Research Proposal

Baryon-Dark Matter Dissociation on Mpc Scales in Cosmological Simulations

Jiaqi Yang

Supervisors: Prof. Danail Obreschkow, A/Prof. Aaron Ludlow and Prof. Simon Driver

Abstract

During the growth of cosmic structure, baryons separate from dark matter (DM) due to multiple processes. These include both radiative processes, like gas cooling and radiative feedback, and non-radiative ones, such as hydrodynamic pressure. Studying baryon-DM dissociation on different scales can help us isolate and constrain these processes, as well as quantify the back-reaction of baryons on the DM density field. The separation of the two forms of matter has been well studied on cosmological (>100Mpc) and galactic (<1Mpc) scales. Comparatively little is known about dissociation on intermediate scales (1-10 Mpc), although cosmological simulations hint at significant and clear differences between the baryon and DM fields on these scales. Structures at these scales do not have well-defined boundaries like halos, making them difficult to quantify theoretically. Observationally, they are either faint or rare. All these challenges make them difficult to study statistically. However, new cosmological simulations and upcoming observations are starting to make this possible. In my PhD project, I aim to understand how baryons and dark matter separate at scales of 1-10Mpc. Building on previous work focused on non-radiative processes, I aim to explore dissociations caused by radiative processes. I will develop a general, quantitative way to describe them and find their observational signatures in X-ray surveys. By using the large volume of the FLAMINGO simulations, I will reconstruct the evolution of baryon-DM dissociation at Mpc scales and find its relation to the radiative processes. This work could provide new insights into how cosmic structure formation are affected by baryonic physics.

Section A: Project Summary

Formation of the Cosmic Web and the Baryon-DM Separation

Galaxies are often the first objects we think of in astronomy. Each one contains billions of stars orbiting its center, but these stars make up only a small part of its mass. Over 80% is believed to be dark matter (DM), invisible material detected only through its gravitational pull[2]. The other 'ordinary' matter, in the form of gas and stars, is called baryon. Although DM hasn't been detected directly, different observations such as gravitational lensing and rotational curve of galaxies, yield similar fractions of DM. The decisive observation comes from the measurement of Cosmic Microwave Background[10], where the distribution of overdensities is determined by the total amount of matter of the Universe, where the fraction of DM can be derived. The most common model of DM is the cold dark matter model, where the thermal velocities of DM particles are low. In this proposal, we refer to DM as the cold dark matter.

Since DM makes up most of the Universe's mass, it largely determines how matter is distributed. Because DM interacts only through gravity, regions that are already dense will grow even denser. On large scales of a few hundreds of Mpc, CDM forms a web-like pattern known as the "Cosmic Web," characterized by strong contrasts between crowded and empty regions (see Fig. 1 (a)). The emptiest regions are called voids, while the densest are known as nodes. These nodes are linked by long, thread-like structures called filaments.

On smaller scales, clumps of DM collapse into self-bound structures called "halos." In denser regions of the Cosmic Web, bigger fluctuations make halos easier to form, so they often appear close to one another. In these regions, especially at the nodes, halos attract and merge, and larger halos can draw in even more halos. This assembly process produces the largest self-bound structures, called galaxy clusters, which can contain hundreds of halos. Baryons, meanwhile, mostly follow the deep potential wells carved out by DM, much like paper drifting on waves. In the early Universe, the distribution of baryons closely matches that of DM.

However, baryons can interact with each other through electromagnetic forces, which can be stronger than gravity in plasmas and gas. Unlike gravity, these forces can be repulsive. As a result, baryons gradually separate from DM. This separation is observed on scales ranging from the cosmic web down to individual halos, as shown in Fig. 1. On cosmic scales, baryons tend to be more extended in the nodes and filaments. Within individual halos, group like halos with masses of $10^{13} M_{\odot}$ contain fewer baryons compared to larger or smaller ones ones(see also [11] [1]). Looking closer at the radial distribution within halos, the inner regions have fewer baryons than the outskirts compared to DM. In the following, We refer to this separation as the "dissociation" of baryons and DM.

The dissociation causes several unique phenomena in the Universe. On cosmic scales, Because of hydrodynamic pressure, baryons collapse more slowly than DM and end up oscillating, creating the so called Baryon Acoustic Oscillations(BAO). The imprint of these oscillations remains visible today in galaxy surveys[8]. On galaxy scales, gas that falls into halos cools down and forms disks at their centers [34]. When these disks become dense and cool enough, stars begin to form and burn within them. These stars eventually make up the galaxies we see today.

Different types of dissociations are caused by different baryonic processes, offering a unique window to isolate and study them. For instance, on cosmological scales, the size of BAO depends on the Universe's overall density and expansion rate [15]. Its observation at the predicted scales in galaxy surveys confirms our robust understanding of cosmic evolution. On galaxy scales, the so-called galaxy–halo connection describes how galaxy form and evolve in their host halos, and how their properties affected by that of halos . Within halos, dissociations also reveal processes we cannot observe directly. For example, if the optical center is offset from the halo center in a cluster, it suggests the cluster has recently experienced disruption [30]. Similarly, a long gas tail behind a galaxy indicates that it has fallen into its host cluster very recently, and is losing gas due to ram pressure stripping [14].

