The DREAMS Project: DaRk mattEr and Astrophysics with Machine learning and Simulations

Alex M. Garcia[1,2], Jonah Rose[3,4], Paul Torrey[1,2]

[1]University of Virginia, [2]NSF Simons Cosmic AI Institute, [3]Princeton University, [4]Center for Computational Astrophysics

This material is based upon work supported by the U.S. National Science Foundation. NSF AIVO is funded via U.S. National Science Foundation (NSF) awards 2231251, 2332864, 2437003, 2528968, and 2536256 and a grant from Google.org. Any opinions, findings, conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of NSF or Google.org.

Dark Matter

Dark Matter is one of the largest systematic uncertainties in all of Astrophysics research. Both observational and simulation campaigns suggest that there *must* be some invisible matter holding galaxies together, yet particle physicists have yet to find experimental evidence for these "dark" particles. Understanding the properties of dark matter on the large scales is therefore critical to understanding what it is and how it behaves.

Astrophysics

Dark Matter is not the only thing that shapes the evolution of galaxies. The physics of stars, gas, black holes, dust, etc. (i.e., "Astrophysics") also has a distinct role in shaping the evolution of galaxies in our Universe.

Quantifying the extent to which astrophysics shapes galaxy evolution compared to dark matter physics is critical to our knowledge of dark matter and its effects.

Machine Learning

Since we have a large, multi-dimensional problem, we aim to use modern machine learning methods to help. The covariance of astrophysics and dark matter physics may be too subtle to measure of quantify with traditional methods.

Simulations

Simulations of galactic evolution have demonstrated a range of success across a range of scales. This work makes use of the popular and successful IllustrisTNG (https://www.tng-project.org/) project as a beginning framework.

The DREAMS Project

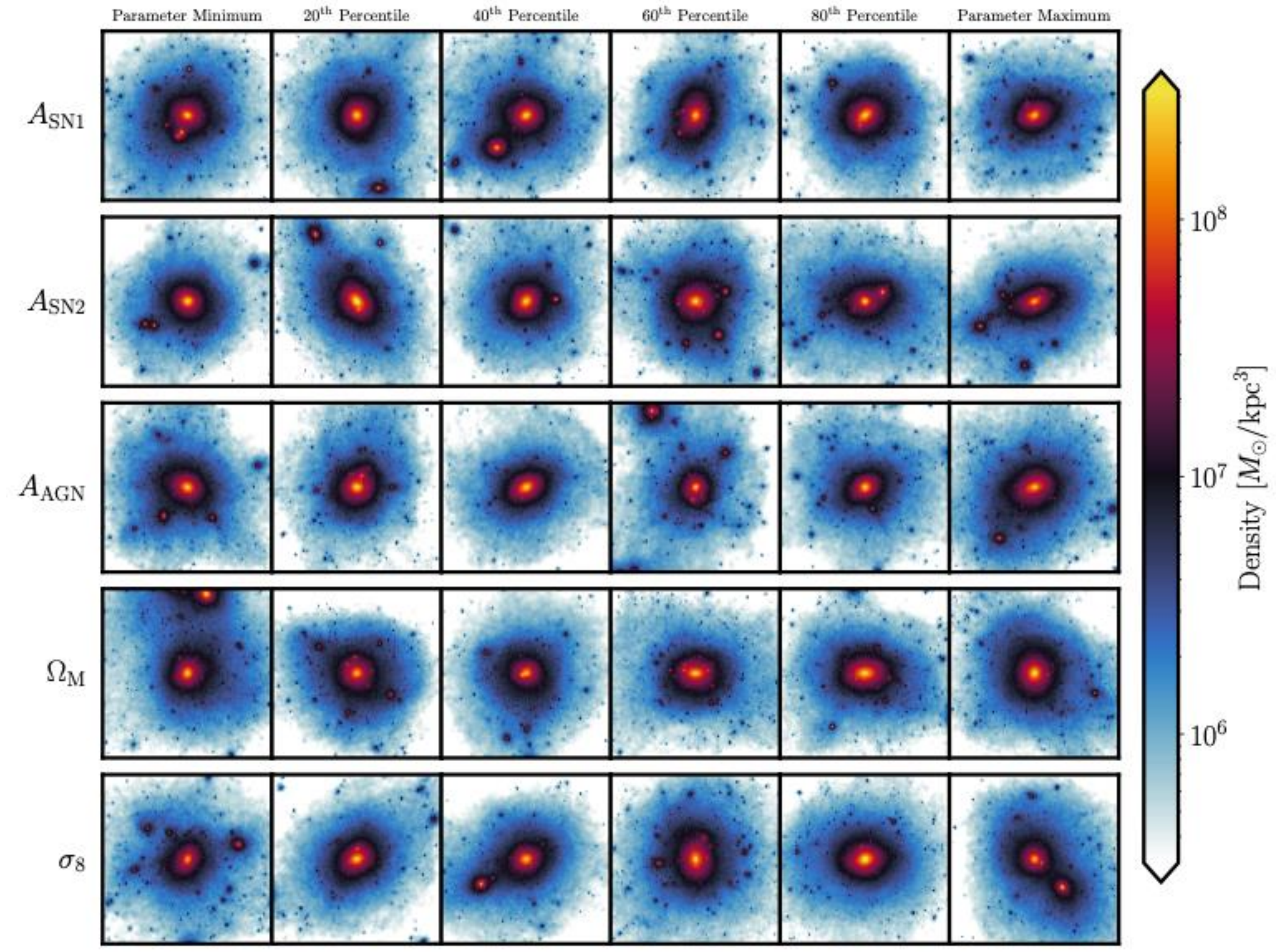
https://www.dreams-project.org/

The Problem

Traditional simulation approaches rely on single implementations of finely-tuned models that reproduce only a few select observed galaxy scaling relations.

