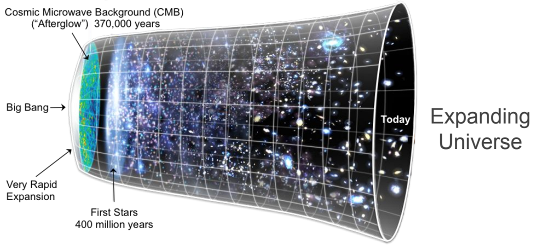
Machine Learning for Cosmic Structure Simulations

Lauren Haley, Raul Hernandez, Maybrit Ortiz, Celine Tan, & Xing Liu   Mentor: Cooper Jacobus

University of California, Berkeley   Undergraduate Lab at Berkeley, Physics & Astronomy Division

Background

Understanding the universe's evolution requires studying the **distribution of matter**. Simulating visible matter is complex and expensive, while **simulating dark matter is more efficient and cost-effective** because it does not interact with itself. Dark matter accounts for 85% of the universe's mass and dominates large-scale dynamics, influencing visible matter. By comparing observational maps with simulations, we can gain insight into the **nature of matter, gravity, and dark energy**.

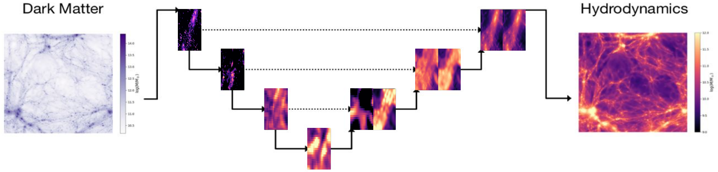


Objective

Our team has developed a novel approach to tackle the challenge of **creating visible matter maps** that are both **accurate** and **cost-effective.** By utilizing machine learning, we have developed a model that can perform **baryon inpainting** and fill in missing of visible matter, using only the dark matter **N-body simulations** as input. Our approach optimizes **statistical similarity** between our less expensive mock simulations and accurate observational maps, **reducing the cost** of generating visible matter maps.

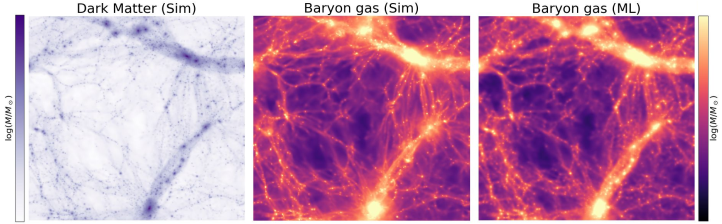
Methodology

We've developed a **neural network** that can **translate between dark matter and hydrodynamic simulations**. The network is trained to learn the features of the simulations and **minimize the differences** between the predictions and the 'true' simulated mocks. We've used a neural architecture called a **UNet**, which can **decompose images** into **feature maps** and **reconstruct them**. This is useful for accurately extrapolating hydrodynamic features from dark matter-only simulations.



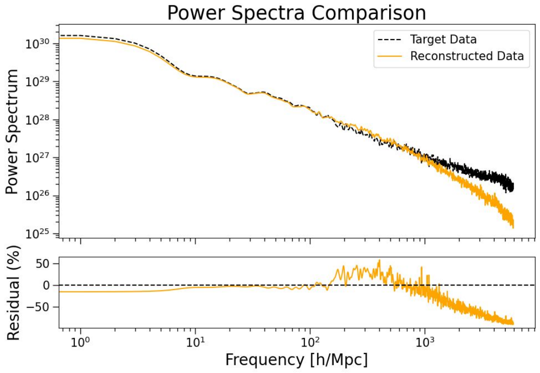
Data Processing

We obtained our training data from a suite of simulations called **IllustrisTNG**, which have been processed for machine learning.Our approach involved creating **pairs of images** by combining **hydrodynamic simulations** with their **dark matter-only counterparts.** We then fed the dark matter-only image into our neural network and gradually optimized (trained) our model so that its outputs were as close as possible to the second image in each pair.



Results

To ensure the accuracy of our machine learning model, we evaluate its performance by comparing the results with the target simulations. Our approach compares both **individual pixels** and critical **statistical properties** of the images to ensure their similarity. We use global statistics, such as the **power spectrum**, to contrast our model with expensive simulations. We strive to reconstruct these properties as accurately as possible.



Conclusion

Overall, we were **able to produce realistic maps** of visible matter using only dark matter simulations. Our surrogate model **accurately captures complex features** and **reproduces global statistics**. Methods like ours can be used to **accelerate expensive simulations**. Future work for similar projects could entail the use of more physically motivated training and architectures.