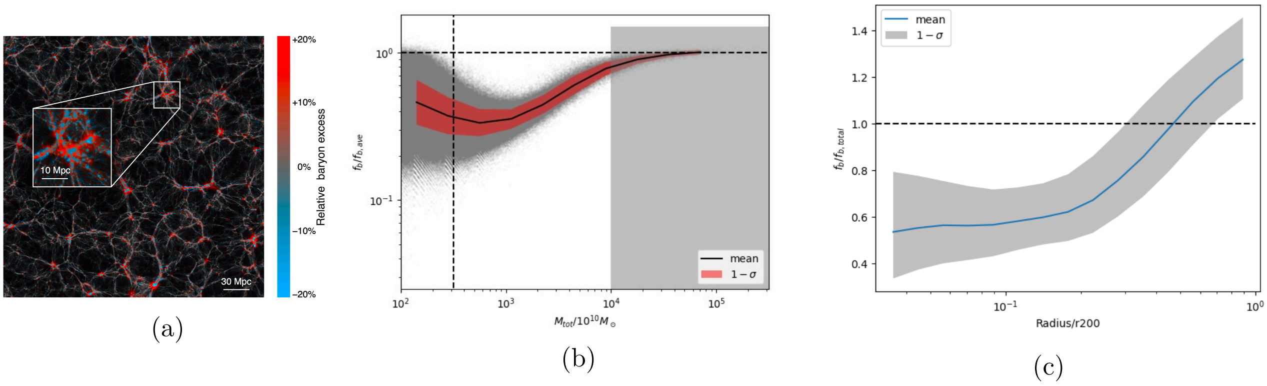

Figure 1: The level of dissociation at different scales from the FLAMINGO simulation. The left, middle, and right panels show dissociation in the cosmic web, within halos, and at different radii inside halos, at scales of 100 Mpc, Mpc, and sub-Mpc respectively. Left: The color indicates the level of baryon density excess compared to DM. Middle: Gray dots represent the baryon fractions (f_b) of individual halos plotted against their masses. The black solid line and red shaded region show the mean and 1-σ scatter within each mass bin. The vertical black dashed line marks the mass where a halo contains 1000 particles in the simulation; below this, matter distribution is considered unreliable. The horizontal dashed line indicates the average baryon fraction of the Universe. The gray shaded area marks the halo sample used in the right panel. Right: The radial distribution of baryon fractions in the 10^4 halos heavier than $10^{14}M_\odot$, as indicated by the gray shaded region in the middle panel. The blue solid line and gray shaded area represent the mean and 1-σ scatter of the sample.

The Dissociation on Mpc scales: Achievements and Challenges

Dissociation on scales of 1–10 Mpc, larger than individual halos but smaller than the BAO scales, is key to understanding baryonic processes in the Universe. At these scales, cluster size halos collide through major or minor mergers, heating the gas significantly and driving turbulence within clusters [26]. They also accrete additional matter via filaments. Recent studies show that processes inside galaxies also influence this scale. Gas in the galaxy clusters can gain substantial energy from feedback effects from central black holes in halos, while lose energy through radiative cooling in both halos and filaments [19]. These energy changes may alter the overall shape of the gas. Moreover, the dissociation at this scale can affect the distribution of CDM itself by a few percent [32], which is significant in current precision cosmology.

The most famous example of using baryon–CDM dissociation to understand baryonic processes is the Bullet Cluster, shown in Fig. 2. In the Bullet Cluster, the gas traced by X-ray emissions lies between two DM dominated clumps , mapped through strong lensing. The gas and DM have strongly separated [7]. A natural explanation is the dissociation results from a recent head-on collision of two halos. Because DM interacts only gravitationally, the DM halos passed through each other, while the gas collided, slowed down, and remained in the middle. This scenario is further supported by bullet-shaped shock fronts in the gas, a clear sign of ram pressure. The existence, especially a non-universal one, provides strong evidence for DM over modified gravity models, which predict uniform features in galaxy clusters. By reconstructing the orbits of the two progenitor halos, researchers have also placed upper limits on possible non-gravitational interactions between DM [25].

As more major merger events like the Bullet Cluster are found and modeled (see the reference in [22]), a natural next step is to study how mergers cause separation with large samples. However, general statistical research faces natural challenges at these scales, both in theory and observation.

On large scales, where matter perturbations are small, the initial matter distribution can be described analytically using linear perturbation theory[19]. This theory predicts the properties of halos formed through gravitational collapse. At galaxy scales, although processes become highly non-linear, they remain confined within halos, which have clear boundaries. These characteristics make modeling somewhat easier.

On the other hand, matter distribution at these scales does not share these conveniences. Structures like halo mergers and filaments are not only too complex to model analytically, but also lack clear boundaries. Filaments are not fully virialized and can not be considered as self-bound structures like halos. Galaxy cluster mergers, on the other hand are continuous events. During a merger, the halos are not in equilibrium. Even after the merger, clusters remain unrelaxed for a long time. Because of this, defining the start and end of a merger, as well as its observational signatures, remains a debate (see [17] and references therein). This makes identifying and quantifying these structures challenging.

Observationally, probing matter distributions at these scales is challenging. Gas in filaments is less dense and therefore dimmer than gas in halos. Gas in cluster mergers is bright, but systems like the Bullet Cluster are very rare, about one per $2Gpc^3$ [22]. Understanding their matter distributions usually requires combining multiple observation methods. For instance, studying the Bullet Cluster needs both lensing and deep X-ray data to reconstruct the total matter distribution.

As a result, previous research on matter distribution at these scales has either focused on its effects on smaller-scale galaxy formation or been limited to individual or a small number of structures. The former group of research aims to build direct links between halos and galaxies, viewing Mpc-scale structures as external influences on galaxy formation. The latter studies the matter distribution of specific cases [3], or examines overall gas features through stacking [31].

Recent advances in cosmological simulations and observations may help overcome these challenges. New instruments with both high sensitivity and large fields of view have appeared. For example, In optical band, Euclid [9] can detect the Intra-Cluster Light, the diffuse stellar component in galaxy clusters not bound to any galaxy. WEAVES[13] can detect much fainter and a much larger sample of galaxies compared to the current generation surveys, providing a more complete sample within clusters. In the X-ray band, the eROSITA telescope [24] can detect X-ray emissions from galaxy groups or satellite galaxies in clusters. Future radio surveys like SKA could observe the 21 cm line emission from neutral hydrogen gas in the outskirts of halos or even in filaments[35]. These surveys will enable more robust statistics on these large, faint structures.

Thanks to increased computational power, realistic hydrodynamical simulations covering large volumes have recently become available. With our knowledge of small-scale processes like feedback and star formation, these simulations successfully reproduce phenomena at Mpc scales. Examples include Illustrius[33], BAHAMAS[21], Magneticum[12], and FLAMINGO [29]. Each simulates regions of a few $100Mpc^3$ consistently. Their large volumes provide extensive samples of galaxy clusters and filaments. This opens the possibility of achieving a universal understanding of matter distribution at Mpc scales.

