UNIVERSITÀ DEGLI STUDI DI ROMA TOR VERGATA
MACROAREA DI SCIENZE MATEMATICHE, FISICHE E NATURALI


CORSO DI STUDIO IN
**FISICA - Curriculum Erasmus Mundus MASS**

TESI DI LAUREA IN
**Cosmologia**

TITOLO

# Cosmic Filament Reconstruction: Insights from Cosmological Simulations

**Relatore:**

**Prof. Marina Migliaccio**
Department of Physics
Tor Vergata University of Rome

**Correlatore:**
**Dr. Javier Carrón Duque**
Institute of Theoretical Physics
Madrid, Spain

**Dr. Simon Prunet**
Université Cote d'Azur

**Laureando:**

**Gauri Shankar Hari**
Matricola: 0325652

Anno Accademico 2023/2024


**Erasmus Mundus Joint Master**
**in Astrophysics and Space Science**

**Master Thesis**

# Cosmic Filament Reconstruction: Insights from Cosmological Simulations

**Supervisors:**

**Prof. Marina Migliaccio**
Department of Physics
Tor Vergata University of Rome

**Dr. Javier Carrón Duque**
Institute of Theoretical Physics
Madrid, Spain

**Dr. Simon Prunet**
Université Cote d'Azur

**Author:**

**Gauri Shankar Hari**
Matricola: 0325652

Academic Year 2023/2024

This master thesis is submitted in partial fulfilment of the requirements for the degree *Fisica - Curriculum Erasmus Mundus* as part of a multiple degree awarded in the framework of the Erasmus Mundus Joint Master in Astrophysics and Space Science −− MASS jointly delivered by a Consortium of four Universities: Tor Vergata University of Rome, University of Belgrade, University of Bremen, and Université Cote d'Azur, regulated by the MASS Consortium Agreement and funded by the EU under the call ERASMUS-EDU-2021-PEX-EMJM-MOB.

*To Achan and Amma*

"*In nature we find our solace,*
*May this endless pursuit of the*
*mankind to understand the*
*nature shall bring us closer and*
*make us more humane.*"

# Abstract

Cosmological N-body simulations have played a pivotal role in the understanding of the Universe, and in testing different cosmological models with different parameters. Moreover, with the current and upcoming large-scale galaxy surveys, which are expected to provide large amounts of data, will allow us to characterise the large scale structure (LSS) of the Universe with an unprecedented precision, allowing for novel tests of the cosmological model. Both simulations and observations show that matter in the Universe is distributed in a complex multiscale network called *cosmic web*, with long *cosmic filaments* being its most defining structure. This motivates the need to study the LSS using filament-finding algorithms to detect and reconstruct filaments, as well as to identify other structures of the LSS. Numerous techniques have been developed so far, with varying applicability and performance, depending also on the analysed data. In this work, we adopt a filament finding algorithm known as the Subspace Constrained Mean Shift (SCMS) algorithm, as it is well understood mathematically and allows us to estimate the uncertainty in filament reconstruction. We apply this algorithm to detect and reconstruct filaments in 2D maps of the distribution of dark matter (DM) halos at different redshift slices between the range $0.0968 < z < 2.0842$ extracted from the DEMNUni N-body simulations with $\Lambda$CDM model and zero neutrino mass. The SCMS method has a free parameter, kernel, and filament reconstruction depends on our choice of kernel size and shape used to smooth the map to regularise the field. We used Gaussian kernels of different angular scales and a fixed physical scale across different redshift slices to understand how the kernel size affects the filament reconstruction and analysed the results using various statistical methods. We also introduce a novel method called the *Two-step Filament Reconstruction* (2sFR) method based on the SCMS algorithm but using two distinct kernels, one larger to initiate the convergence of the filament search points on a map and then using these already moved points for better reconstruction of filaments on the same map smoothed with the desired smaller physical scale.

**Keywords**: *N-body simulations, Large Scale Structure, Cosmic Filaments, Filament Finders, SCMS, DEMNUni, neutrino mass.*

# Contents

# Acknowledgements

With this thesis work, a beautiful journey that started two years ago is coming to an end. The pursuit of knowledge, though tremendously overwhelming and tiring at times, does provide some comfort when you reach the end, sometimes with more questions than answers, but with teamwork and encouragement that are more than enough for a curious mind to keep going forward. Throughout the course of this Erasmus Mundus degree programme, I would like to say that I have learnt some science and a lot about myself and I am grateful to countless people who have supported me throughout this journey and made this work possible.

