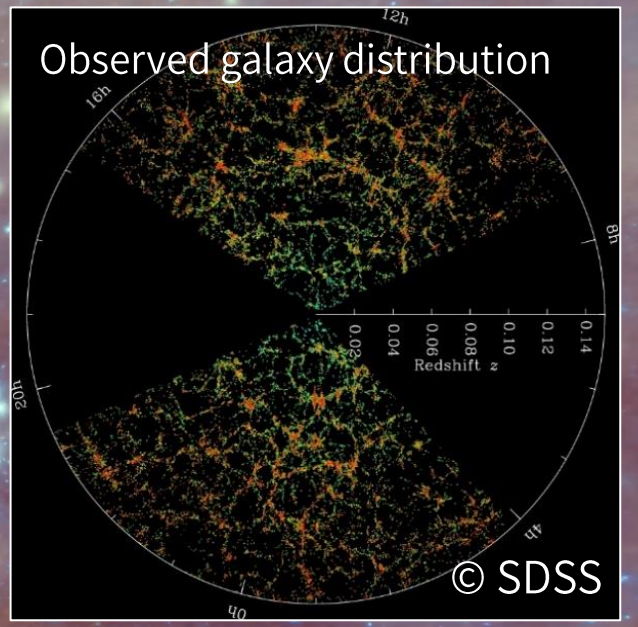
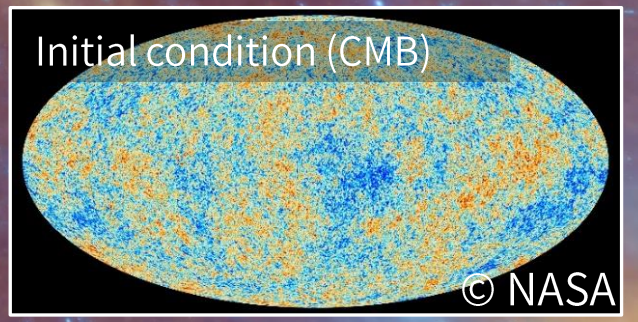
Investigating dark matter structure by cosmological simulations on the supercomputer Fugaku and machine learning

Tomoaki Ishiyama (Chiba University)

Publicly offered research 23H04002

LCDM cosmological model





4D2U (NAOJ) movie
「Formation and Evolution of Dark Matter Halos (III. Formation of Cosmic Web and Void Structures)」

https://www.youtube.com/watch?v=R7nV6JEMGAo

Halo structure

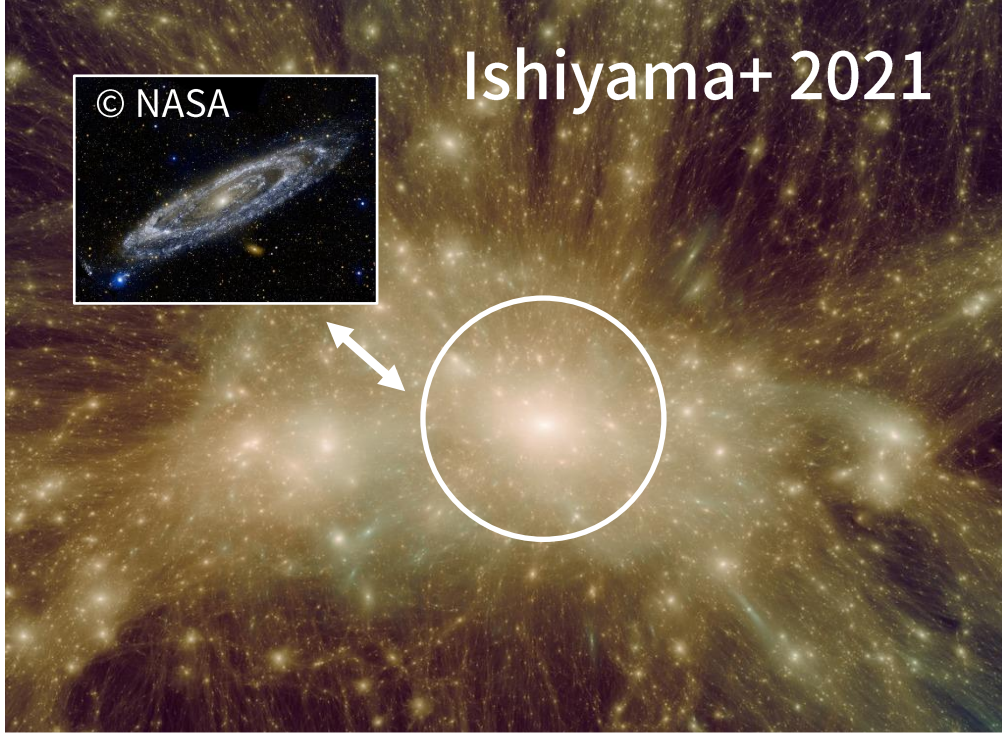
- **Central cusp**
  - Einasto profile or NFW profile

