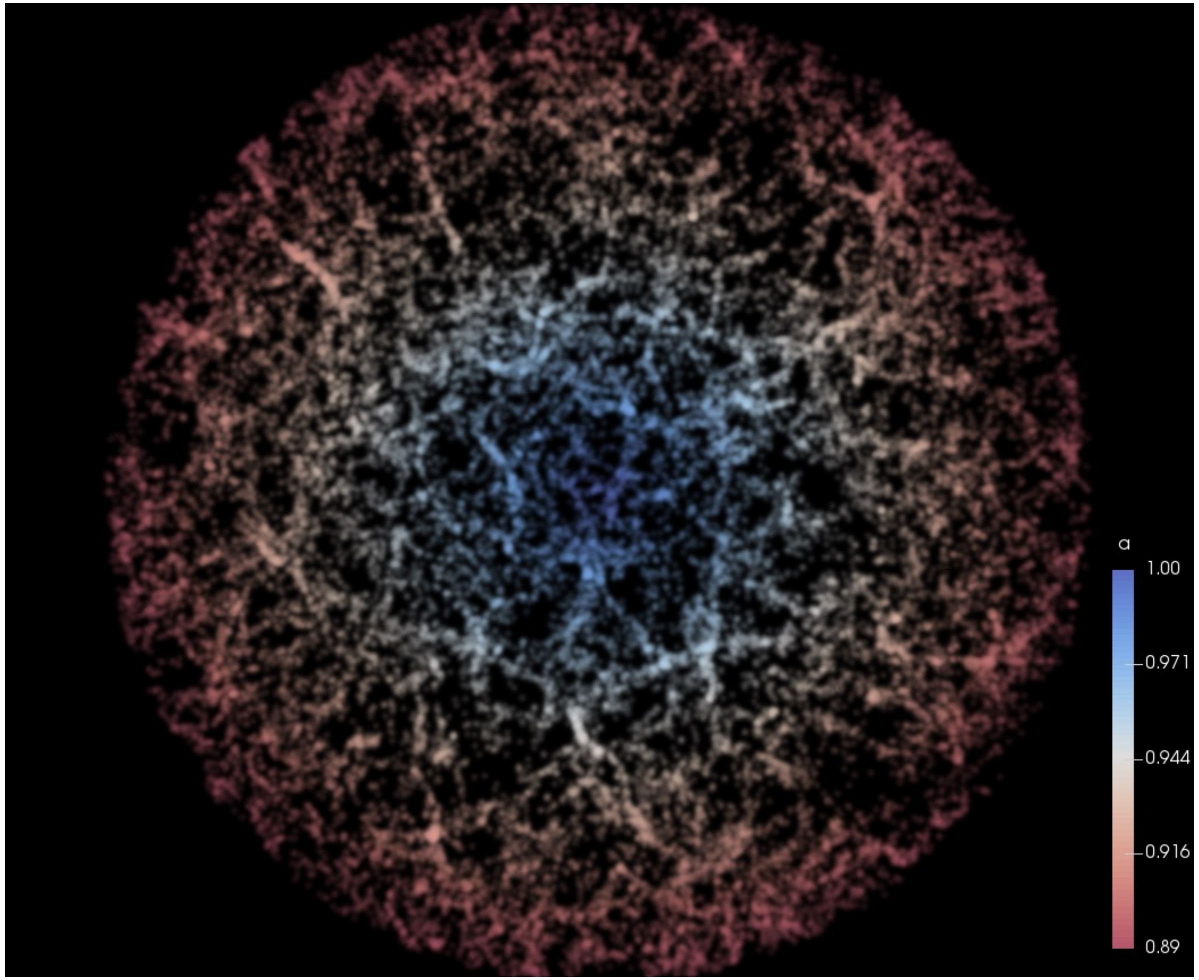
- Very inner part of halo is a tightly-bound core of particles that is not easily disrupted during halo-halo interactions
- Track sub-structure within halos by continuing to track cores even after halos merge
- Core positions are likely good proxies for galaxy locations



Physical trajectories of cores that merge into 1 halo.
OuterRim: https://arxiv.org/abs/1904.11970