First of all, I would like to thank my supervisors from the bottom of my heart, Prof. Marina Migliaccio (UNITOV) and Dr. Javier Carrón Duque (IFT), for guiding me with great kindness and patience for the entire duration of the internship and thesis. I am deeply grateful for their expertise, which has deepened my interest in the subject, and for their valuable time reviewing this work.

I also thank Dr. Simon Prunet (OCA) for his collaboration and helpful insights.

I extend my gratitude to the MASS consortium for this wonderful opportunity and sincerely appreciate the efforts of the MASS Project Office, especially Dr. Adalia Caroli, for all the help, which made life a lot easier during the course of this programme.

I would also like to thank my parents, brother, friends and teachers who have supported and motivated me throughout my life. None of this would have been possible without their love, kindness and advices.

Last but not least, I would like to thank myself for hanging in there, not giving up, and keeping alive that dream of the 10 year old boy of becoming an astrophysicist, even when going through the toughest times.

# Introduction

Cosmological probes like the Cosmic Microwave Background (CMB) and Large Scale Structures (LSS) provide an opportunity to test the cosmological models and constrain its parameters. With the advent of space missions like the Wilkinson Microwave Anisotropy Probe (WMAP) and Planck to map the full microwave sky, there have been significant improvements in understanding the early universe. Similarly, the LSS can tell us more about the structure formation and evolution in the later stages of the Universe. Large-scale galaxy surveys along with large volume and high-resolution cosmological N-body simulations provide a promising probe of LSS and can be used to verify the estimated cosmological parameters from other cosmological probes.

Matter in the Universe is distributed in a complex intricately interwoven weblike structure on large scales, called the cosmic web. This cosmic web can be classified into its structural components on the basis of its morphological features and properties: halos, filaments, sheets and voids. Highly dense halos and empty voids have been extensively studied and understood in the literature. The long tubular cosmic filaments that act as a highway for matter and gas to fall into halos, possibly aiding star formation, form the crux of this thesis work. Like formation and evolution of structures on small scales, their evolution is also driven by gravitational collapse. Their growth is affected by the accelerated expansion driven by dark energy (DE), the presence of massive neutrinos, and dark matter (DM) density. Thus, they represent a promising cosmological probe that can help us constrain cosmological parameters and improve our understanding of the cosmological model.

Studying the LSS has its own challenges; the large multi-scale morphology of these structures demands statistics and techniques beyond the usual correlation functions and matter power spectrum. These statistics do not give any information on the anisotropy of the LSS. Thus, we require methods which can factor in these anisotropies and the complex nature of the LSS. The filament finding algorithms are used to detect and reconstruct filaments and other cosmic web substructures from the maps of tracers of matter distribution such as galaxies and clusters from large-scale galaxy surveys or dark matter (DM) halos from simulations. In this thesis, we focus on the cosmic filament reconstruction on state-of-the-art N-body cosmological simulations to study their properties and morphological features. Various algorithms based on different mathematical techniques and catering to different goals have been developed to identify and reconstruct filaments in different tracer maps.

We use the Subspace Constrained Mean Shift (SCMS) algorithm based on the ridge formalism, where ridges are one-dimensional curves tracing the distribution of

matter. Unlike many other existing filament finding algorithms, uncertainty estimation in filament reconstruction using SCMS has been explored in the literature. The filament detection and reconstruction using SCMS requires the tracer field to be regularised using a kernel of our choice, which is Gaussian in our case. Therefore, the filament reconstruction is subjective to the effects of kernel scales used to smoothen the distribution map of galaxies or halos. We perform filament reconstruction on the N-body simulation called Dark Energy and Massive Neutrino Universe (DEMNUni) simulations with a larger aim to study the effects of various neutrino masses on the structure formation and evolution.

Current and upcoming large-scale galaxy surveys such as DESI, Euclid, Vera Rubin-LSST, and SKA will provide us with huge amounts of data, which will help us understand structure formation and evolution. Thus, it is very important to have an optimised filament finding algorithm to recover filament properties without any bias arising from the algorithm itself and create filament catalogues. With this goal in mind, we explore and study the dependence of filament reconstruction on the smoothing kernel size and suggest some improvements.

In this thesis, we start by providing an overview of the cosmological framework in which we work in Chapter 1. We explain the ΛCDM model, the cosmological principle, and the dynamics of expanding Universe. We then briefly explain the timeline of the Universe and parameters of ΛCDM cosmology. We then discuss the two imporatant cosmological probes: Cosmic Microwave Background (CMB) and Large Scale Structure (LSS) with much more emphasis on the formation and evolution of LSS, and its statistics. We finish the chapter with a detailed discussion of the cosmic web and its components, which form the crux of the thesis.