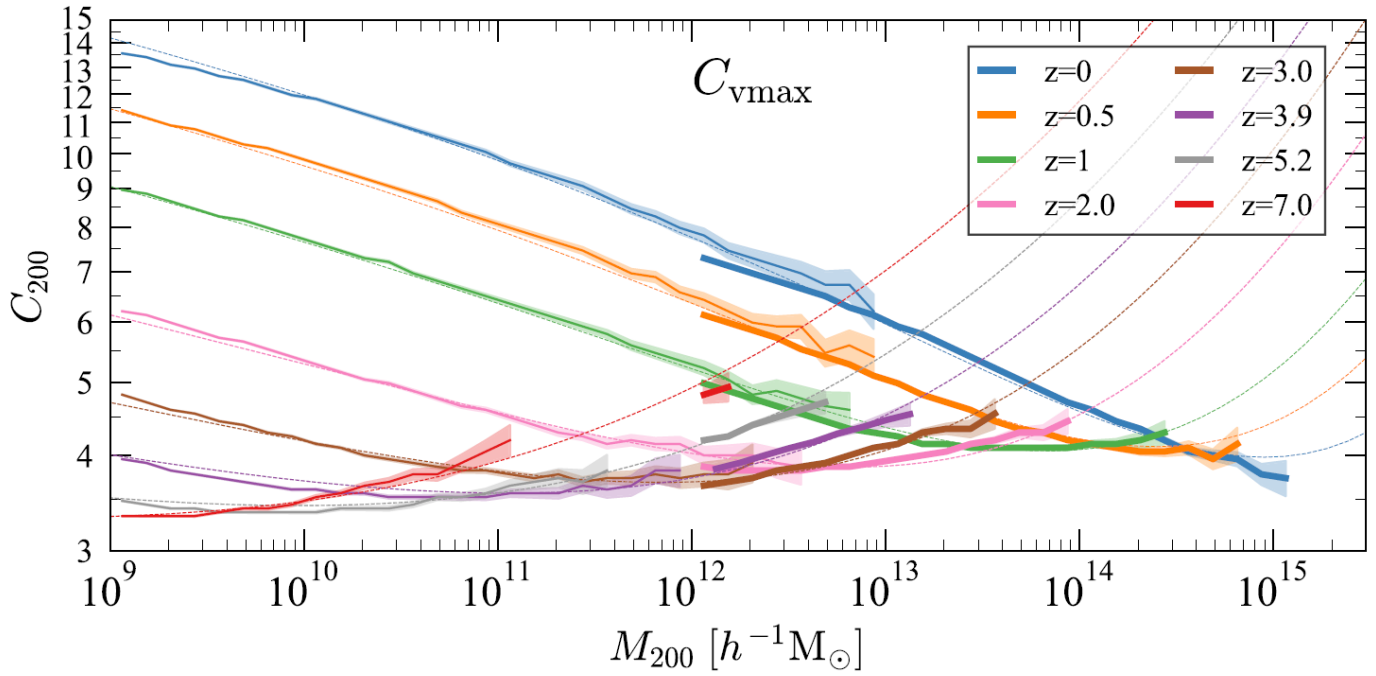
$$\rho(r) = \frac{\rho_s}{(r/r_s)[1+(r/r_s)]^2}$$

Concentration
$c = r_{vir}/r_s$, r_{200}/r_s

- **Numerous subhalo**
  - $dn/dm \sim m^{-(1.8\sim2)}$
- **Triaxial**
- **Non Universality**
  - Weak dependence on the halo mass
  - halo to halo variation

Halo structures have large impact on galaxy formation physics and dark matter detection





Predicting galaxy cluster's mass and concentration using convolutional neural network

- Galaxy clusters are the most massive virialized objects
- The cluster abundance is sensitive to cosmological parameters
- The cluster mass-concentration relation can also be used as a cosmological probe
- Those quantities are not directly observable
- This work: Predicting those quantities using mock galaxy catalogs (Uchuu-UniverseMachine; Aung, TI+ 2023) and convolutional neural network

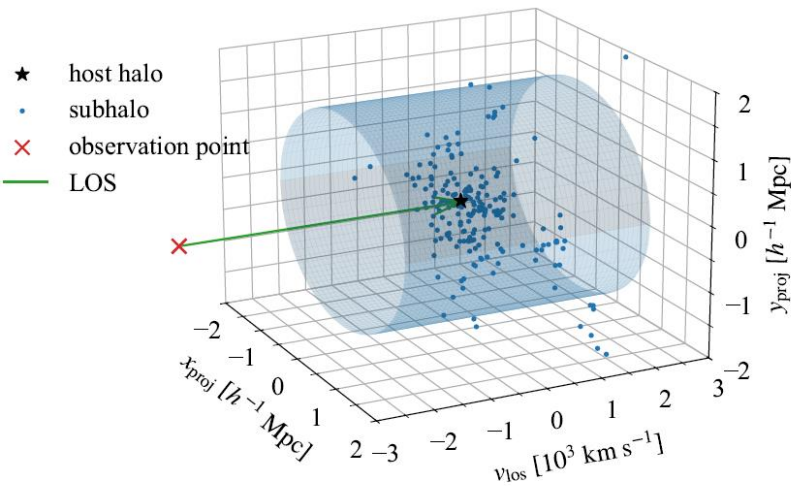
$$\rho(r) = \frac{\rho_s}{(r/r_s)[1+(r/r_s)]^2}$$

