Abstract

The present thesis is the result of the work I have conducted together with my supervisor Robert Feldmann and colleagues over the last four years at the Institute for Computational Science (ICS). I decided to structure this thesis in a way such that readers are guided from the bigger picture to the details, where most of our contributions are connected to. The thesis begins with a brief historical introduction to the vast field of cosmology and galaxy formation acting as a motivation for further chapters. In the following part of the thesis we discuss fundamental theoretical concepts starting from cosmological derivations and connect them to galaxy formation on astrophysical scales. I tried to condensate and formulate with my own words the most fundamental physical processes involved in the formation of galaxies and to present them in a manner that is easy to understand. That includes many figures displaying conceptually the physics that are at play at the different formation stages. Then I introduce important concepts in Deep Learning, since they are the foundation of the algorithms that I developed during my projects, followed by the main contributions of my work to the field.

The different regimes of galaxy formation models can roughly be separated into empirical models, (semi-) analytical models (SAMs) and hydrodynamical simulations. While SAMs are statistical approaches that typically include coarse grained information (like e.g. the halo mass), hydrodynamical simulations yield detailed insights into the process of galaxy formation. However, due to the small number of systems that are typically simulated and the large parameter space of subgrid models, simulating samples that are large enough to make statistical statements is expensive. Naturally, there is a trade-off between the two regimes. In recent years, Machine Learning has been proven to be able to capture complex correlations in datasets, while operating on time scales that are much shorter than simulations. In this thesis I explore the use of neural networks and Deep Learning to emulate aspects of structure formation on the field level.

In Chapter 5, we present a neural network model (termed EMBER-1) that is capable of mapping atomic hydrogen and total gas on top of dark matter fields. We also investigate the capabilities to upsample the gas fields in resolution showing that Machine Learning models can in fact synthesize complex high-resolution gas fields from very limited input information. In Chapter 6 and Chapter 7 we extended the datasets to multiple baryon fields and trained a more powerful model (termed EMBER-2) on large redshift ranges and multiple resolution levels. The results from these projects prove that neural network models are valuable alternatives that can synthesize realistic astrophysical fields. Moreover, EMBER-2 is designed in a way that allows to modulate the synthesized fields as a function of external parameters that encode global information like e.g. the redshift. EMBER-2 also encompasses a new strategy to emulate physical fields on varying length scales connecting the environment of the large cosmic web with small scale structure like galaxies, which is especially suitable for the next era of large scale observational efforts that the field is moving into.