Dark Matter Halo Density Calibration

Titus Nyarko Nde

Computing Ph.D.

September 14, 2022

Contents

General Introduction

A dark matter halo is a hypothesized region dominated by gravitationally bound invisible matter in which galaxies, galaxy clusters, and groups form and live. Dark matter halos are fundamental building blocks of large-scale structures and the cosmic web in general.

According to the Lambda Cold Dark Matter (hereafter ΛCDM) model, our universe is dominated by dark energy ($\sim 68\%$) and dark matter ($\sim 27\%$ and $\sim 85\%$ of all matter in our Universe). Baryons constitute only $\sim 5\%$ of the Universe but $\sim 15\%$ of the total matter in the Universe.

The Big Picture

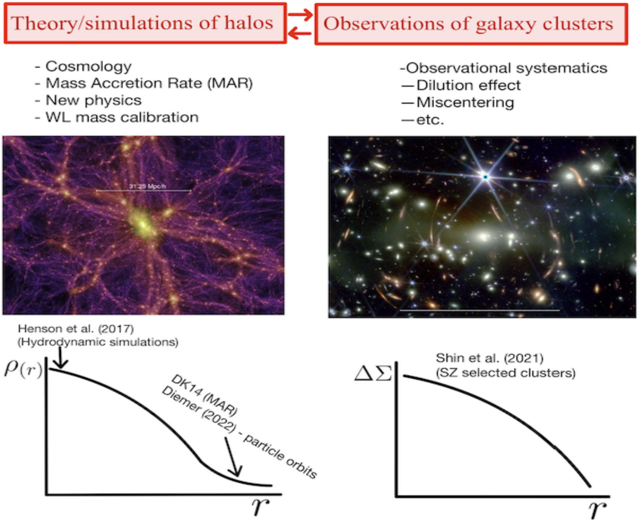

Figure: The big picture

Seed Papers

This synthesis considers the following seed papers:

- **Paper I:** Dependence of the outer profiles of halos on their mass accretion rate (Diemer Kravtsov, 2014).
- **Paper II:** The mass and galaxy distribution around SZ-selected clusters (Shin et al., 2021).
- **Paper III:** A dynamics-based density profile for dark halos – I. Algorithm and basic results (Diemer, 2022).
- **Paper IV:** The impact of baryons on massive galaxy clusters: halo structure and cluster mass estimates (Henson et al., 2017).

Summary of Seed Papers

Paper	Objective	Support	Analysis	Result	Implication
Paper I	Measure density profiles focusing on the outer region	N-body simulation. NFW and Einasto profiles poorly fit outer density profiles	Propose DK14 profile	Detect the splashback feature for the first time in N-body simulations	Observationally detect the splashback feature
Paper II	Measure density profiles	Observational data from ACTDR5 and DESY3. SZ data is more accurate	HOD and DK14 with 2PT	Detects the splashback feature for the first time in SZ data	Splashback feature is not yet well understood
Paper III	Understand the splashback feature	N-Body simulations. Splashback term not well understood	Disentangle orbiting and infalling matter and model them separately	Profiles are dependent on different spatial scales	Single-scale models are insufficient
Paper IV	Measure the impact of baryons on high-mass halos	Hydrodynamic simulations. Ignoring baryons leads to systematic bias in mass estimation	Uses NFW and Einasto profiles	Detects minor baryonic effects on high-mass halos	Baryonic effect remains uncertain in high-mass halos

BOISE STATE UNIVERSITY

Cosmology and Simulations

Table: Cosmological parameters from different astronomical surveys.

Cosmology	$\Omega_{m,0}{}^{a}$	h^{b}	$\sigma_8{}^{c}$	$n_s{}^{d}$
Einstein-de Sitter	1.0000	0.7000	0.8200	1.0000
WMAP1 (2003)	0.2700±0.04	0.7200±0.05	0.9000±0.1	0.9900±0.04
WMAP7 (2011)	0.2743±0.007	0.7020±0.014	0.8160±0.024	0.9680±0.012
Planck 2018	0.3111±0.0056	0.6766±0.0042	0.8102±0.006	0.9665±0.0038

[a] Matter density parameter
[b] Hubble constant
[c] Density fluctuation
[d] Scalar spectral index