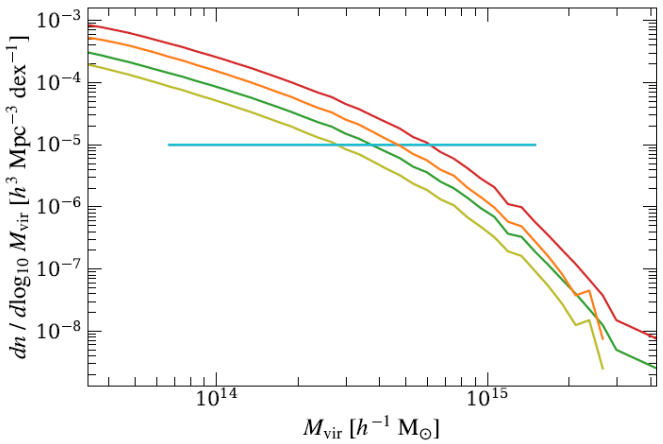
Concentration
$c = r_{vir}/r_s$, r_{200}/r_s

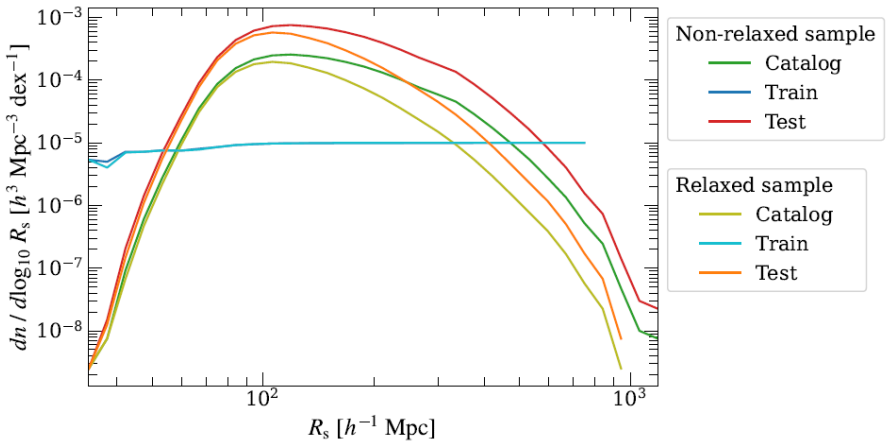
Tominaga, Nakamura, TI, Abdullah, in prep

Data and method

- Galaxy clusters at z=0.09 and member galaxies in Uchuu-UniverseMachine mock catalogs (Aung, TI+ 2023)
- Input for CNN: 2D histogram of member galaxies' observable quantities, observed by observers at z=0.0
  - Projected distance from the center of cluster
  - Line of sight velocity
- Mimicking observations taking interloper galaxies into account
- Flattening the data distribution by down- and upsampling (multiple projections)