HACC Analysis: Lightcones

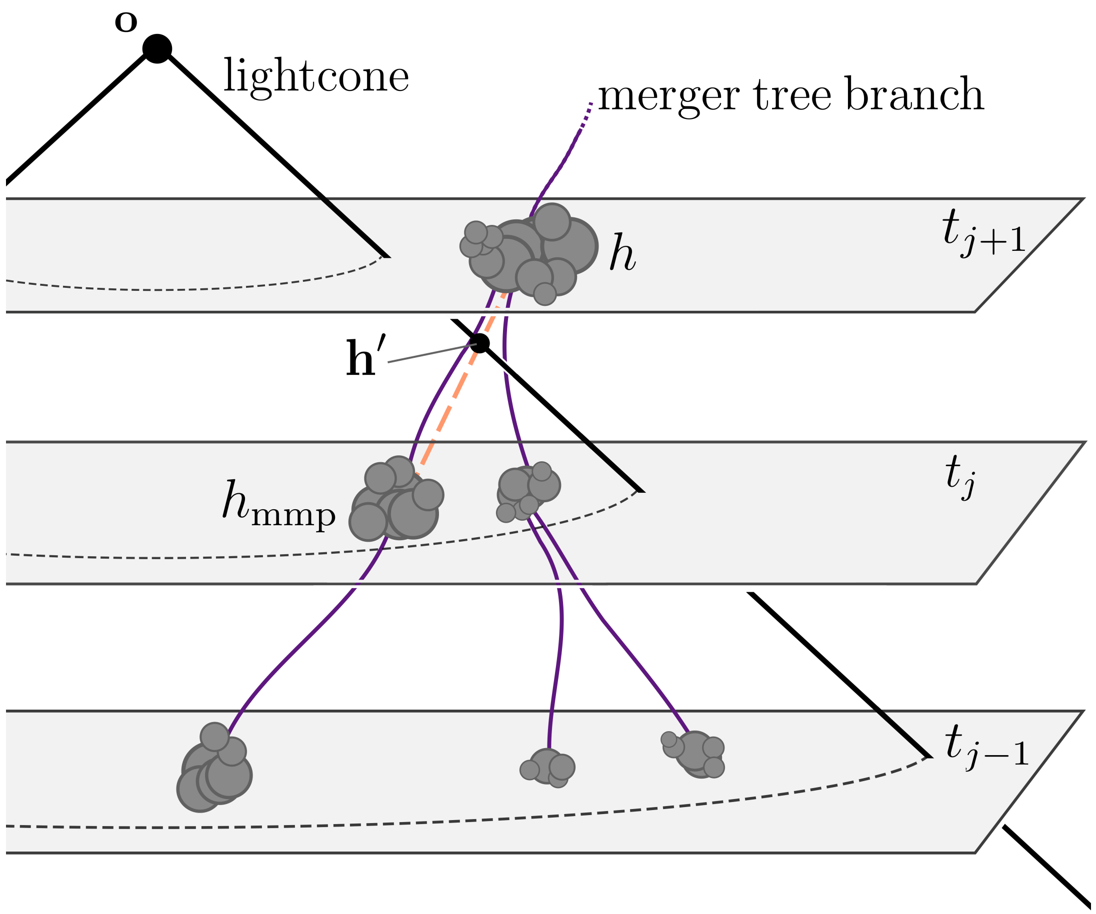
- N-body simulations operate in a comoving gauge, observations are not in same gauge
- Finite speed of light, we observe objects as they were when the light that we now collecting left the object
- Objects that are farther away have a longer lookback-time
- HACC runs in a fixed-sized box (in comoving/expanding units) with periodic boundary conditions, but we can create a lightcone around an observer by saving the correct spherical shell from each time step



Particle lightcone from a HACC simulation with observer at center; color indicates distance/lookback-time. OuterRim: https://arxiv.org/abs/1904.11970