Cosmology and Simulations — The σ_8 Tension

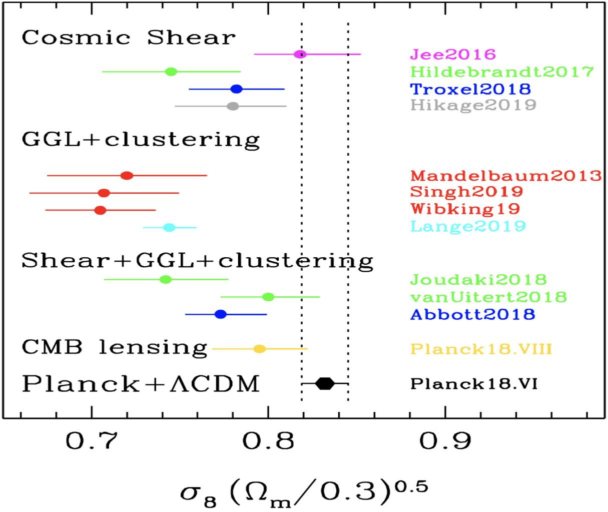

Figure: The σ_8 Tension

Analytic Fitting Functions — 3D Density Profiles

$$\rho_{NFW} = \frac{\rho_{crit}\delta_{crit}}{\frac{r}{r_s}\left(1+\frac{r}{r_s}\right)^2}, \quad \rho_{crit}\delta_{crit} = \rho_s, \quad r_s = \frac{R_{200}}{c}. \tag{1}$$

$$\rho_{Einasto} = \rho_s \exp\left(-\frac{2}{\alpha}\left[\left(\frac{r}{r_s}\right)^{\alpha} - 1\right]\right). \tag{2}$$

$$\rho_{DK14} = \underbrace{\rho_s \exp\left(-\frac{2}{\alpha}\left[\left(\frac{r}{r_s}\right)^{\alpha} - 1\right]\right)}_{\text{Einasto}} \times \underbrace{\left[1+\left(\frac{r}{r_t}\right)^{\beta}\right]^{-\frac{\gamma}{\beta}}}_{\text{Transition term}} + \underbrace{\rho_{outer}}_{\text{Outer term}}. \tag{3}$$

$$\rho_{outer} = \rho_m \left[b_e\left(\frac{r}{5R_{200}}\right)^{-s_e} + 1\right]; \quad s_e > 0. \tag{4}$$

Analytic Fitting Functions — 2D Density Profiles

The 2D density equation is given as follows:

$$\Sigma(R) = \int_{-l_{max}}^{l_{max}} \rho\left(r = \sqrt{R^2 + l^2}\right) dl, \tag{5}$$

The excess surface density is given as:

$$\Delta\Sigma(R) = \bar{\Sigma}(< R) - \Sigma(R). \tag{6}$$

Observationally, we can compute $\Delta\Sigma$ directly from WL tangential shear, γ_t, via the following equation.

$$\Delta\Sigma(R) = \bar{\gamma}_t(R)\Sigma_{crit}(z_l, z_s). \tag{7}$$

3D Density Plots

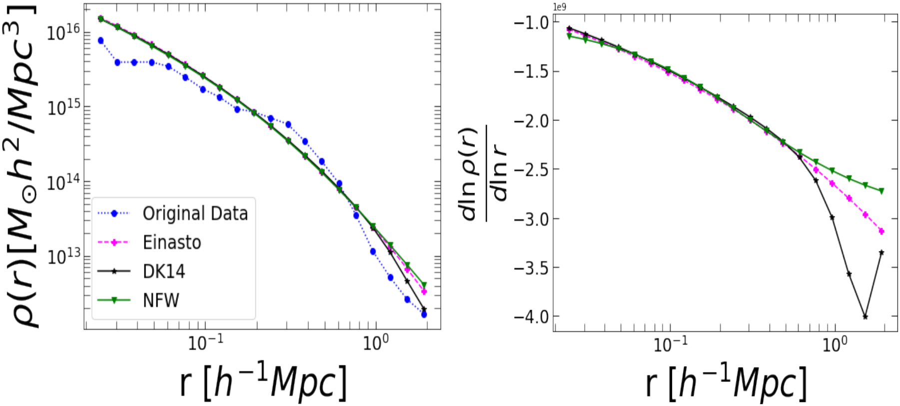

Figure: *Left:* 3-dimensional DK14, NFW, and Einasto density profiles vs a halo from the Buzzard N-body simulation. The halo plotted above has the following properties; $M_{vir} = 3.9 \times 10^{14}, z = 0.39, c = 3.97, R_{vir} = 1.21$, and $= 0.24$. *Right:* The logarithmic slope of the analytic density profiles.