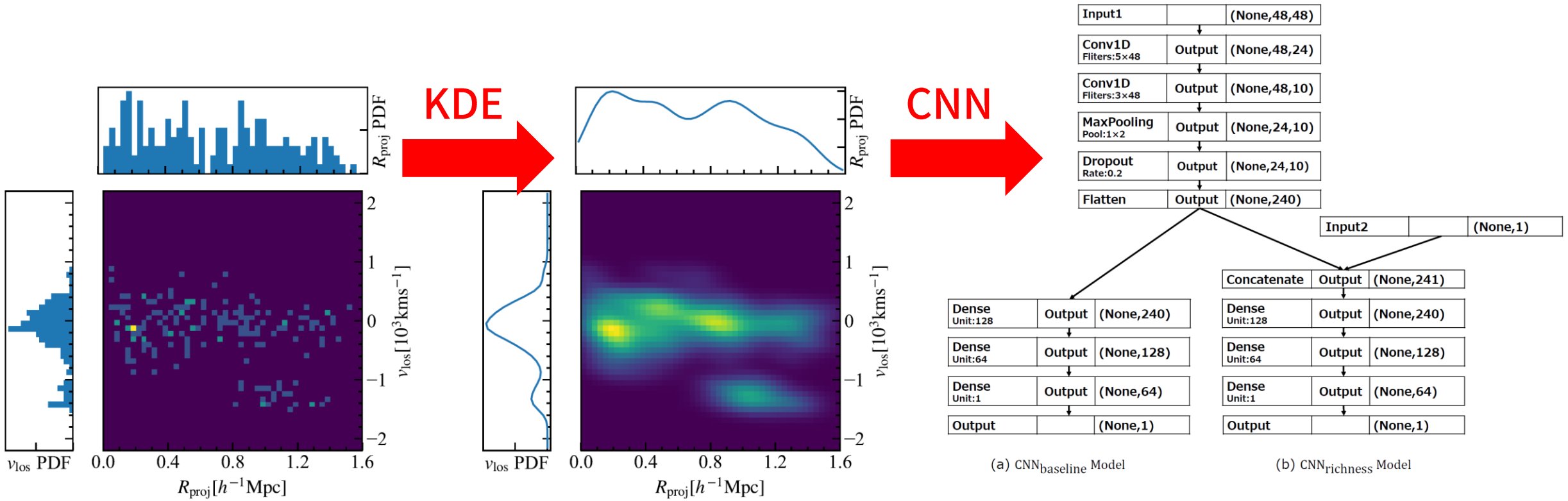




Tominaga, Nakamura, TI, Abdullah, in prep

CNN

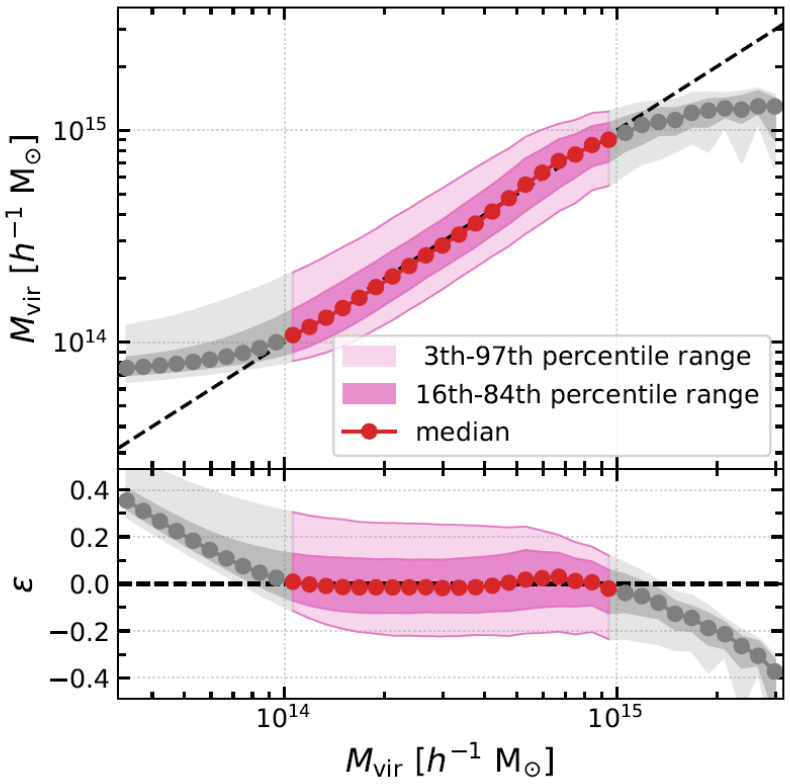
- Input: 2D histogram of member galaxies' observable quantities projections, taking interloper galaxies into account
- Smoothed by kernel density estimation (KDE)
- Richness can also be used as additional input parameter for a dense layer



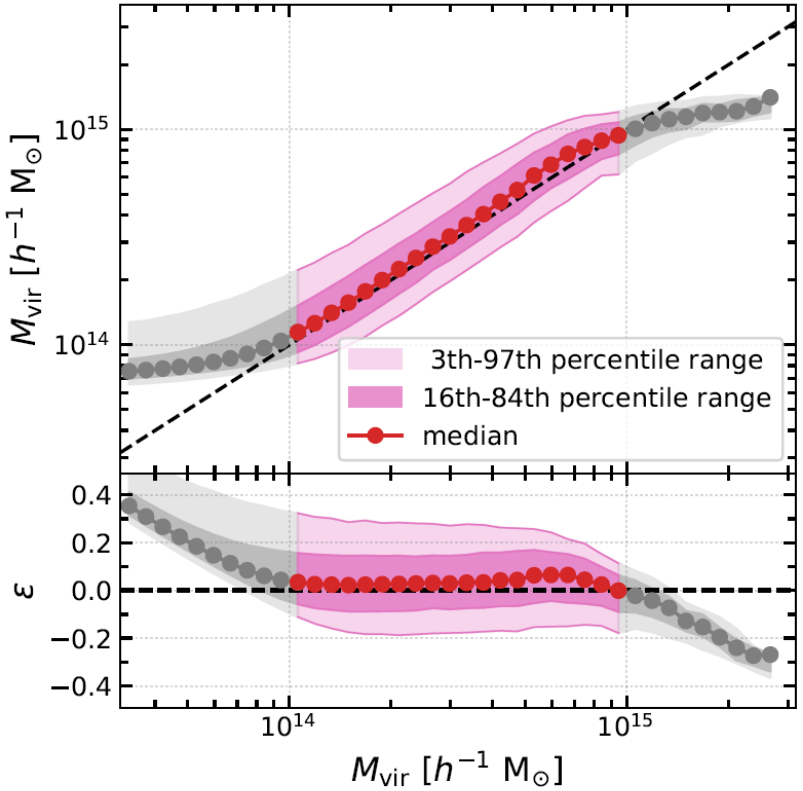
Halo mass

ε = log10 (Predict/true)

