Abstract
Cosmological large-scale structure encodes a wealth of information about the nature, the origin and the evolution of our Universe. Its study requires comparisons of observations with theoretical predictions. In the last decade or two, we have entered the epoch of precision cosmology. Modern surveys such as the Planck mission, Euclid and others collect large amounts of high quality measurement data. In order to exploit the full potential of that data, their precision and accuracy have to be matched by those of the theoretical predictions they are compared to. Analytical models fail to provide the required level of accuracy and precision when small, so called "non-linear", spatial scales are studied. In this regime, computer simulations (called "N-body simulations") have served as the de-facto theory as for years the precision and accuracy of their results have been unchallenged. While providing high-fidelity predictions, computer simulations are notoriously expensive and require specialized high-performance computing infrastructure. As a result, certain tasks as e.g. Bayesian parameter inference or forecasting still rely on the faster but lower-fidelity analytical models. Only in the last couple of years data-driven tools, in the field called "emulators", have entered the game. They have the potential to revolutionize the field by providing highly accurate and precise predictions within fractions of a second on a standard commercial lap top.

The core of this thesis is the discussion of EuclidEmulator, a new, efficient and highly accurate emulator I have developed to predict one of the key statistical quantities of cosmological large-scale structure: the matter power spectrum. This emulator has been developed having principles of both machine learning and uncertainty quantification in mind and was trained to nearly perfectly reproduce the simulation results of pkdgrav3, an N-body simulation code well known for its very high precision and one of the main simulation codes of the European Space Agency (ESA) mission Euclid. The accuracy requirements imposed by the Euclid Consortium have been extremely challenging. EuclidEmulator, however, is able to meet them thanks to highly optimized training data and learning pipelines.

EuclidEmulator is public and being used by several research groups of the Euclid collaboration.