In Chapter 2, we briefly explain the N-body simulations and their importance in studying and testing various cosmological models, and introduce and explain our choice of simulation, DEMNUni, for this work. Then we briefly explain the main software and data analysis tools used in this work.

We introduce and briefly discuss various filament finders or reconstruction algorithms based on various mathematical approaches and motivate our choice of using the SCMS algorithm based on the ridge formalism in Chapter 3. We explain the SCMS algorithm, the uncertainty estimation in the framework of SCMS. We then discuss how varying angular scales for smoothing affects filament reconstruction at different redshifts, the use of constant physical scale as a smoothing scale, and propose the Two-step Filament Reconstruction (2sFR) method as an improvement to the SCMS algorithm for filament reconstruction using small physical smoothing kernels. We end the chapter with two statistical methods, the stacking method and distance estimation between halos and reconstructed filaments, to analyse the filament reconstruction.

We discuss in detail the main results of our work, using the SCMS algorithm to reconstruct filaments in DEMNUni simulations in Chapter 4. We discuss how filament reconstruction performs at different redshifts for varying angular smoothing kernels with uncertainty estimation, staking method, and estimating the distance from halos to filaments. We also compare the results of the SCMS method and the 2sFR method for the same physical smoothing kernel.

In Chapter 5, we conclude our findings with summarised main results and suggestions for improvement, and outline the future scope of this thesis work.

After the main body of the text, Bibliography and Appendix are given. Then, a list of acronyms, figures, and tables are provided, respectively, at the end for the reader's convenience.

# Chapter 1

# ΛCDM Cosmology

The current standard cosmological model is the most widely accepted and successful model of the Universe, explaining phenomena of widely varying scales with great accuracy for most of its lifetime. The model assumes General Relativity (GR) with a cosmological constant $\Lambda$, representing dark energy (DE), believed to be the driving factor behind the rapid expansion of the Universe, and consists of matter mostly in the form of Cold Dark Matter (CDM), a non-relativistic, non-baryonic, non-luminous matter that only interacts gravitationally and with some amount of baryonic matter, neutrinos, and photons. This model assumes that the Universe is homogeneous and isotropic on large scales, and to a very good approximation. This means that there are no preferred locations or directions in our Universe, i.e. it is the same everywhere and in all directions. This important idea is now recognised as a principle in cosmology, popularly known as the *Cosmological Principle* (see, Hawking and Ellis 1973; Ellis and Harrison 1974). There is observational evidence supporting this claim and also suggesting some deviations (see Park et al. 2017; Alam et al. 2021; Kim et al. 2022; Kumar Aluri et al. 2023;).

The ΛCDM model is successful in modelling a Universe that accounts for the cosmic expansion history. Its predictions on cosmic microwave background radiation, and its anisotropies agree with observations with great precision. It explains the formation and evolution of structures at large scales from primordial density perturbations and the periodic acoustic oscillation observed in the galaxy clustering data, known as the Baryon Acoustic Oscillations (BAO). Despite its success in explaining most of the observable features of the Universe, it is still limited or does not give definite answers to certain questions, such as the nature of dark matter and dark energy, and the inconsistency in the estimations of the Hubble parameter from various probes, which gives the expansion rate of the Universe. This discrepancy in cosmology is known as the Hubble tension. The ΛCDM model predicts way more small dwarf galaxies than observed, and this is known as the small-scale structure problem.

The horizon and curvature problems, arising from the homogeneity and isotropy of the CMB temperature map on large scales and the flatness of the Universe respectively, have been resolved to a great extent by the cosmic inflation theory (Guth 1981), suggesting that the Universe expanded enormously since its formation which made the CMB uniform and isotropic on large scales and the Universe to be flat.

In this chapter, we will briefly discuss the dynamics of the expansion, and the Friedmann equations and density parameters governing this expansion. Then we briefly describe the timeline of our Universe and introduce the important parameters of the ΛCDM model. We then look at two important cosmological probes: Cosmological Microwave Background (CMB) and Large-Scale Structure (LSS), with more emphasis on the structure formation and evolution, and its statistics. Finally, we close this chapter with a discussion on the cosmic web and its structural components. A detailed treatment of these topics can be found in most standard textbooks for cosmology, for example: Dodelson (2003); Weinberg (2008); Vittorio (2017); etc.

## 1.1 Dynamics of Expanding Universe

From GR, we know that the matter and energy distribution in the Universe determines its geometry, which in turn determines the dynamics of matter and energy. The Einstein's field equations can be written as:

$$G_{\mu\nu} + \Lambda g_{\mu\nu} = \frac{8\pi G}{c^2} T_{\mu\nu} \tag{1.1}$$

where $G_{\mu\nu}$ is the Einstein tensor, $g_{\mu\nu}$ is the metric tensor, $G$ is the gravitational constant, $c$ is the speed of light in vacuum and $T_{\mu\nu}$ is the stress energy tensor.