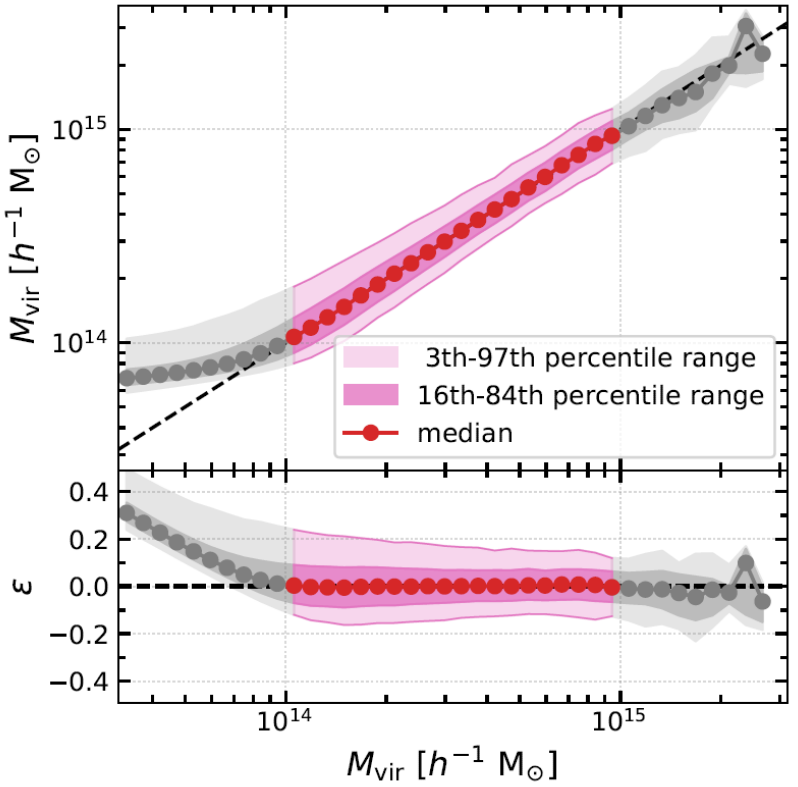
All clusters



relaxed



relaxed + richness

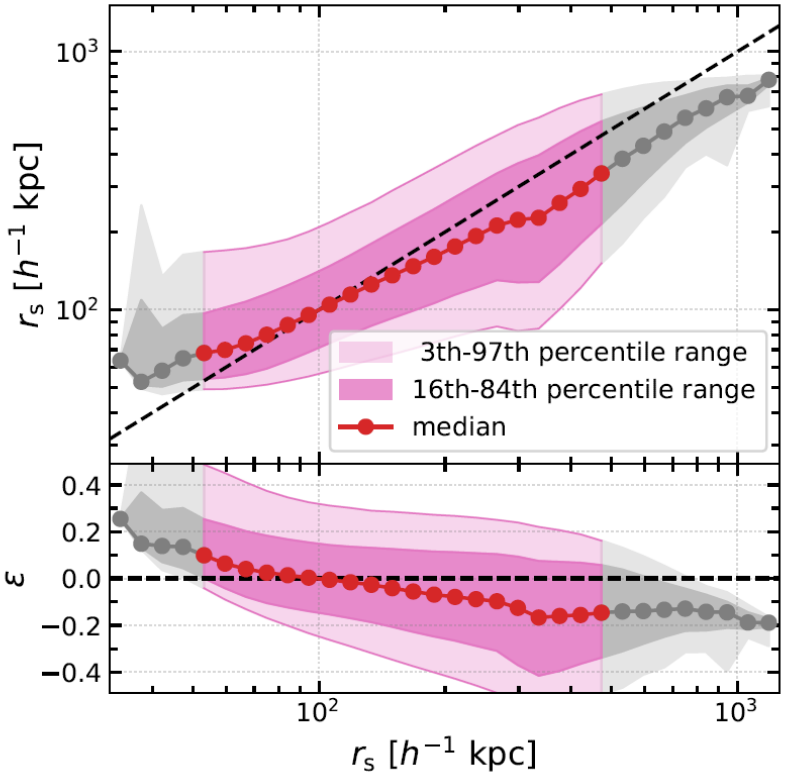


- The median error and 1sigma scatter are less than ~0.03 and ~0.15 dex, respectively
- Including richness can further decrease scatter down to ~0.1 dex
  - Larger effect than using only relaxed clusters

