Cosmology Simulation of Dark Matter Scientific Visualization Project Report

Fengshi Zheng, Cyrill Imahorn, Robert Veres and Leonardo Salsi

December 21, 2023

1 Introduction

The cosmos offers an unending tableau of mysteries, chief among which is the phenomenon of dark matter. Dark matter, an unseen and unidentified form of matter, makes up about 85% of the universe's substance, and makes direct detection impossible, revealing its presence only through its gravitational influence on visible matter [5]. Understanding dark matter and its evolution is not just a curiosity-driven pursuit of science, but it holds key insights into the fundamental nature of our universe and its origins.

As we continue to refine our digital competencies and augment our computational prowess, the visual analysis of such large data sets presents fresh avenues for gaining insight into the cosmos. Consequently, this project stands as a leap towards evolving the study of cosmology from being predominantly observational to a more engaging and interactive field. By enabling an in-depth exploration of dark matter evolution, our goal is to illuminate some of the most enigmatic corners of the universe, inching us closer to unravelling the intricate tapestry that weaves reality itself.

2 Data

Our project utilizes an intriguing and rich data set that originates from the Hardware/Hybrid Accelerated Cosmology Code (HACC) team from the Argonne National Laboratory [4]. It consists of an array of simulated cosmological information including positions, velocities, mass, internal energy, the smoothing length parameter of the smoothed-particle hydrodynamics (SPH), molecular weight, density, and gravitational potential of particles. Their simulation was conducted in a manner that accurately emulates the fundamental laws and principles governing our universe. To gain some specific insights and a deeper understanding of the underlying mechanism of the development of our universe we developed a framework that simulates the evolution of cosmological structures, providing us with a deeper understanding of our universe.

The data can be classified into several types of celestial matter, i.e., we can distinguish between stars and gas within said data set. It is organized on a grid system, allowing us to accurately model the vast expanse of the universe and track the evolution of various celestial bodies over time. As for the attributes, they vary across a range of types - positions, velocities, and the SPH smoothing length are quantitative attributes. These are numerical in nature and can be measured on an interval scale. The data also includes nominal attributes such as the distinction between dark matter, baryon, star, wind, and star-forming gas particles, and active galactic nuclei (AGN). These are categorical and do not have a quantitative relationship with each other.

Regarding its nature, the data set comprises both scalar and vector data. Scalar data includes values like mass, internal energy, molecular weight, density, and gravitational potential. These represent quantities that have magnitude but no direction. On the other hand, vector data includes parameters like positions and velocities, which encompass both magnitude and direction.

By leveraging this extensive and multifaceted data set from the HACC team, we can dive into the study of dark matter evolution, illuminating one of the universe's most profound mysteries.

3 Goals

Our project, spread over an eleven-week timeline, was built on a meticulously planned road map that was designed to address several ambitious goals. These objectives were aimed at developing a sophisticated tool capable of providing insightful visualizations into the fascinating world of dark matter and its evolution. In this section, we outline our project goals, their time frame, and the outcomes achieved.

Weeks 0 - 3: In the first phase of our project, we concentrated on constructing a solid groundwork to guide our subsequent efforts. Three principal goals were targeted in this phase.

The first was to enable interactive browsing of data, a crucial requirement for any user-friendly analytical tool. This included providing an overview of the rich data set and enabling slicing of the data based on time, providing a temporal perspective. We are proud to say that this goal was fully achieved.

Our second objective was to interpolate a continuous scalar field, using the Smoothed Particle Hydrodynamics (SPH) as an example. Interpolating this scalar field was a critical step in effectively utilizing our data set, and we successfully achieved this objective.

Our third goal during this phase was to develop capabilities for data type filtering and value thresholding. This functionality was aimed at providing users the ability to filter the data based on specific types or values, a crucial feature for detailed and targeted analysis. We were able to successfully implement this feature, marking the successful completion of all our goals for this phase.

Weeks 5 - 7: Moving into the second phase of the project, we sought to build upon the strong foundation laid in the first phase.

Our primary focus was to visualize the distribution of various particle types, such as baryons, dark matter, stars, wind, and active galactic nuclei. This functionality would allow users to gain a spatial perspective of these particles' distribution, a key insight into understanding their behavior and interaction. We were successful in realizing this goal.

Our next aim was to create a density-based clustering mechanism for stars and AGNs. By achieving this, we enabled a more nuanced understanding of the spatial relationships and groupings among these entities. We're pleased to report that this goal was successfully met.

During this phase, we also aimed to develop an interactive GUI to enhance user experience and interaction with the data. We were able to implement a fully interactive GUI, providing a user-friendly interface that significantly enhances the project's overall utility.