Current Status of the Field

Observations of Halo Density Profiles

- Weak Lensing (WL) — bending of light from source galaxies
  - richness-mass relation
  - concentration-mass relation
- Halo Occupation Distribution (HOD) — the number of galaxies per halo of a given mass
- Galaxy number per unit area — the number of galaxies per unit area or volume

Observational data from astronomical surveys come with several systematic effects. These systematic effects contaminate the data to some extent.

Syetematic Effects

Dilution Effect

This is the situation where galaxies that do not belong to the source/background sample are mistakenly counted as though they belong to it. This mistake is typically associated with foreground galaxies or galaxies that are at the lens. This leads to the underestimation of the WL signal.

Miscentering Effect

Density profiles are calibrated around halo centers. Therefore, correctly identifying the halo center is very important. Usually, the location of the Brightest Cluster/Central Galaxies (BCGs) is taken to be the halo center. However, BCGs are not always located at the center of galaxy clusters.

Summary of Results from Seed Papers — Synthesis Matrix

Theme	Paper I	Paper II	Paper III	Paper IV
Outer profile	Depends on MAR or PH	Not well calibrated by 2PT and HOD	Depends on MAR. Dominated by infalling matter	Dark matter dominates at large radii
Splashback feature	Steeper in halos with high MAR or PH	Location consistent in WL and galaxy density	Steeper and occurs faster in high-mass halos	—
MAR and PH	Density profiles highly depend on total MAR or PH	MAR can help us understand halo formation and evolution	Density profiles highly depend on total MAR or PH	—
Halo boundary	Sets R_{vir} as halo boundary	Sets R_{500c} as halo boundary	Sets $R_{200m,all}$ as halo boundary	Sets R_{200m} as halo boundary
NFW and Einasto fit	Splashback and profile dependence on MAR not captured	Do not capture the splashback feature	Splashback and profile dependence on MAR not captured	Einasto better than NFW; but both underestimate masses

The Splashback Feature

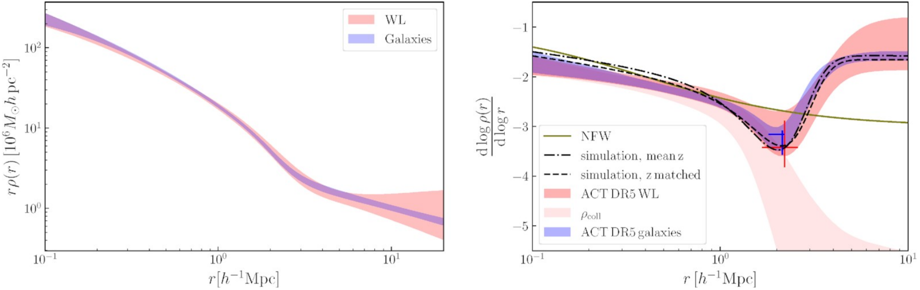

Figure: *Left:* 3D profiles calibrated from galaxy number density and WL. *Right:* Logarithmic slope of the density profiles. The truncation in the orbiting term, ρ_{coll}, occurs before the location of the splashback radius. (Credit: Shin et al., 2021.)

Gaps in the Literature and Future Paths

- **Poor predictions:** As shown in Fig. 3, analytic models usually result in predictions that do not capture the detailed shape of halo density profiles.
- **Uncertainty quantification:** Quantifying prediction uncertainty has been a major concern. None of the analytic models in the literature can do this.
- **Halo evolution:** We cannot observe the entire evolution of a ΛCDM halos in a lifetime but we can simulate them — this can help us predict halo masses.
- **Splashback feature:** The mass distribution within the splashback feature is still poorly understood. The location of the splashback radius is also poorly understood.
- **Solving observational systematics:** There is no unambiguous way of handling systematic effects. We need more efficient modeling strategies to fill this gap.
- **Low-mass halos:** Most studies impose a mass cut to eliminate low-mass halos from their analysis.

My Ongoing Research: Uncertainty Quantification

The general intuition is to build an army of models whose aggregate performance is significantly better than any of the models in isolation. **Uncertainty quantification techniques help to measure a model's confidence in its predictions.**

Some Uncertainty Quantification Techniques

- Monte Carlo Dropout
- Deep Ensembles
- Bayesian Neural Networks
- Masksembles
- Random Forest

This presentation covers only the first two models in the list above. Note that the list of models is not limited to the one above.

Types of Uncertainty

Below are some of the types of uncertainty that can affect predictions.

Aleatoric Uncertainty

Also known as data uncertainty, captures noise that arises from data and therefore is irreducible because the natural complexity of the data directly causes it.