HACC Analysis: Halo Lightcones

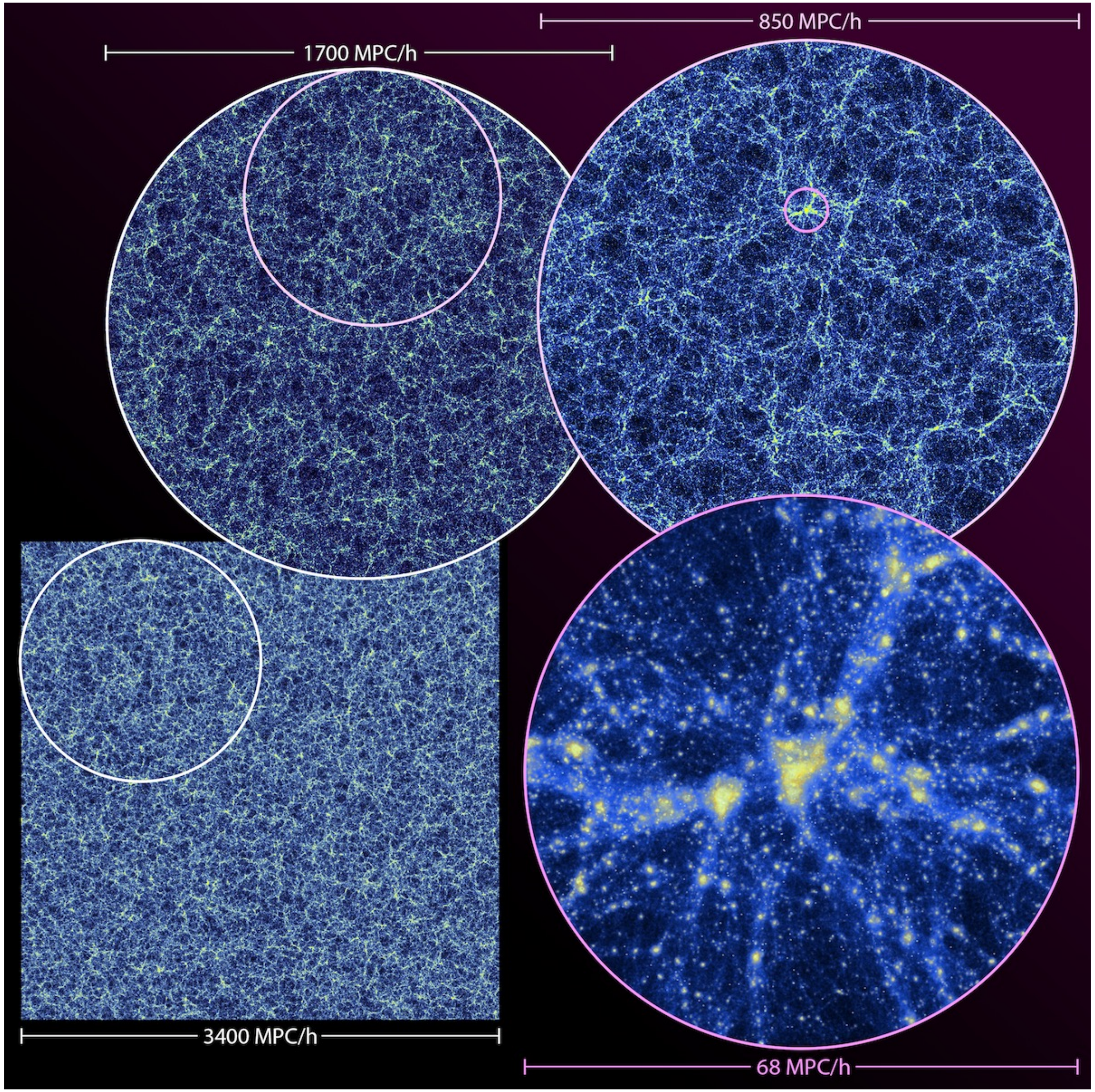
- Construct halo merger trees to the end of the simulation in the entire simulation volume
- Can go back and figure out where a merger tree intersects an observer's lightcone in order to display information from the merger tree in the right place at the right time



Interpolating merger trees onto an observer's lightcone.
CosmoDC2: https://arxiv.org/abs/1907.06530

HACC: Limitations

- HACC does not directly produce galaxies
  - —Position, size/shape, luminosity/color
- Dynamic Range
  - —Need very large volumes (~Gpc^3) for reliably realistic structure formation
  - —Galaxies are ~10 kpc in size, memory limited in number of tracer particles per galaxy
- Computational Intensity
  - —Total workload increases dramatically as physical resolution increases
- Self-consistent creation of realistic galaxies within a cosmological-volume n-body simulation is a beyond-exascale problem