In the study of bullet-like clusters [22], the authors selected halos from the SURFS simulation [16] and tracked their merging histories. They found that gas pressure during mergers fully explains the distribution of quadrupole dissociation. Using similar SURFS halo samples, [20] examined the rotation separation between gas and DM. They successfully explained how gas ends up rotating faster than DM, given the spins, geometries and the collapse rates of halos. Dissociation in filaments is also drawing attention. For example, [3] defined a universal way to identify filaments in simulations, and discussed the possible origins of the dissociation. These works highlight the strength of statistical methods and the importance of describing separations quantitatively.

Both [22] and [20] focused on the effects of gas pressure on the dissociation, while that radiative processes, like gas cooling and feedback remain unexplored. Nevertheless, these studies provide useful references. Before finding the connections between the gas pressure and the dissociation, they firstly identified the dissociation quantitatively. The dissociations identified can not be described in simple ways shown in Fig. 1 and are not spherically symmetric. This suggests that the key to connect the processes and the dissociation is requires finding correct ways of describing dissociation, which is likely to be non-trivial.

Some practical issues lying in observations also worth considering. The information in observations is 2D, which can not be compared directly with the 3D one in simulation. Moreover, the limited angular resolution of X-ray surveys makes it hard to resolve fine details. Strong lensing events, which are the only direct way to probe DM distribution, are relatively rare. To build the connect from simulations and observations, clear signatures of the dissociation visible to future surveys are necessary.

Figure 2: Left: The distribution of gas (red) and DM (blue) in the Bullet Cluster, constructed from data in [7] by NASA. Right: A similarly DM-elongated halo from the FLAMINGO simulation in this work. The stars are colored in white, and the other color codes are the same as the left panel.

Section B: Research Project

Aims and Proposed Projects

Motivated by the theoretical and observational interest in the dissociation, I aim to understand how baryons and DM separate on scales of 1–10 Mpc, as well as their observational signatures. Building on previous research focusing on gas pressure, I will study the effects of radiative processes on the dissociation. I will analyse state-of-the-art hydrodynamical simulations, improving on previous work with a larger dataset. I will work towards a general, quantitative way to describe this dissociation. Specifically, the aims of the three projects I will conduct are as follows:

Project 1: **The dissociation on Mpc scales caused by radiative processes.**

In this project, I aim to find how radiative processes, for example gas cooling or feedback, can cause the separation of gas and dark matter, in particular:

- What kind of dissociation can radiative processes cause on Mpc scales, in galaxy clusters or filaments, and how can it be quantified?
- What is the distribution of the dissociation in FLAMINGO simulations, and how it is different from that in simulations with ideal gas only?
- What can we learn about the evolution of the Universe from this dissociation?

Project 2: **Characterizing the signatures of Mpc dissociation in mock X-ray observations.**

After understanding the relationship between radiative processes and the dissociation in Project 1, I will look for potential signatures of dissociation in X-ray surveys. In particular, I will focus on:

- What are the differences in the X-ray luminosity profiles between structures, i.e., galaxy clusters or filaments, with large and small dissociation?
- How can we probe dissociation with future X-ray surveys like eROSITA?
- How observational effects, such as projection effects and the limited angular resolution of X-ray surveys, influence the observational signatures.

Project 3: **Some potential directions to explore:**

- The dissociation between stars and DM on Mpc scales, especially how well Intra-Cluster Light (ICL) can trace the distribution of DM during major mergers.
- Looking for observational signatures with other datasets, such as WEAVES, Eucild and SKA.
- Looking for other processes that can cause Mpc-scale dissociation.

By the time Project 3 begins, a new set of simulations, Colibre, is expected to be ready [27]. This simulation has higher resolution, allowing it to resolve smaller objects and more diffuse components. If necessary, I will analyze Colibre data with approaches similar to Projects 1 and 2.

Methods

To achieve a universal understanding of these structures in the Universe, we need to study many of them simultaneously and consistently. This can be done using cosmological simulations. These simulations treat matter as discrete mass points and calculate their positions based on their interactions—gravity

Box name	**L1_m10**	L1_m9	L1_m8	L2p8_m9
Box length$//Gpc$	1	1	1	2.8
Resolution	Low	Intermediate	High	Intermediate
Number of Particles$/Gpc^3$	900^3	1800^3	3600^3	5040^3
$m_{DM}(/10^9 M_\odot)$	45.2	5.65	0.706	5.65
$m_{gas}(/10^9 M_\odot)$	8.56	1.07	0.134	1.07

Table 1: The boxes I will use in this project in FLAMINGO simulations, the corresponding resolutions, the numbers of particles for both DM and gas particles, and the particle masses. For the highest resolution, a cluster with $10^{14}M_\odot$ has 10^5 particles. The box in bold is the one I used in this proposal and in pipeline testings.

for DM, and gravity plus fluid dynamics for baryons. In this way, we can obtain the matter distribution in a region in the most consistent way, with the fewest assumptions.

Simulations were restricted to DM for a long time because gravity is easier to model than fluid dynamics. In the past 10 years, realistic hydrodynamical simulations that include gas components have started to appear, for example [28]. However, due to limitations in computational power, simulations must make compromises in two ways. They can only cover a finite region (usually hundreds of Mpc), and the mass of a single particle (usually greater than $10^5 M_\odot$) is much larger than that of individual stars, which is necessary for achieving the same accuracy as in our Universe.

The first effect causes simulations to miss processes on scales larger than the box size and limits the number of clusters within the box. For example, there is roughly one Bullet Cluster per 2 Gpc^3 [22], making it a rare object that requires a much larger box size to study. The second effect causes simulations to miss physical processes on smaller scales. Therefore, models of these small-scale processes are added manually to estimate their effects on the scales well resolved by the simulations. Moreover, the finite mass of particles makes it difficult to resolve small structures due to shot noise caused by limited particle numbers. Typically, at least 10^3 particles are needed to resolve a structure well.