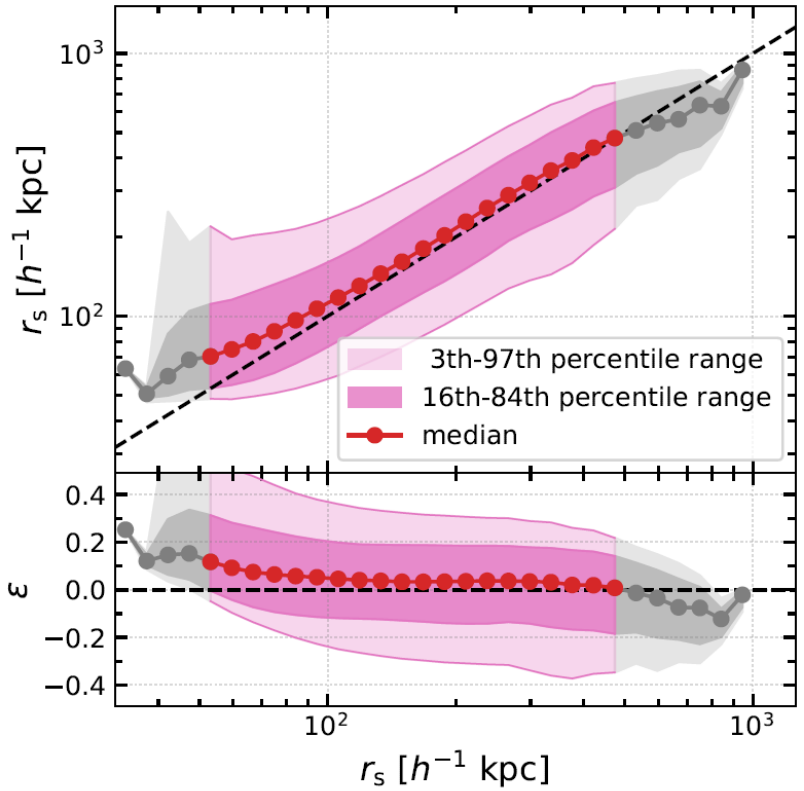
Scale radius (concentration)

ε = log10 (Predict/true)

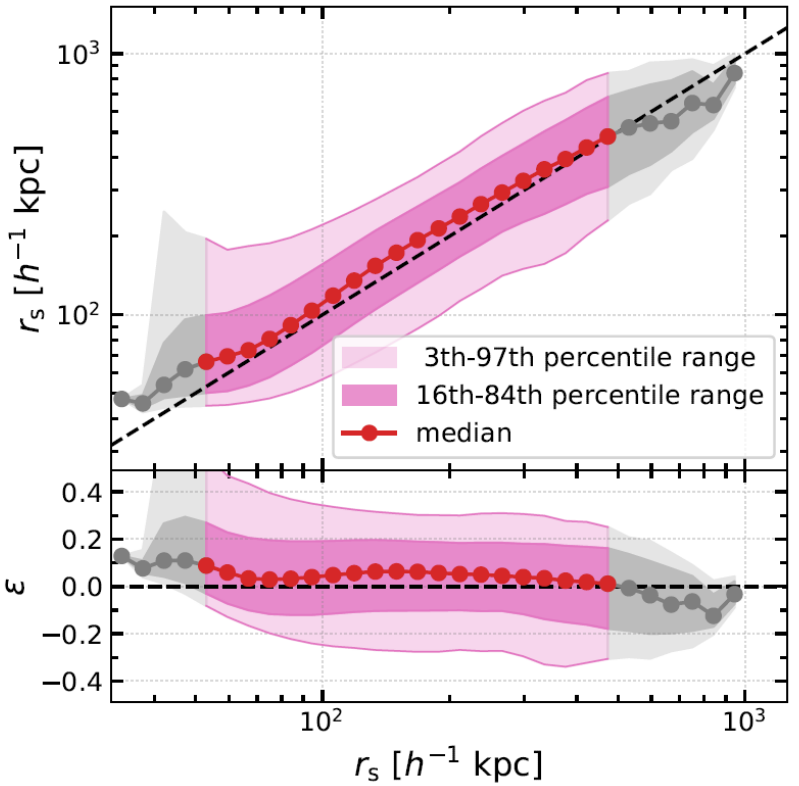
All clusters



relaxed



relaxed + richness



- The median error and 1sigma scatter are less than ~0.04 and ~0.17 dex, respectively
- Using only relaxed clusters can modifiy the overall slope to the true one
- Including richness doesn’t have much effect