Weeks 8 - 11: As we embarked on the final phase of our project, our goals became more ambitious.

Firstly, we sought to visualize the temporal evolution of the cosmos over time in a single image. This feature would provide a unique time-lapse view of the cosmos, giving a holistic perspective of the evolutionary processes. We were successful in achieving this, providing a powerful tool for understanding the temporal dynamics of the universe.

The interactive GUI was also enhanced during this phase, adding more capabilities for a richer user experience. Furthermore, we introduced animation and video export capabilities, providing more ways for users to engage with the data and share their findings. Both these goals were successfully achieved.

However, one of our aspirations in this final phase was to visualize the cosmic web, such as the streakline visualization. Despite our best efforts, we were unable to fulfill this objective within our project timeline.

In conclusion, while some our goals remains unrealized, the accomplishments over our eleven-week journey are substantial. The ability to interactively explore, filter, and visualize the cosmological data marks a significant advancement in our study of the creation of filaments. The work completed sets a good foundation for future exploration and for potential improvements, including the visualization of the cosmic web. The journey through these ten weeks has been rewarding, bringing us closer to unraveling the mysteries that the cosmos holds.

4 Visualizations

The data was initially processed by parsing .vtp files, followed by sorting the points according to their bitmask. This allowed us to distinguish between various forms of celestial matter effectively.

Following the preprocessing phase, a Visualization Toolkit (VTK) [2] agent was created for each type of data. The agents were designed to be approached and modified singularly, providing the flexibility to represent the data in diverse ways. A GUI was built by integrating VTK into PyQT5 [1] to facilitate the data representation process. Multiple data views were constructed based on the data's requirements, including Data View, Type Explorer, and Volume View. The

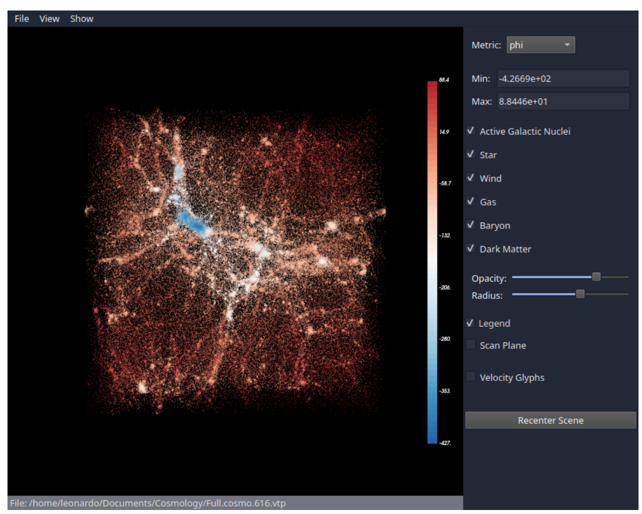

Figure 1: Data view of the visualization system for the cosmology data.

Data View visualizes the properties of each point relative to the selected property's magnitude, such as energy potential or velocity.

The Type Explorer provides the ability to filter points by their type, like baryons, dark matter, stars, or even an aggregate view.

The Volume View implements voxelization of the data space and applies volume rendering to present a more spatially comprehensive view.

Our system also supports snapshot capture of the current data view and video compilation of all time steps, providing a dynamic look into the celestial data's evolution.

4.1 Part 1 - Dark Matter Clustering

This first visualization focused on revealing the characteristics and dynamics of dark matter based on its gravitational potential. The choice of focusing on dark matter was motivated by our intention to delve into the intricacies of how dark matter clusters are formed as a result of their gravitational potential.

To represent the gravitational potential of the dark matter, we utilized a color scale where cooler colors represent lower potential and warmer colors represent higher potentials. This color mapping was strategically chosen as it enables an intuitive understanding of the data, given that people often associate cooler colors with lower values and warmer colors with higher values.

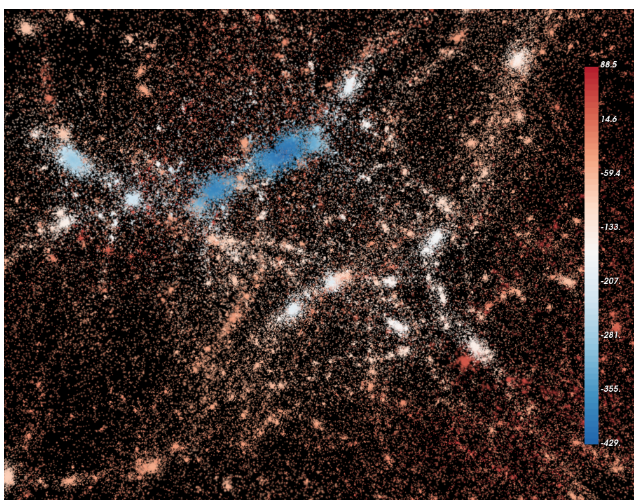

Figure 2: Dark matter color coded depending on its gravitational potential.