Thus, when choosing simulations, we need to ensure the box is large enough to contain a sizable sample of galaxy clusters, and that each cluster has enough particles to be properly resolved. The simulation that fits these requirements best is FLAMINGO [29], known for its extremely large box size. The largest simulated box, L2p8_m9 covers a volume of $(2.8Gpc)^3$ at $z = 0$, providing a large sample of about 10^6 massive clusters with $M > 10^{14}M_\odot$, including roughly 10 bullet-like clusters, which are very rare. Three resolutions are available. The highest resolution available in the largest box is the intermediate resolution, corresponding to about 10^4 particles per cluster. Although this resolution is lower than that of some other simulations, traded for the box size, it is still sufficient to resolve cluster substructures and mass distributions. A summary of the boxes available can be seen in Table 1

The simulation includes all the hydrodynamical processes required for our study, such as gas cooling and feedback. Since implementing AGN feedback remains an open question, FLAMINGO offers multiple AGN feedback models for comparison [29]. In particular, it provides X-ray emission models in the eROSITA band, with metal cooling effects taken into account [5], making it well-suited for studying gas-related processes. The underlying cosmology follows the best-fit parameters determined by the Planck satellite [10].

The output data of the simulations are snapshots that record all the kinetic and thermal of particles at specific moments in time. A halo finder then determines which halo each dark matter particle belongs to and identifies subhalos, which are halos bound by other halos. After that, baryon particles are assigned to halos. A halo catalogue is then created. To compute halo properties accurately, the FLAMINGO team uses HBT halo finder [18], which is good at detecting subhalos, even if they are close to each other. It is also free from mistakenly assigning the halo center to a satellite, compared to conventional halofinders[6]. The halo catalogue is then arranged in the format of the SOAP catalogue[1], which makes it easy to access the properties in the central and satellite halos.

[1]https://github.com/SWIFTSIM/SOAP

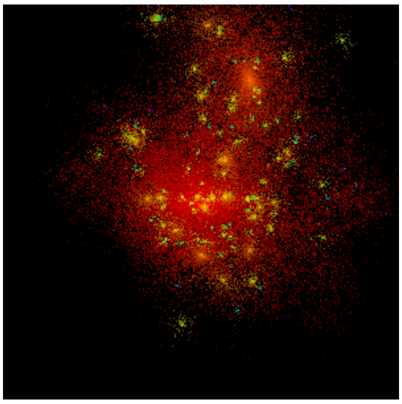

Figure 3: An example of the visualization of the most massive cluster in FLAMINGO L1_m10 box. The subhalos in the cluster are color coded by their masses. Less massive halos are bluer.

Due to the complex shape of galaxy clusters and filaments, we do not have a universal way of describing dissociation. Visual inspection is therefore crucial to build the intuition about how the DM and baryons separate. I use the R package "gadgetry"[2] to visualize the snapshots. The package is able to highlight the particles we want to focus on, as well as color code them with assigned weights. It also implements kernel density smoothing to create realistic, smooth distributions, which is also a necessary step when compare our results with observations. An Example of the visualization by the package is shown in Fig 3

To finish the three projects proposed above, I will conduct the following:

- Extract the matter distributions around halos up to a few Mpc in FLAMINGO(for project 1, 2)
- Reconstruct the evolution of the dissociation in the regions above(for project 1, 2)
- Understand how radiative processes affect the dissociation around one halo, if necessary, by controlled simulations(for project 1)
- Understand how radiative processes affect the dissociation in the cosmological context, and find the distributions of the dissociation(for project 1)
- Construct the X-ray emission and its evolution of the clusters(for project 2)
- Learn about the observational ability of future X-ray surveys(for project 2)
- Find potential signatures of dissociation in X-ray emission in FLAMINGO(for project 2)
- Repeat step 1-4 for other structures or process in FLAMINGO or Colibre(for project 3)

Current Progress

FLAMINGO snapshot files are extremely large, as they store many particle properties to support a wide range of research. A single snapshot can require at least 100 GB of memory to load fully, even for the low-resolution L1_m10 box. The intermediate and high-resolution boxes require approximately 8 and 64 times more memory, respectively. Fortunately, we only need a subset of properties and particles. To efficiently handle the data, we use the Swiftsimio [4] and Swiftgalaxy [23] packages. These tools avoid loading the entire dataset all at once and instead load only the specified particle properties when needed.

However, reading all the properties from the raw data remains time and memory consuming. For L1_m10 box with 900^3 particles, a single property stored as arrays with units can require about 5 GB of memory. In terms of time, loading all the necessary data usually takes hours, limited by disk read speeds. Therefore, it is unrealistic to load the full dataset every time we perform an analysis.