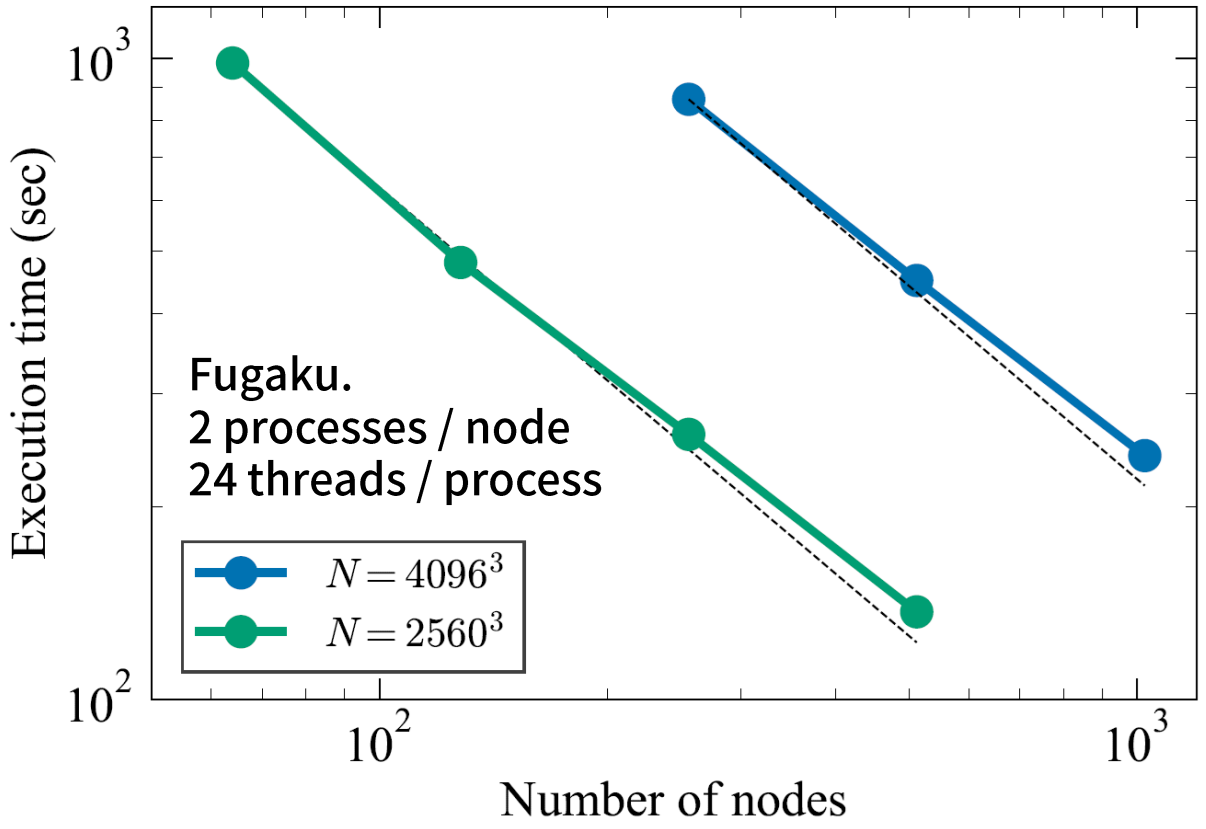
Other achivements (digest)

MPI-Rockstar: a Hybrid MPI and OpenMP Parallel Implementation of the Rockstar Halo finder

Tokuue, Ishiyama, Osato, Tanaka, Behroozi, arXiv: 2412.18629, submitted to JOSS

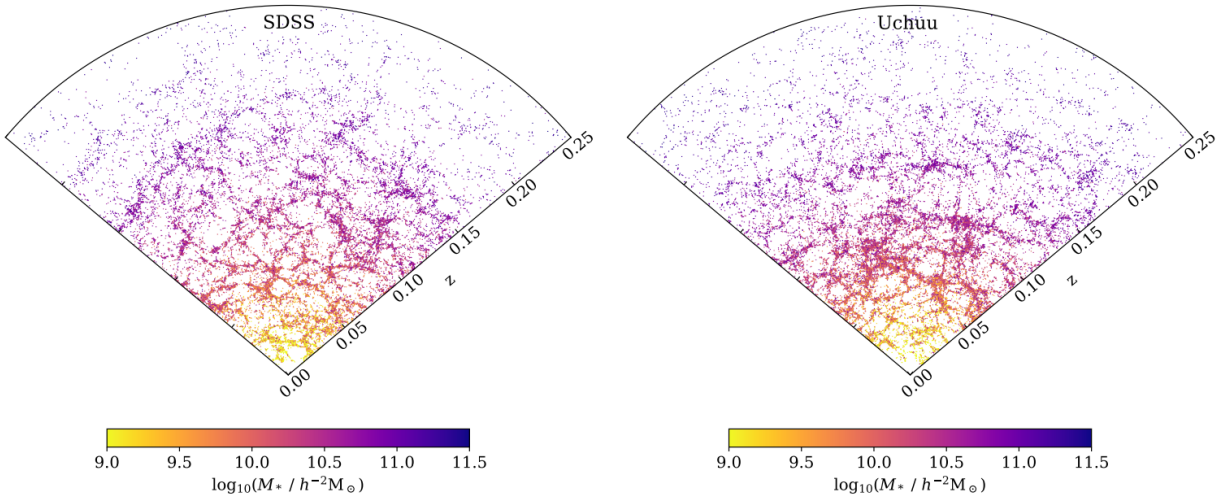
https://github.com/Tomoaki-Ishiyama/mpi-rockstar

