CASPEN Exit Report
Niall Jeffrey (University College London)

Host: Francisco Villaescusa-Navarro (CCA, Flatiron Institute)
Dates: 22nd-26th May 2023

Purpose of visit
The visit to the Flatiron CCA had two interconnected aims. The first was to collaborate with Francisco Villaescusa-Navarro and the CAMELS (Cosmology and Astrophysics with MachinE Learning Simulations) team to apply a new methodology, Evidence Networks*. This technique, developed by myself and Ben Wandelt (IAP/CCA), performs high-dimensional Bayesian model comparison and could be applied to hydrodynamic simulations. CAMELS is a suite of cosmological simulations that vary both astrophysical and cosmological parameters, in conjunction with different astrophysical process prescriptions, which can serve as a rich source of training data for model comparison. The second aim was to present Evidence Networks as a new methodology as part of "Cosmic Connections: A ML X Astrophysics Symposium at Simons Foundation", an interdisciplinary forum attended by machine learning researchers from academia and industry.

Scientific background
The Evidence Network approach enables Bayesian model comparison for implicit (likelihood-free or simulation-based) inference for applications where the likelihood or prior is intractable or unknown. Even when the statistical model is known, Evidence Networks are independent of the dimensionality of the parameter space and of the complexity of the posterior probability density function. They are a simple solution for high-dimensional Bayesian model comparison, with applications to scientific and real-world model inference tasks. Bayesian model comparison using the CAMELS suite of simulations is currently intractable – we expect Evidence Networks to be a widely used and impactful tool for these analyses.

Outcomes
My visit to the CCA, in partnership with the CAMELS team, resulted in the initiation of a project aimed at contrasting different astrophysical models, specifically various hydrodynamical simulation prescriptions, using observational data. Specifically, we take advantage of existing work that maps simulated galaxy properties to simulated photometric data to generate mock Sloan Digital Sky Survey (SDSS) catalogues. Utilizing these mock SDSS catalogues, which employ distinct choices of astrophysics (e.g., Illustris TNG, SIMBA; for more information, visit www.camel-simulations.org), we aim to employ Evidence Networks for Bayesian model comparison. The project is ongoing, with a concentrated effort being made to ensure the robustness of the eventual results based on data (the current work is strictly performed using simulated data).

As a notable outcome of this project, a MAPS summer intern at UCL has begun conducting related work, having expressed an interest in participating.

Upon successful completion, this methodology will offer a novel and efficient approach to constraining the currently unknown astrophysical prescriptions in hydrodynamics simulations using observational data.

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

*Jeffrey & Wandelt: arxiv:2305.11241