However, these approaches do not often generalize to *all* scaling relations, nor are they guaranteed to be unique solutions in a model's parameter space.

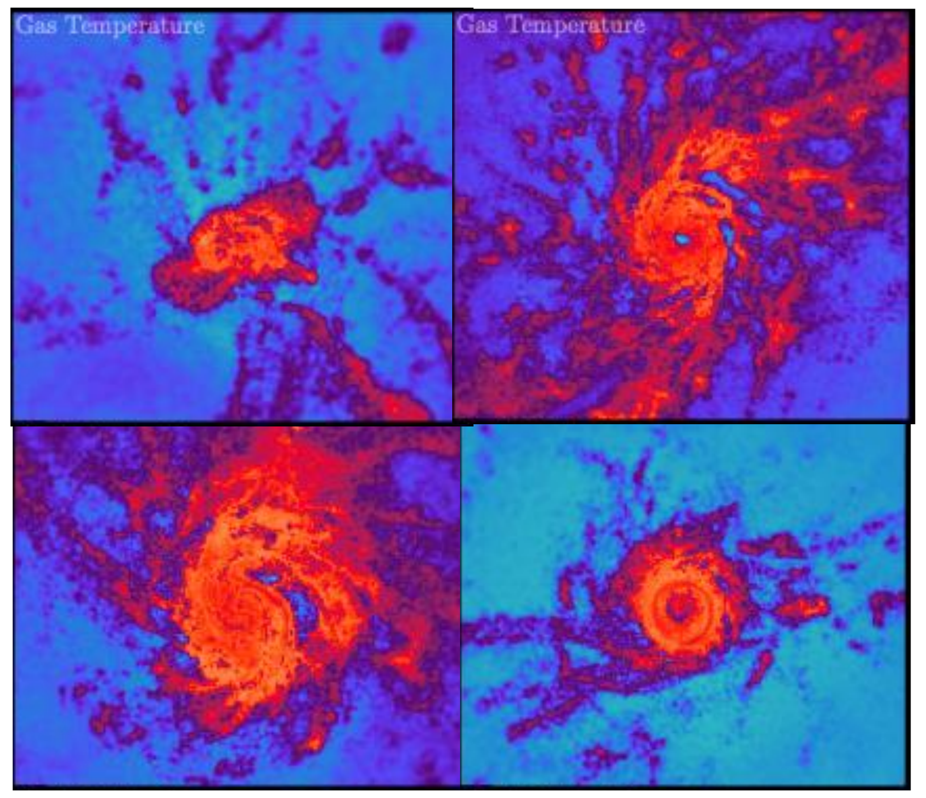
Since the simulation models themselves are using a fixed realization of galaxy physics, it is not clear how the assumptions we make map into the observable properties of the galaxies themselves. Moreover, these simulations are expensive to run.



A sample of 30 Dark Matter halos from the new suite of 1024 DREAMS galaxies. For a sense of scale, each image is approximately 1.6 million light years across

The "galaxy" (all of the gas, stars, dust, and black holes) are all deeply imbedded in the center of these dark matter halos. These halos have "substructure" (A.K.A. satellites, like the small and large Magellenic clouds).

Our DREAMS sample represents the largest sample of Milky Way mass galaxies at the highest resolution ever run.



The gas temperature distribution of four DREAMS galaxies

Our Approach

Instead of relying on a single implementation of galaxy physics, we use a 1000s of variations in astrophysical, numerical, cosmological, and dark matter parameters to measure their relative impact on galaxies.

We therefore sample our model parameter space more completely. Using this method, we can understand the extent to which individual model assumptions matter for individual properties of galaxies.

This approach fundamentally requires Machine Learning architectures (e.g., Neural Network Emulators, Graph Neural Networks, Diffusion Models) but has the potential to save millions of CPU hours of compute running simulations.

Simulations To Date

To date, we have run 5000 simulations of Milky Way mass galaxies over two simulation suites. The two suites comprise of a standard "Cold" Dark Matter (CDM) and "Warm" Dark Matter (WDM) models.

In the coming months, we are expanding our suite to include 1000 more simulations. This new suite will include variations in our target halo mass: ranging from dwarf galaxies up to small galaxy clusters.

Quick Results

While the DREAMS collaboration efforts are still in their early days, we have already published a few papers using these new simulations. These works range from more astrophysics to machine learning.

- We find that the number of satellite galaxies corresponds with your Dark Matter Type using Neural Network Emulators (Rose et al., 2024; arXiv.2405.00766)
- We build a diffusion model that is able to generate samples of millions of galaxies consistent with the parent simulations (Nguyen et al. 2024; arXiv.2409.02980)
- We find that Graph Neural Networks can learn global properties of our simulations when conditioned upon merger histories (Leisher et al., In Preparation)

End Goal

Eventually, we want to generalize this approach to not just Dark Matter searches, but to many different properties of galaxies. In doing so, we aim to save millions of CPU hours in training the next generation of galaxy evolutionary simulations

Award numbers will go here.