[2] https://github.com/obreschkow/gadgetry

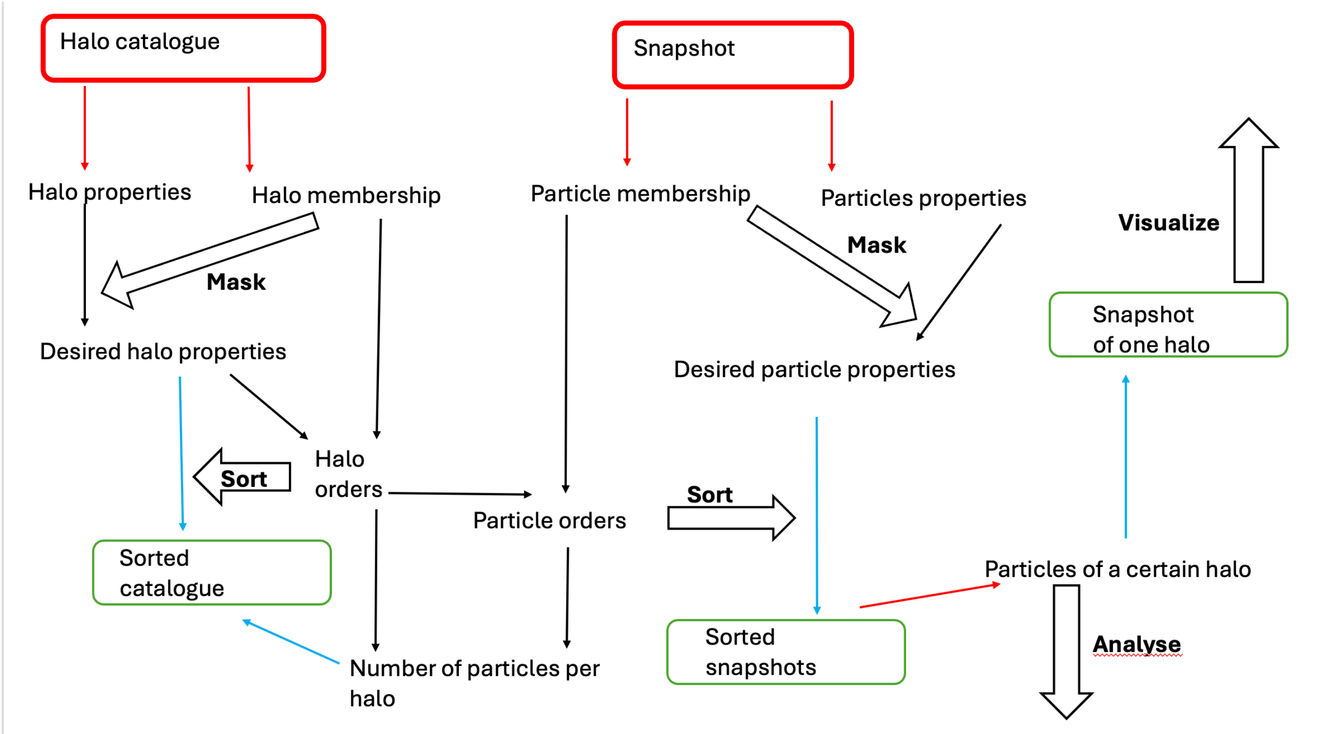

Figure 4: The flowchart of the pipline to reduce the dataset and extract the information I need from the raw FLAMINGO snapshots. The red, green boxes shows the original and resulting dataset. The red, blue and black arrow means read, write and manipulate data. The bold arrow shows the key algorithms used when processing the data, and the input data of these algorithms.

As a result, I developed the pipeline shown in Fig. 4 to pre-load data and save it in reduced snapshot files. The halo properties in the halo catalogues are organized by halo IDs. For each halo property, I sort the data by mass and membership, then save it into a new halo catalogue. To save memory, I load particles in chunks and sort them in the same order as the halo catalogue. In this way, I can find all particles within a halo without the particle membership data. With halos and particles sorted consistently, I not only create smaller snapshots but also avoid loading entire snapshots, thanks to Python's built-in index slicing. When loading 15 properties of 10^4 clusters in L1_m10 box, divided into 30 chunks. The pipeline uses about 500MB of memory at peak and takes 40 CPU minutes. The resulting dataset has the size of 1.8GB. The speed of loading data scales linearly as the number of particles. Therefore, the same reduction takes 5 and 40 CPU hours, for intermediate and high resolution boxes , with the memory use of 5 and 50GB at peak respectively. The memory can be further reduced by loading the data in more chunks.

I calculate the same distribution of the dissociation index defined in [22] in FLAMINGO L1_m10 box , using 10^4 massive halos of $> 10^{14} M_\odot$ and 3×10^4 smaller halos of $10^{13} M_\odot$(see Fig. 5). For massive halos, The dissociation distribution in FLAMINGO is consistent with that in the SURFS simulation [16] used by [22], despite SURFS treats gas as ideal gas without radiative processes. This agreement is expected because radiative processes are inefficient in massive halos. However, in less massive halos, the peak of the dissociation distribution shifts closer to zero and the width of the distribution increases. This new feature differs from [22], where the distribution was independent of mass due to the self-similar nature of halos and their ideal gas. This result demonstrates that FLAMINGO captures the effects of radiative processes on baryon distributions.

Fig. 5 also implies that gas cooling and halo collisions can produce similar effects on the dissociation. Therefore, to study cooling with clean samples, we may need to select halos that are not undergoing collisions. Halos with ongoing collisions can be identified by their "dynamical state" (for a review see [17]). There are several ways to define dynamical state; Fig. 6 shows one by determining whether the most massive galaxy lies within the central halo. Fig. 6 illustrates it is equivalent to the distance between the center of the most massive galaxy and the halo center. With this criterion, the distance shows a