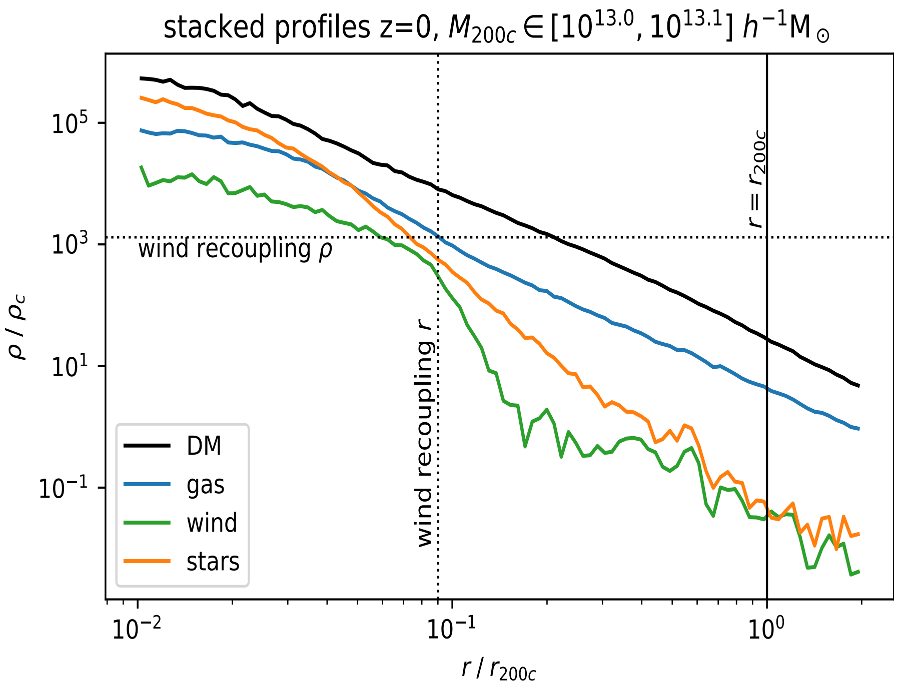
LastJourney: https://arxiv.org/abs/2006.01697

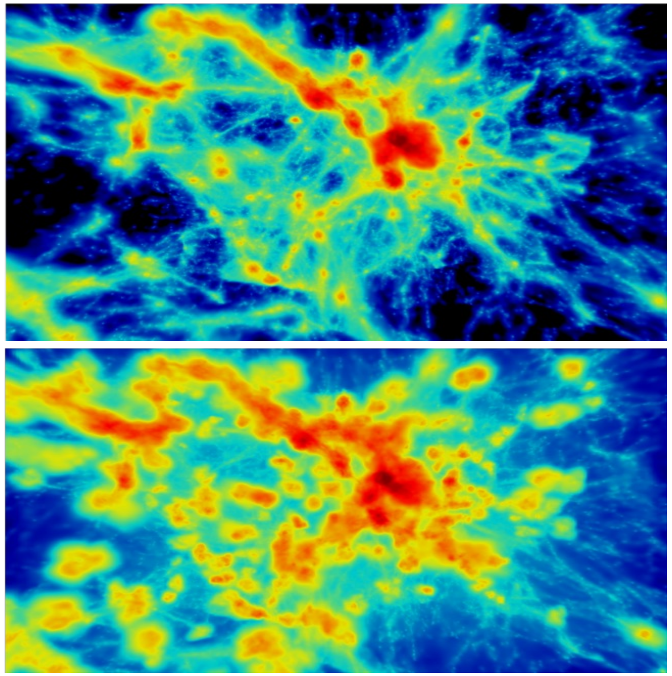
Sub-resolution modeling for modern cosmological survey predictions

- **Goal**
  - — Develop and calibrate models that emulate astrophysical processes at scales smaller than those resolved by simulations.
  - — Utilize results from high-resolution galaxy formation simulations to parameterize models suited for cosmological predictions. Collaborative project with Prof. Claude-Andre Faucher-Giguere's group at Northwestern.
  - — Implement the new models within the CRK-HACC cosmology code
  - — Final goal is to fully calibrate and validate the currently implemented models to our relevant scientific targets.
- **Scope of Work**
  - — Subgrid modeling of galaxy formation and feedback
  - — Sensitivity analysis of model parameters
  - — Calibration and validation to observations
  - — Performant models on accelerated hardware
  - — New integration techniques that are scalable for cosmological volume simulations