- Fully refactoring Rockstar using MPI+OpenMP instead of the socket library
- Most communications (e.g., particle re-distribution) is replaced by fast collective (e.g., MPI_Alltoallv) communication
- Halo/subhalo finding is done by MPI+OpenMP hybrid way
- Additional halo properties: Inertial tensor, intermediate axis, chi square of the NFW fitting
- N>10^{12} can be analyzed



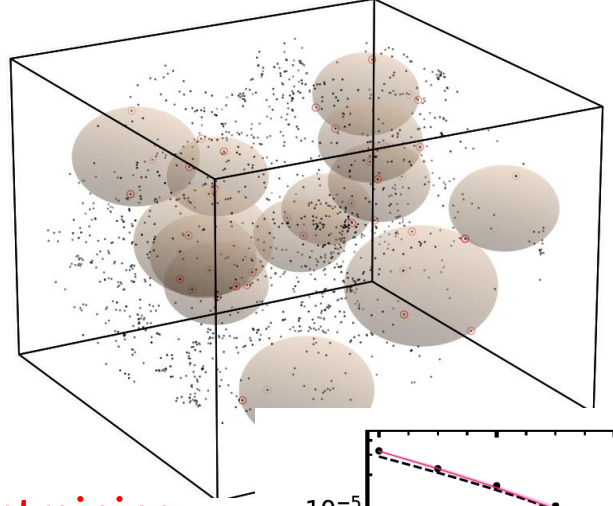
Uchuu lightcones

- Lightcones mock catalogs based on Uchuu simulation
- Abundance matching (V_{peak} – luminosity function) + color assignment
- Reproduce statistics of SDSS main galaxy survey

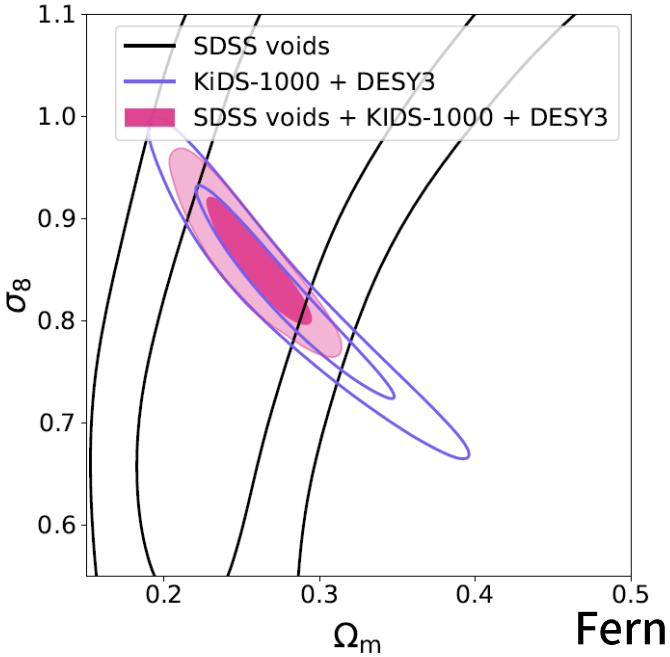


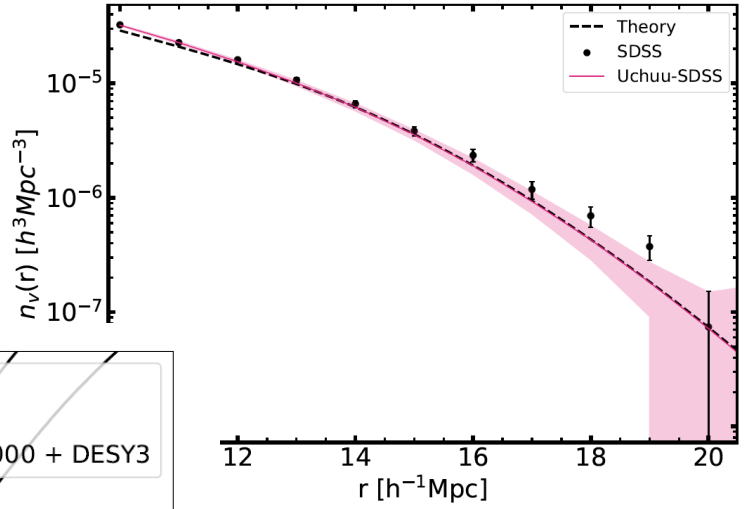
Dong-Páez, TI+ 2024, MNRAS

Uchuu-BOSS lightcones are also available (Ereza, TI+ 2024, MNRAS)



Constraining cosmological parameter using SDSS void





Fernández-García, TI+ 2025

Summary

- Predicting galaxy cluster’s mass and concentration using mock galaxy catalogs and convolutional neural network
- The accuracy is quite well
  - Including richness increases the accuracy for mass prediction
  - Using only relaxed sample increases the accuracy for concentration prediction
- Caveat: rely on Planck18 cosmology