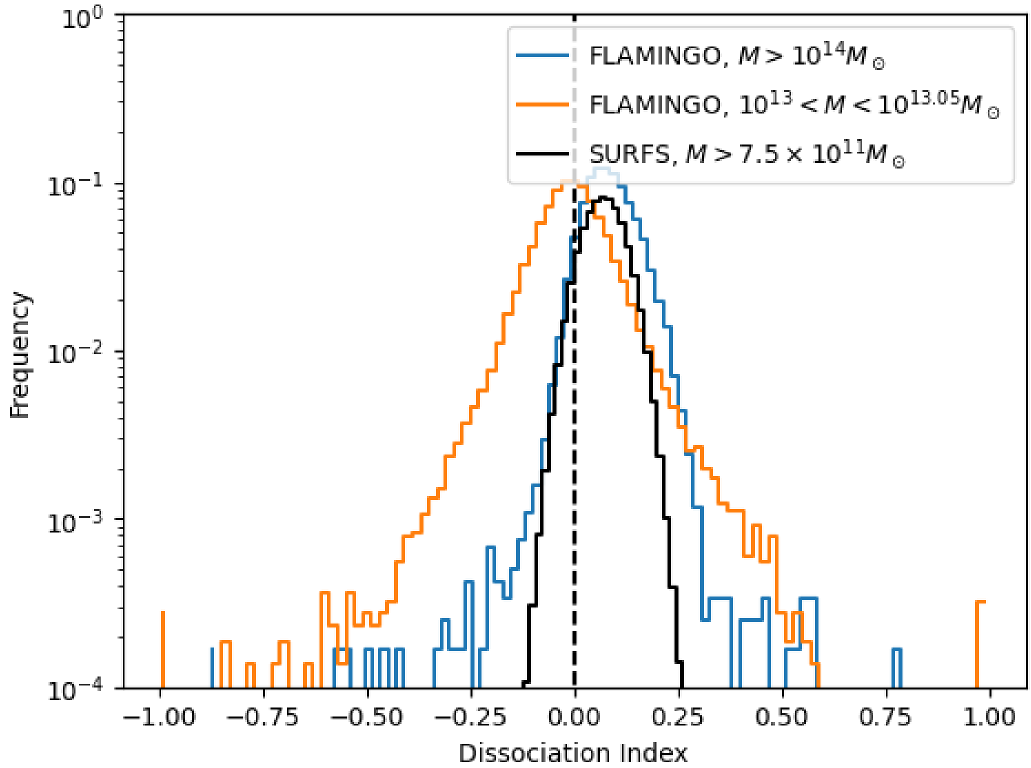

Figure 5: The frequencies of the dissociation indexes defined in [22], for all the halos with masses $M > 10^{14} M_\odot$(blue, $N_{halo} = 10^4$) and $10^{13} < M < 10^{13.05} M_\odot$(orange, $N_{halo} = 3 \times 10^4$) in FLAMINGO L1_m10 box. The black curve is the distribution obtained in SURFS simulation in [22](see Fig. 5 there). The black dashed line indicate the zero line, where there is no dissociation.

two-peak distribution, dividing the halos into clear two groups. The distance is a well-accepted criterion in the literature, but this criterion is easier to compute in simulations.

Besides, I also managed to develop the code to do the following:

- Build a KDTree of the particles in the snapshots and find the density of particles given a center and a radius, as shown in Fig 1,
- Select the particles of a certain halos or all the halos in a cluster from the reduced snapshot, or save then in a single snapshot. This piece of code prepares for further analysis the visualization with gadgetry.
- Smoothing the snapshots based on density, as shown in Fig 7. In low density region, the diffuse X ray emitting gas is represented by discrete particles. Around halo center region, the manually added AGN feedback model can lead to over heating of some particles. Smoothing can address these 2 problem and results in more realistic X ray profiles. Besides, X ray surveys usually have limited angular resolutions, smoothing can also help mimic the X ray profiles in observations
- Find the brightest pixel in a png file of a snapshot, this is to determine the luminosity center and then calculate the center offset
- Find the halos that are cross-boundary. FLAMINGO uses the periodical boundary condition at the edge of the boxes, where some halos lie there and are cut into half and need special care.

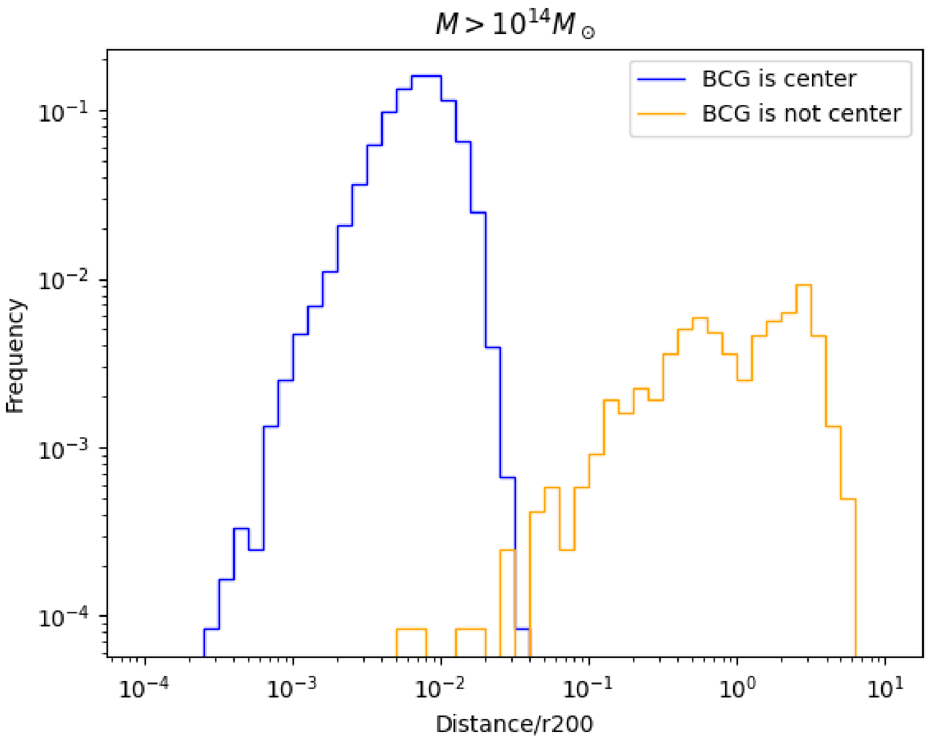

Figure 6: The distribution of distances between the Brightest Cluster Galaxy(BCG), defined as the most massive galaxies in stellar mass, and the halo potential centers, for 10^4 $M > 10^{14} M_\odot$ halos in FLAMINGO L1_m10 box. The most massive galaxy is defined by summing up all star particles within $100kpc$ from the star center of mass. The halo where the BCG is at the center(blue) and is not(orange) is divided clearly into two groups. For a more discussions, see Fig 4 in [36].