Suite of subgrid models

- UV/Background using latest model by Faucher-Giguere 2020
- Consistent metal-line cooling treatment (Wiersma)
- Chemical enrichment rates taken from FIRE simulation measurements
- AGN kinetic and momentum feedback (utilizing a GPU FOF finder).
- Supernova Feedback and Galactic Winds, informed by outflow rates measured in FIRE.

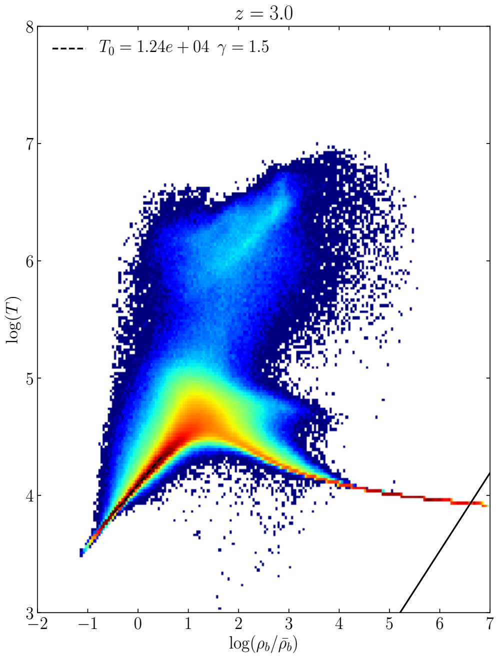




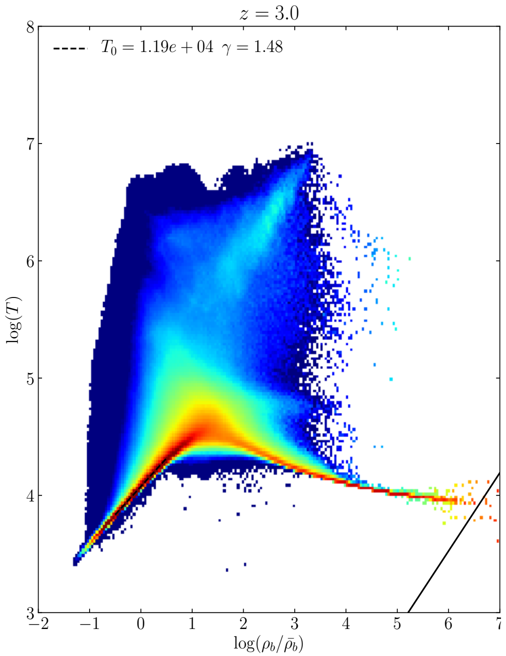
Argonne NATIONAL LABORATORY

Simulation Validation and Comparisons

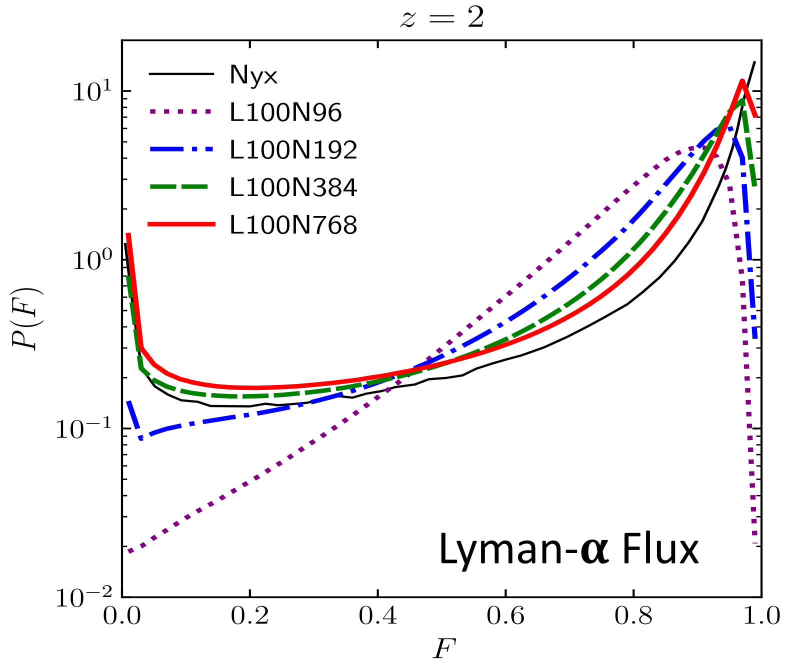
- Validation measurements using many different solvers
  - Nyx, Gadget, AREPO, FIRE
- Lyman-α science runs
- Subgrid modeling comparisons