The visualization is indeed interactive, albeit with a few restrictions. While the color mapping for the values is fixed, users can select different time steps and perspectives to view the data from various angles. This interactivity also extends to the choice of different metrics for examination, offering the flexibility to delve deeper into specific aspects of the data.

The primary narrative we aim to convey through this visualization is the speed and process by which dark matter clusters form due to their gravitational potential. This is evident in the image where we see data points depicted in blue, symbolizing dark matter with negative potential. Negative potential implies an attractive gravitational force that forces these data points to cluster together.

4.2 Part 2 - Filament Creation from Trajectories

In this second part we focused on other aspects of dark matter data; the movement patterns of dark matter particles. We found these to be crucial in understanding the dynamic behaviors of dark matter in the universe and in the creation of filaments.

To visualize these complex aspects, we implemented a method that introduces vectors according to the velocities of the data points. These vectors, effectively arrows, represent the trajectories of the particles, thus illuminating the movement patterns. This was coupled with the use of the same color-coding scheme as before, which was designed to represent the gravitational potential of each point.

The visualization method was chosen with a specific objective in mind. We sought to uncover and illustrate the intricate interplay of particles possessing different potentials. More specifically, we aimed to depict how these particles are influenced by the forces of gravity, both pulling and pushing, that they exert upon each other. The use of arrows signaling their trajectories provides a clear and intuitive way to observe these effects.

As we follow these trajectories over time, we can see an emergent pattern. The dynamic interplay of forces and movement patterns leads to the formation of filaments, revealing a structure that evolves over time due to the movement and positions of the particles.

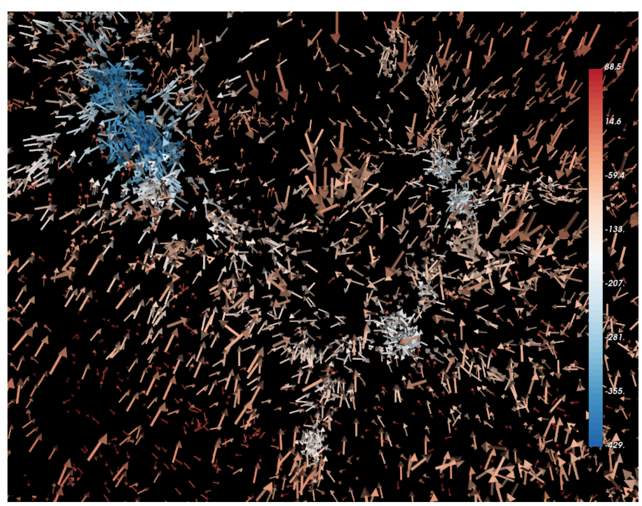

Figure 3: Movement vectors of dark matter including color coding for gravitational potential.

This visualization is not only interactive in terms of choosing how many arrow glyphs we want to sample and how articulated they need to be, allowing for an exploration of the data in a meaningful way, but it also provides vital insights. It gives us a detailed picture of how dark matter behaves and evolves in the universe, contributing significantly to our understanding of its role in cosmic structure formation. Thus, the visualization method serves as a potent tool for unveiling the otherwise hidden intricacies of dark matter dynamics.

4.3 Part 3 - Temperature Distribution and Potentials

In this part we specifically focused on discerning potential relationships between gravitational potential and temperature within our dataset. Our primary aim was to detect and visually express any correlations between these two variables, thereby providing insights that might not be easily identifiable through numeric or tabular data alone.

We opted for a volumetric rendering approach, a choice primarily driven by the three-dimensional nature of our data[1]. By mapping temperature distributions onto a voxel grid, we were able to create an intuitive and representative three-dimensional visualization. This was further augmented by employing color coding, with blue and red hues designating colder and hotter temperatures, respectively. This choice of visualization was justified due to its capacity to clearly illustrate complex spatial relationships and variations within the data.