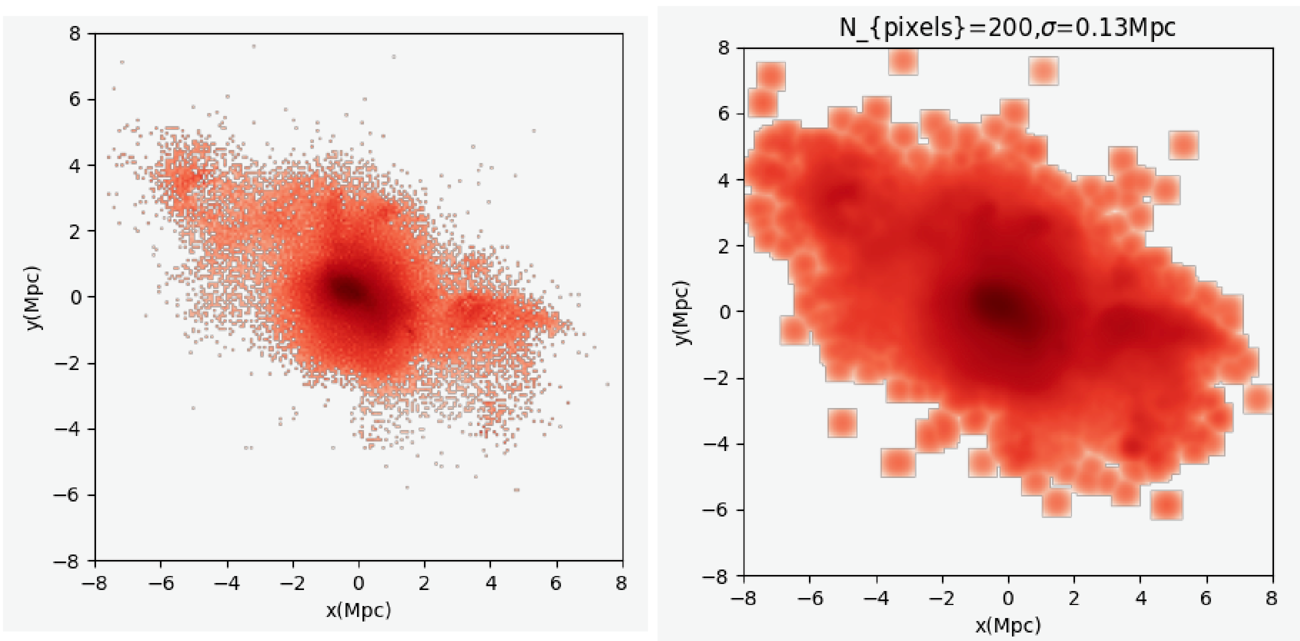

Figure 7: The X-ray luminosity distribution of a galaxy cluster in FLAMINGO before(left) and after(right) smoothing. The smoothing filter size is determined assuming the cluster is placed at $z = 1.5$ and observed by eROSITA[24]

Skills	Current level	Desired level
Code Development	Competent	Proficient
Scientific Writing	Basic	Proficient
Oral Presentation	Competent	Proficient
Understand Literature in depth	Competent	Proficient
Make academic decisions independently	Basic	Proficient

Table 2: The estimated timeline of

Section C: Supervision

- **Primary Supervisor: Prof. Danail Obresckhow(50%)**

  Danail will guide me during the project and help me understand the results, with his expertise on galaxy evolution. I will meet Danail weekly to discuss what I have done and how we can understand them, as well as any scientific or technical questions we find interesting .

- **Co Supervisor: A/Prof. Aaron Ludlow(30%)**

  Aaron will also give me advice about how to understand our results, with his expertise on simulations and galaxy assembly histories. He will also help me analyze data from the simulations. Aaron and I will meet weekly on the same meeting with Danail.

- **Co Supervisor: Prof. Simon Driver(20%)** Simon will help me understand the current state and capabilities of X-ray and optical surveys, such as eROSITA and WEAVES. He will also assist in comparing our simulation results with observations. Simon and I will meet once a month, or more often if necessary.

Budgets and Approvals

I obtained a Mac laptop for data analysis in the ICRAR office. The budget came from UWA ($1500) and ICRAR ($1200). I will use travel funding from the GRS ($1500 per year) to attend the Harley Wood School in the first year and the ASA annual conference in the second and third years, to stay informed about the work of other scientists in Australia. The estimated budget is $1200 per year. If I need to attend an international conference (estimated budget of $3000), I will skip the ASA conference that year and seek extra funding from ICRAR. The code I use is all public on GitHub, and no extra approval is currently needed from the GRS.

Research Plan

Fig 8 shows what I plan to do across my three projects and the rough timeline for completing them. It also includes my plans for travel and conferences, plus the presentations I'll give each year as progress check-ins for my PhD. I'm giving myself extra time on the first project since I need to build the pipelines that all three projects will use. Table 2 lists the skills I would like to acquire during my PhD.