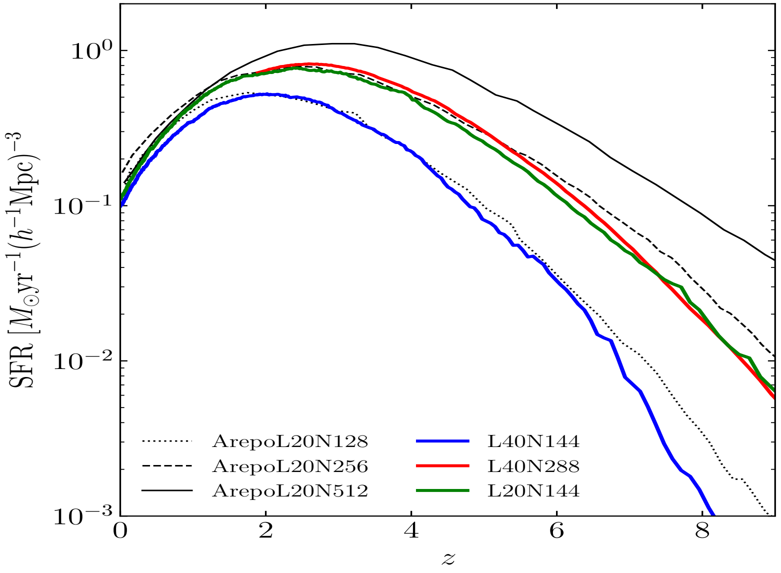


HACC



Nyx



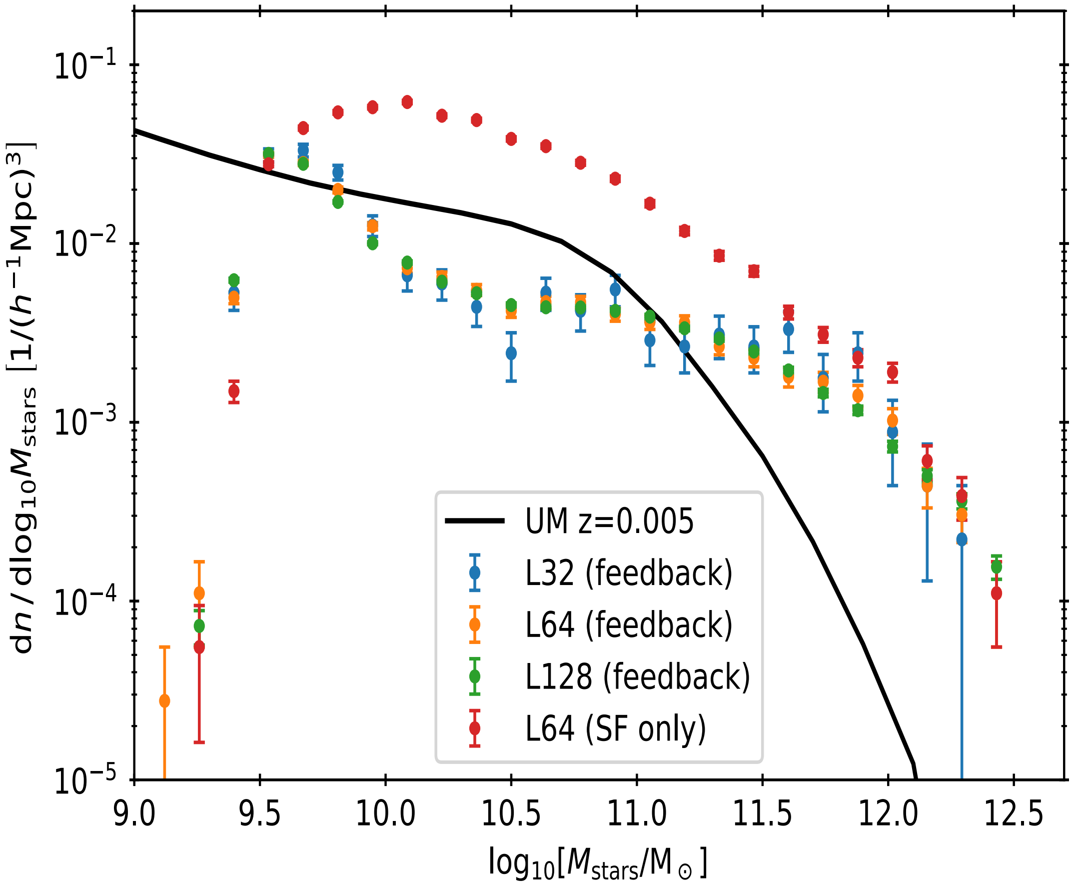


Star Formation Convergence

Argonne NATIONAL LABORATORY

Calibrated Simulations

- Improved subgrid models that are performant on exascale hardware
  - ➢ Radiative and metal-line cooling, star formation and chemical enrichment, active galactic nuclei feedback