Invoking the cosmological principle, at large scales, matter and energy could be said to be homogeneously distributed. Considering matter and energy as perfect fluids, we can write the following equation of state.

$$p = w\rho c^2 \tag{1.2}$$

$$\rho \propto a^{-3-3w} \tag{1.3}$$

where $p$ is the pressure of the fluid, $\rho$ is the density, $w$ is a time-independent parameter depending on the fluid considered, and $c$ is the speed of light in vacuum. The following fluids are considered in the ΛCDM model:

- **Matter:** It comprises all non-relativistic particles, which are, the baryonic matter i.e., stars, planets, brown dwarfs, dust, gas, black holes, and most of the luminous matter in the Universe and the non-baryonic cold dark matter (CDM), they are non-luminous and interact only gravitationally. The pressure and density of the matter fluid can be written as (and $c = 1$);

$$p = 0 \tag{1.4}$$

$$\rho \propto a^{-3} \tag{1.5}$$

- **Radiation:** Radiations consists of all relativistic particles: photons and neutrinos. The pressure and density of the fluid is given by:

$$p = \frac{\rho}{3} \tag{1.6}$$

$$\rho \propto a^{-4} \tag{1.7}$$

- **Dark Energy:** In the ΛCDM model, the term Λ represents the dark energy, a fluid which drives the accelerated expansion of the Universe at large scales. It has a negative pressure, as given by:

$$p = -\rho \tag{1.8}$$

The metric of an expanding homogeneous and isotropic Universe is described by the Friedmann-Lemaître-Robertson-Walker metric or the FLRW metric, in short (Friedman 1922; Lemaître 1931; Robertson 1935; Walker 1937). It is given by:

$$-c^2 d\tau^2 = c^2 dt^2 + a^2(t) \left[ \frac{dr^2}{1 - kr^2} + r^2 d\Omega^2 \right] \tag{1.9}$$

where $a(t)$ is the scale factor that determines the evolution of space in time, $t$ is the time coordinate, $\tau$ is the proper time, $cd\tau$ is the proper distance, $d\Omega^2 = d\theta^2 + sin^2\theta d\phi^2$, and $k$ represents the curvature of the Universe with the following three possible values:

$$k = \begin{cases} +1, & \text{closed.} \\ 0, & \text{flat.} \\ -1, & \text{open.} \end{cases} \tag{1.10}$$

### 1.1.1 Friedmann Equations

To understand the dynamics of the expansion of the Universe, we must use the field equations with various fluids and our choice of the metric. The Einstein field equations applied to the FLRW metric provide us with the Friedmann equations. The first Friedmann equation can be written as;

$$H(t)^2 \equiv \frac{\dot{a}^2}{a^2} = \frac{8\pi G}{3}\rho + \frac{\Lambda c^2}{3} - \frac{kc^2}{a^2} \tag{1.11}$$

where $H(t) = \frac{\dot{a}(t)}{a(t)}$ is the Hubble parameter (the current expansion rate is given by the Hubble constant, $H_o \equiv H(t_o) = \dot{a}(t_o)$), $a$ and $\rho = \rho_m + \rho_r$ (density of matter and radiation) are functions of time, and $\dot{a}$ is the first derivative of the scale factor with respect to time, and the second Friedmann equation is given by:

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}(\rho + 3p) + \frac{\Lambda}{3} \tag{1.12}$$

It is possible to simplify these equations further by reexpressing them in terms of the density parameters of fluids for easy interpretation, which we shall define as follows:

$$\Omega_m = \frac{8\pi G}{3H^2}\rho_m \equiv \frac{\rho_m}{\rho_o} \tag{1.13}$$

$$\Omega_r = \frac{8\pi G}{3H^2}\rho_r \equiv \frac{\rho_r}{\rho_o} \tag{1.14}$$

$$\Omega_\Lambda = \frac{\Lambda c^2}{3H^2} \tag{1.15}$$

$$\Omega_k = -\frac{kc^2}{H^2 a^2} \tag{1.16}$$

where $\rho_o = \frac{3H_o^2}{8\pi G} = (1.88 \times 10^{-29} h^2) g \cdot cm^{-3}$ is the critical density and $h = \frac{H_o}{100} \frac{kms^{-1}}{Mpc}$ is the normalised Hubble constant. Inserting these density parameters into the first Friedmann equation yields the following constraint.

$$\Omega_m + \Omega_r + \Omega_\Lambda + \Omega_k = 1 \tag{1.17}$$

$$\Omega_o + \Omega_\Lambda + \Omega_k = 1 \tag{1.18}$$

where $\Omega_o = \Omega_m + \Omega_r$.