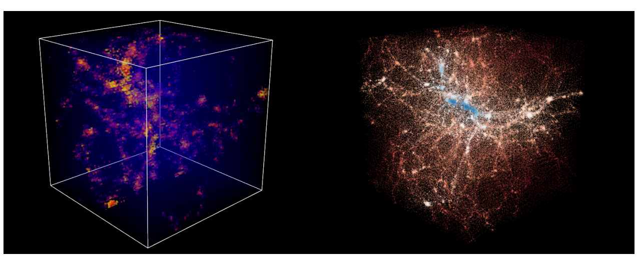

Figure 4: Temperature distribution within the data set. For comparison, same view on the point data with color coded gravitational potential.

In terms of implementation, our visualization method involved the creation of a voxel grid. For each cell, we gather nearby points to determine their attribute value assigned. We then render this three-dimensional grid with VTK's "VolumeRayCastMapper", because of its perfect GPU optimization [3]. Using the temperature as our attribute of interest, this creates a visual map of temperature variations within the data set. The granularity of this representation provided a comprehensive depiction of temperature fluctuations, thus enabling easier identification of trends and patterns.

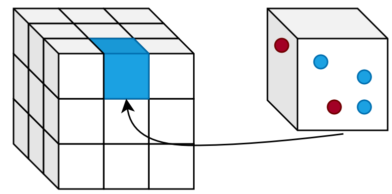

Figure 5: Our volume rendering relies on gathering points that were clustered into a certain grid cell based on their position

The visualization method offered limited interactivity, primarily allowing for alterations in perspective and time steps. Despite these restrictions, the method proved effective in visualizing our data set and identifying significant trends. From our observations, we inferred a strong correlation between higher gravitational potentials and increased temperatures. This insight, gleaned from the effective utilization of volumetric rendering, demonstrates the power of sophisticated data visualization in uncovering complex relationships within multi-dimensional data sets.

[1] One could use more reliable statistical metrics such as the Pearson correlation or an arbitrarily complex regression analysis, but since this would be outside of the scope of this work, it is left as an exercise to the reader

5 Contributions

5.1 Fengshi Zheng

- Plane interpolation
- Particle type visualization
- Velocity field visualization
- Point Gaussian visualization with SPH radius scale

5.2 Cyrill Imahorn

- Interactive Animation with preloading (not used in final product)
- Worked on stream- and streaklines (with no success)

5.3 Robert Veres

- Data visualisation with matplotlib as a fallback for our internal use
- Gathering data points into three-dimensional grids

- Creating color and opacity mapping for mass and temperature
- 3D-volume rendering

5.4 Leonardo Salsi

- Integration of VTK into PyQT5
- Particle type coloring
- GUI controls and interactive data manipulation
- Data point filtering and thresholding
- Temporal interpolation of .vtp data
- Animation and video creation

6 Discussion

6.1 Limitations

Our current data representation is point-based, lacking structural information that could allow for the extraction of visual cues regarding filaments. Rendering, constrained to a single viewpoint, is time-consuming due to the necessity of data loading and preprocessing. Animation playback is not real-time capable and necessitates prior rendering. Despite the implementation of temporal interpolation, we do not perform computations across multiple time steps.

6.2 Future Work

There are multiple additional things we would like to add if we had more time:

1. **Contour Tree**:
   Contour trees would be an interesting addition, since it would allow to get more insights about the formation of the filaments. It would be much easier to grasp the extent of them and therefore work with them.

2. **Measurement Capability**:
   Especially combined with the contour tree, but not only, measurement capability would allow to quantify aspects of the filaments. Like the volume or the mass of them. This would again likely generate interesting insights about the filaments.

3. **Particle Tracing**:
   More focused on the formation of the filaments, particle tracing could be interesting, since we hope to see the flow of the particles converging to the filaments and generally see how different sectors in space behave.

4. **Further Work on Animation**:
   We are quite happy with our animation so far, but there are some additional functionalities that would be nice to have. For example control over the camera (over different time steps) would be good, so you could really emphasise a region you want to look at.

5. **Bugfixing**:
   Although we are able to produce satisfactory visual results, we still encounter some singular cases in which our implementation encounters some unforeseen bugs, e.g. when switching between from and to the Volume View.

References

[1] Qt for python.
`https://doc.qt.io/qtforpython-5/`.

[2] Visualization toolkit.
`https://vtk.org/`.

[3] vtkSmartVolumeMapper Class Reference.
`https://vtk.org/doc/nightly/html/classvtkSmartVolumeMapper.html`.

[4] A. N. Laboratory. Cosmology dataset.
`https://cpac.hep.anl.gov/`.

[5] R. Leay. Dark matter particle that may finally shed light on cosmic mystery the 'best of both worlds,' scientists say.
`https://www.space.com/dark-matter-new-model-hyper-particles`.