Epistemic Uncertainty

Also known as model uncertainty. It accounts for uncertainty in the parameters of the model that we learned.

Distributional Uncertainty

Also known as dataset shift. This type of uncertainty arises when there is a mismatch between the training and testing data distributions, and a model is confronted with unfamiliar samples.

Deep Ensembles (DEs)

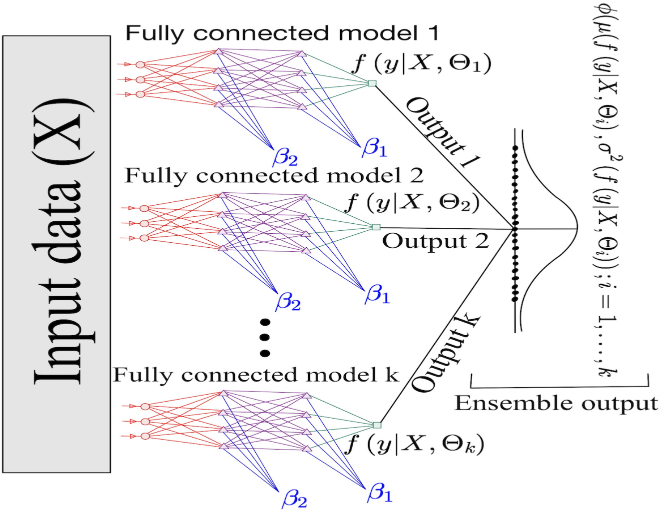

Figure: Illustration of DEs.

Monte Carlo Dropout (MCD)

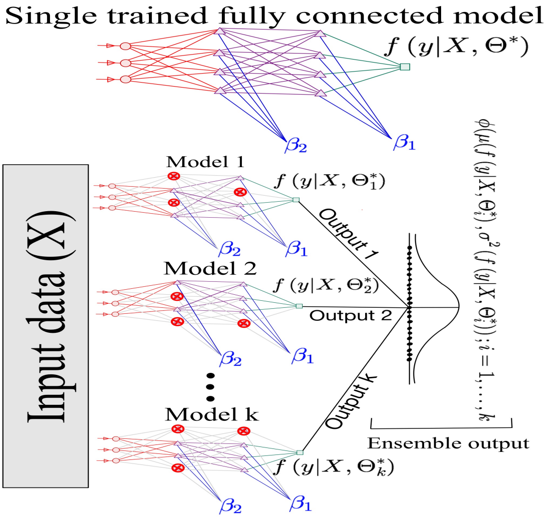

Figure: Illustration of MCD.

Some Results

I present a computing artifact that uses Machine Learning to calibrate dark matter halo density profiles: The aim of this artifact is to:

- Get better predictions than existing analytic models.
- Quantify uncertainties on dark matter halo density predictions.

The computing artifact is available on a public GitHub repo via this link.

Sample Distribution of Prediction (MCD)

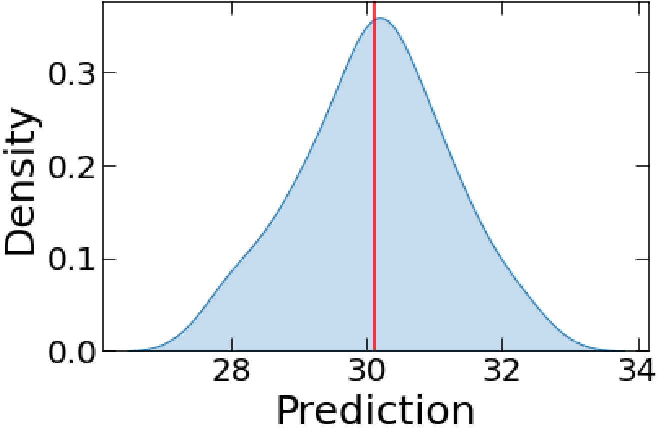

Figure: Sample distribution of a single prediction. Not all the distributions look symmetric. The red vertical line indicates the position of the mean. Halo properties: $z = 0.64078$, $M = 1.85 \times 10^{13}$, $R_{vir} = 0.3834$, $r_{mid} = 0.8455$.