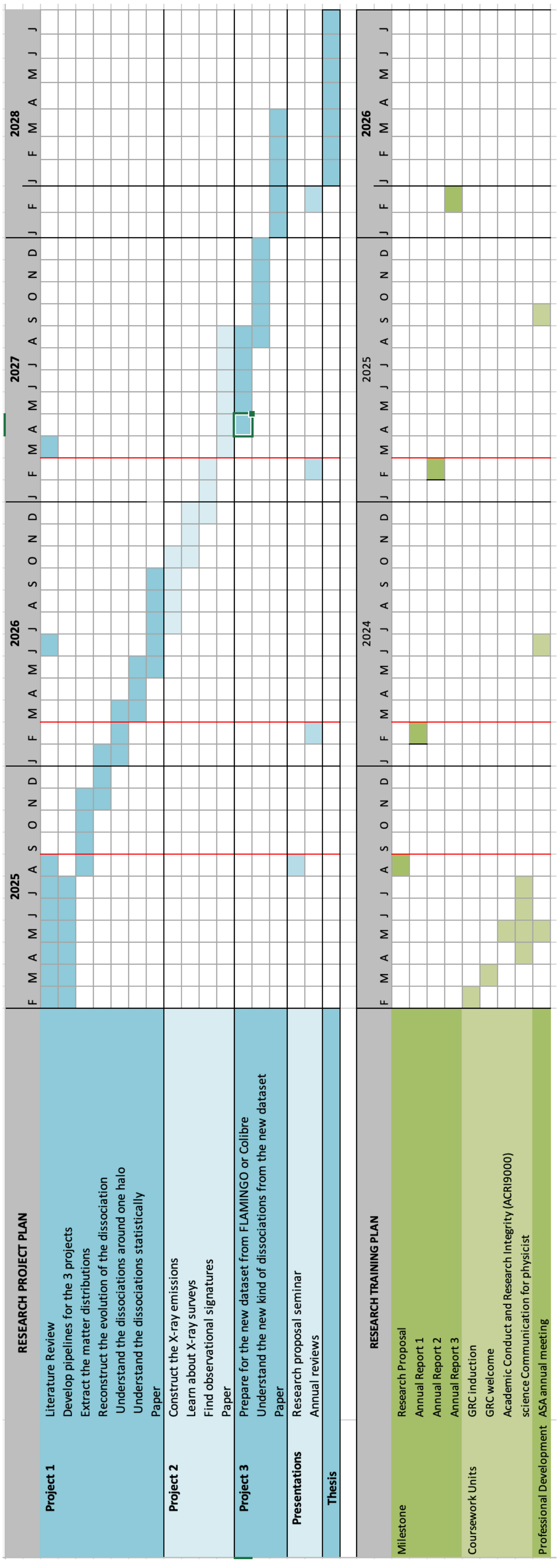

Figure 8: The list of activities involved in my research and the estimated timeline

References

[1] Daichi Akino et al. HSC-XXL: Baryon budget of the 136 XXL groups and clusters. February 2022.

[2] Heinz Andernach and Fritz Zwicky. English and Spanish Translation of Zwicky's (1933) The Redshift of Extragalactic Nebulae, November 2017. arXiv:1711.01693 [astro-ph].

[3] Yannick M. Bahe and Pascale Jablonka. Galaxies in the simulated cosmic web: I. Filament identification and their properties. February 2025.

[4] Josh Borrow and Alexei Borrisov. swiftsimio: A python library for reading swift data. 2020.

[5] Joey Braspenning et al. The FLAMINGO Project: Galaxy clusters in comparison to X-ray observations, June 2024.

[6] Ángel Chandro-Gómez et al. On the accuracy of dark matter halo merger trees and the consequences for semi-analytic models of galaxy formation. April 2025. arXiv:2501.07677 [astro-ph].

[7] Douglas Clowe et al. Weak lensing mass reconstruction of the interacting cluster 1E0657-558: Direct evidence for the existence of dark matter. April 2004. arXiv:astro-ph/0312273.

[8] Shaun Cole et al. The 2dF Galaxy Redshift Survey: power-spectrum analysis of the final data set and cosmological implications. September 2005.

[9] Euclid Collaboration. Euclid Definition Study Report, October 2011.

[10] Planck Collaboration. Planck 2015 results. XIII. Cosmological parameters. October 2016. arXiv:1502.01589 [astro-ph].

[11] Ajay Dev et al. The baryon census and the mass-density of stars, neutral gas, and hot gas as a function of halo mass. December 2024.

[12] Klaus Dolag et al. Encyclopedia Magneticum: Scaling Relations from Cosmic Dawn to Present Day. April 2025.

[13] Simon Driver et al. in preparation.

[14] Harald Ebeling et al. Jellyfish: Evidence of extreme ram-pressure stripping in massive galaxy clusters. January 2014. arXiv:1312.6135 [astro-ph].

[15] D. J. Eisenstein. Dark energy and cosmic sound [review article]. November 2005.

[16] Pascal Elahi et al. SURFS: Riding the waves with Synthetic UniveRses For Surveys. 2018.

[17] Roan Haggar et al. Reconsidering the dynamical states of galaxy clusters using PCA and UMAP. July 2024.

[18] Jiaxin Han et al. HBT+: an improved code for finding subhalos and building merger trees in cosmological simulations. February 2018. arXiv:1708.03646 [astro-ph].

[19] Andrey V Kravtsov and Stefano Borgani. Formation of Galaxy Clusters.

[20] Jie Li et al. Spin transfer from dark matter to gas during halo formation. September 2022.

[21] Ian G. McCarthy et al. The BAHAMAS project: calibrated hydrodynamical simulations for large-scale structure cosmology. March 2017.

[22] William McDonald et al. Dissociation of dark matter and gas in cosmic large-scale structure. September 2022.

[23] Kyle A. Oman. swiftgalaxy, May 2025.

[24] P Predehl et al. The eROSITA X-ray telescope on SRG. March 2021.

[25] Scott W. Randall et al. Constraints on the Self-Interaction Cross Section of Dark Matter from Numerical Simulations of the Merging Galaxy Cluster 1E 0657-56. June 2008.

[26] Craig L. Sarazin. The Physics of Cluster Mergers. June 2002.

[27] Joop Schaye et al. in preparation.

[28] Joop Schaye et al. The EAGLE project: Simulating the evolution and assembly of galaxies and their environments. January 2015. arXiv:1407.7040 [astro-ph].

[29] Joop Schaye et al. The FLAMINGO project: cosmological hydrodynamical simulations for large-scale structure and galaxy cluster surveys. October 2023. arXiv:2306.04024 [astro-ph].

[30] R. Seppi et al. Offset between X-ray and optical centers in clusters of galaxies: Connecting eROSITA data with simulations. March 2023.

[31] H. Tanimura et al. X-ray emission from cosmic web filaments in SRG/eROSITA data. November 2022.

[32] Marcel P van Daalen et al. The effects of galaxy formation on the matter power spectrum: a challenge for precision cosmology. August 2011.

[33] Mark Vogelsberger et al. Introducing the Illustris Project: simulating the coevolution of dark and visible matter in the Universe. October 2014.

[34] S. D. M. White and M. J. Rees. Core condensation in heavy halos: a two-stage theory for galaxy formation and clustering. May 1978.

[35] Linfeng Xiao et al. Forecasts on interacting dark energy from the 21-cm angular power spectrum with BINGO and SKA observations. February 2022.

[36] A. Zenteno et al. The dynamical state of eROSITA clusters and its impact on the brightest cluster galaxy luminosity. June 2025.