CASPEN EXIT REPORT

Visitor: Luisa Lucie-Smith (UCL)
Host: Risa Wechsler (KIPAC Stanford)
Dates: 11/02/2019 — 01/03/2019

Machine learning the profile of dark matter halos

Dark matter halos are the fundamental building blocks of cosmic large-scale structure. In our current paradigm for the formation of structure in the Universe, halos formed hierarchically through mass accretion and mergers with smaller structures. The mass accretion history of individual halos can vary drastically — halos of fixed final mass may have formed through numerous small merger events or a small number of larger, violent mergers. Despite the large variety of non-linear physical processes characterising halo formation, N-body simulations revealed that the radial density profiles of N-body halos can be well described by an r^{-3} decline at large radii and a cuspy profile of the form r^{-1} near the centre, known as the NFW profile. Remarkably, this functional form provides a good fit to halos over two decades in radius for a large range of halo masses and for several different cosmological models.

Together with Risa Wechsler and Susmita Adhikari, we wish to provide some physical insight into the physics determining the density profiles of halos predicted by simulations using a qualitatively new method based on machine learning. Previous approaches mainly focused on investigating correlations between physical quantities from large samples of halos, as for example that between concentration and halo formation time. Correlations of this type provide a qualitative understanding of the complex processes seen in simulations. However, a fundamental question which remains open is — how do individual halo density profiles build up over time? To what extent can one predict the final halos' profiles from properties of the initial density peaks and how do the mass accretion histories impact the final halo structure?

We train a machine learning algorithm to predict the final profile of halos, given different choices of inputs about the initial density field and the mass accretion history (MAH). The aim is to test whether information on the initial density field and the mass accretion history are able to fully determine the final profile of halos. Using a quantity known as feature importances, we will be able to further investigate which of the inputs that we provide to the algorithm are predominantly responsible in determining the profile. Our method differs from previous approaches as the machine learning algorithm will be able to pick out non-linear correlations between the final halo profile and the different inputs. We also plan to compare to existing theories which form the basis of our current understanding of halo growth. One example is that we will test the hypothesis that the accretion history at early times is the main driver of the inner halo profile, whilst the outer profile is determined by late-time accretion.

During my first week at Stanford, we agreed on the plan of the project and carried out the first steps to set-up the N-body simulation we will use to train the machine learning algorithm. We used an N-body simulation of box length L=400 Mpc/h and N=2048^3 particles. We restricted our analysis to a small range of massive halos (1 — $2\text{x}10^{14}$ Msol/h), but intend to extend the analysis to halos of lower mass. In the following weeks, we proceeded in building the machine learning framework. We trained the algorithm to predict the final profile of halos, given different choices of inputs about the initial density field and the mass accretion history. The training was done on a subsample of dark matter particles from the N-body simulation. For each dark matter particle living in these halos, the algorithm predicts the ratio M(r)/M(vir), where M(r) is the mass enclosed within a given radius r from the centre of the halo and M(vir) is the viral mass.

First, we provided as input spherical overdensities around each dark matter particle of 30 different smoothing mass scales. The resulting predictions showed a rather low degree of accuracy, which we plan to compare to existing methods which also predict the final halo profile given initial spherical overdensities. We then added information about the mass accretion history of the halos. We first parametrised the MAH of halos with a two-parameter functional form and used the two best-fit parameters of each halo as machine learning inputs. This yielded a large improvement in the mass profile predictions. Predictions became near to perfect when using the actual MAH, i.e.

the value of the virial mass at different redshifts, rather than the two parameters of the MAH functional form. This showed that the initial density and the MAH fully determine the final profiles of halos. Future work will involve interpreting these results using feature importances to understand what aspects of the MAH impact different radial ranges of the profile. We expect to write a paper on this work in the next months.

During my stay at Stanford, I gave a tea talk to a broad audience of the astrophysics group and an informal presentation about my previous PhD projects at Risa's group meeting. I had interesting discussions with daily visitors who were invited to give the astrophysics colloquium, such as Kevin Bundy and Renee Hlozek, and other members of Risa's group such as Ethan O. Nadler and Arka Banerjee.

This work was supported by collaborative visits funded by the Cosmology and Astroparticle Student and Postdoc Exchange Network (CASPEN).