- First results from analysis pipelines measurements
  - ➢ GSMF, stellar-halo relation, star formation density history
- Next steps
  - ➢ Complete a suite of simulations to calibrate models to observations
  - ➢ Measure outflow rates and local effects for A.I. models

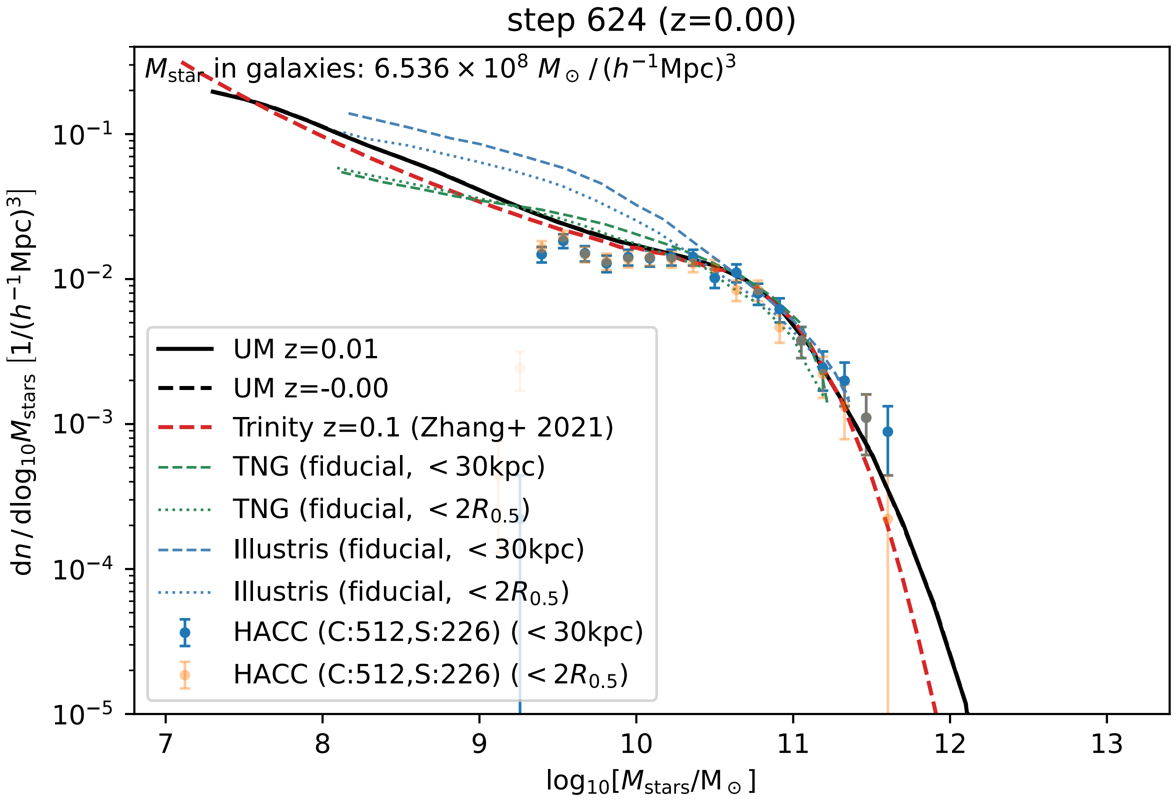


Calibration runs for the cosmic star formation density and galaxy stellar mass function