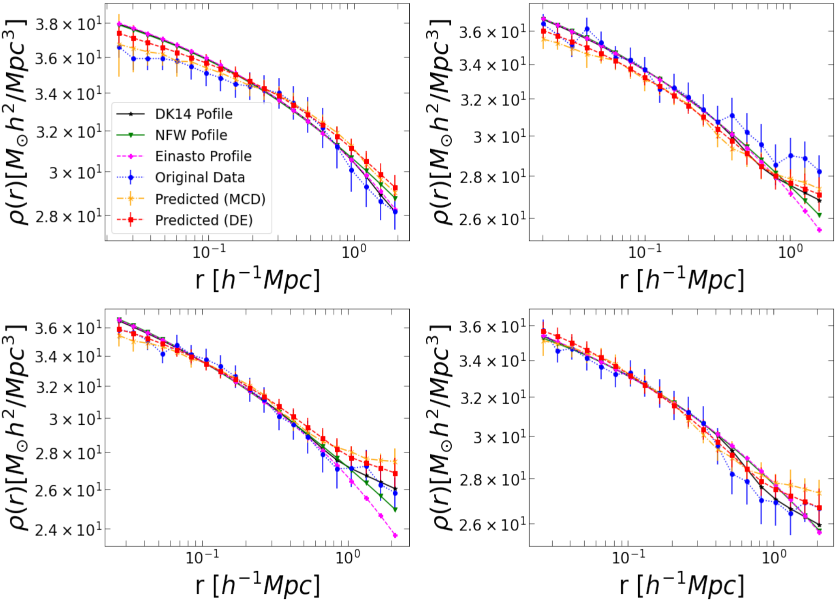

Figure: Some test halos

My Ongoing Research: Actual vs Predicted Values

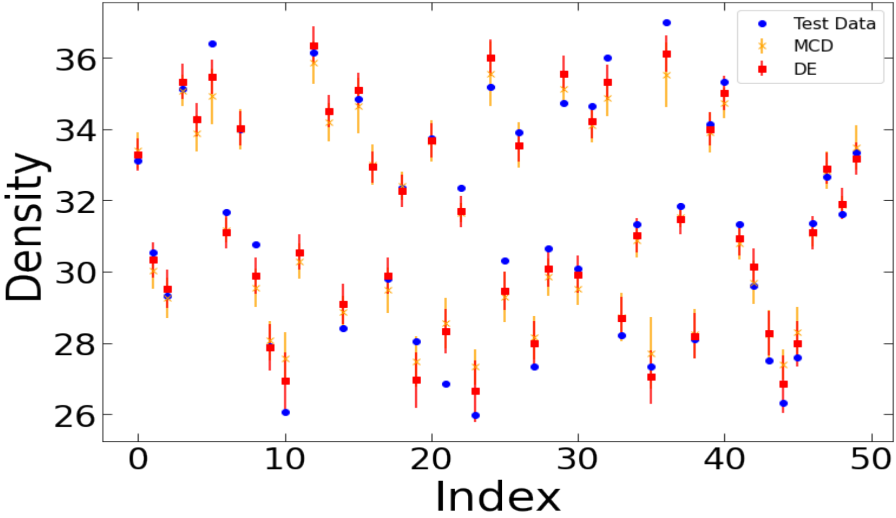

Figure: Predicted vs actual values.

My Ongoing Research: Some Numeric Metrics

Layers	Nodes	NFW MSE	Einasto MSE	DK14 MSE	DEs MSE (Chi-sq)	MCD MSE (Chi-sq)
3	8-8-1	0.3807	0.5160	0.3548	0.2799 (0.8043)	0.2895 (1.2576)
4	8-16-16-1	0.3807	0.5160	0.3548	0.2442 (0.7970)	0.3128 (0.9516)
5	16-30-50-20-1	0.3807	0.5160	0.3548	0.2418 (0.7561)	0.3201 (1.06143)
8	8-64-128-256-256-128-64-1	0.3807	0.5160	0.3548	0.2512 (0.5050)	0.3653 (0.5121)

MCD and DE have adjusted R-squared values of 0.9573 and 0.9664, respectively.

BOISE STATE UNIVERSITY

Recap

- Understanding the dynamics of dark matter halos goes a long way to help in understanding our Universe. However, there is a need to reconcile theory with observations in cosmology.
- Uncertainty quantification on density predictions has been a widely asked question in cosmology. The existing analytic models cannot do this.
- Machine learning ensemble models can help us quantify prediction uncertainties.
- Structure and mass distribution around the splashback feature is still poorly understood
- We are yet to discover any significant impact of baryons in high-mass halos.
- As part of a bigger project we are considering more models and more data (such as the 2D density profiles).

Acknowledgement

Many thanks to my supervisor, Dr. Wu for the great support in my preparation for my comprehensive exam.


Thank you