Finding Suitable Subgrid Parameters

Series of 32 Mpc h^{-1} simulations, run on Polaris 2-node jobs (3-4 hr. wall time)

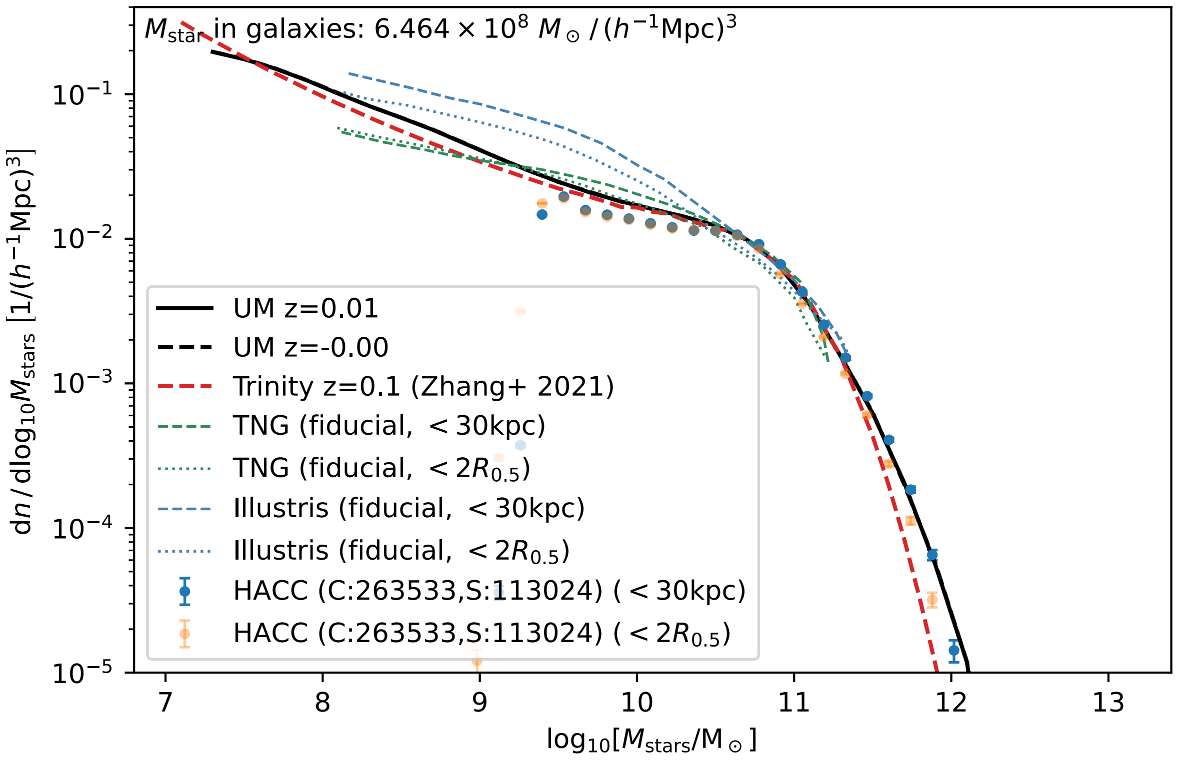
STARAGN/KAPPA_3.7_EGW_0.5_FOF_1e11_SEED_1.25e6_NPERH_AGN_2_GALSEED_SOD_RERUN



Argonne NATIONAL LABORATORY

Results for Tuned Subgrid Models

256 Mpc h^{-1} simulation, run on Polaris 128-256 nodes



Argonne NATIONAL LABORATORY

Thank you!

erangel@anl.gov