It is known that the contribution of radiation to the total density in the present-day Universe is negligible and this will allow us to rewrite the geometrical quantity tied with the curvature of the Universe, $\Omega_k$ in terms of the other two observable density parameters. If $k = 0$, which seems to be the case, then $\Omega_k = 0$, all the contributions to the total density come from matter and dark energy.

### 1.1.2 Evolution of Density Parameters

Now, we will see how these density parameters scale with the scale factor, $a(t)$. We will examine the quantity $H^2\Omega_i$, where $\Omega_i$ represents the density parameter, in general.

In the case of matter, the energy density $\rho_m$ is directly proportional to the number of particles per unit volume. Although the number of particles remains the same, the volume increases with expansion by a factor of $a^3$. Therefore,

$$H^2\Omega_M \sim \frac{n}{V} \sim a^{-3} \tag{1.19}$$

In the case of radiation, the particles are relativistic; therefore, the energy density is proportional to the number of particles (photons or neutrinos) and their frequency and inversely proportional to the volume.

$$H^2\Omega_R \sim \frac{n\nu}{V} \sim a^{-4} \tag{1.20}$$

This could be interpreted as the along with the volume of the Universe scaling as $a^3$, the wavelength also scales as $a$, thus decreasing the energy density by a factor of $a^{-4}$ as the Universe expands.

For dark energy, the density parameter is independent of the scale factor in the $\Lambda$CDM model. Thus,

$$H^2\Omega_\Lambda \sim a^0 \tag{1.21}$$

The density parameter from the curvature of the Universe will be zero for a flat Universe. Otherwise, it scales as

$$H^2\Omega_k \sim a^{-2} \tag{1.22}$$

It is possible to estimate the value of these density parameters at any time, given that we know their value in the present day Universe by the following relations;

$$H(t)^2\Omega_m(t) = \frac{H_o^2\Omega_{m,o}}{a^3} \tag{1.23}$$

$$H(t)^2\Omega_r(t) = \frac{H_o^2\Omega_{r,o}}{a^4} \tag{1.24}$$

$$H(t)^2\Omega_\Lambda(t) = H_o^2\Omega_{\Lambda,o} \tag{1.25}$$

$$H(t)^2\Omega_k(t) = \frac{H_o^2\Omega_{k,o}}{a^2} \tag{1.26}$$

Figure 1.1 shows the energy density of different fluids as of today estimated from Planck measurements and during the time of recombination.

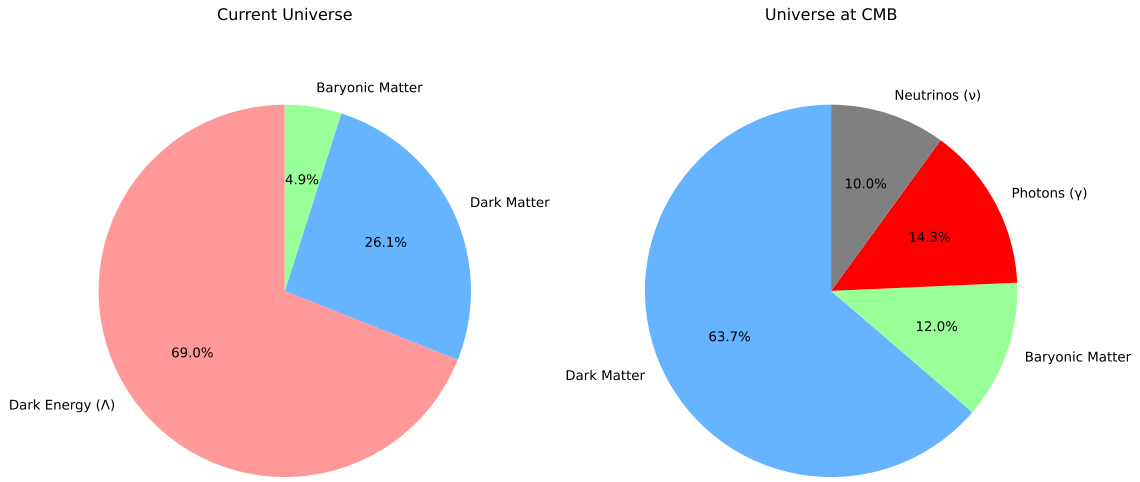



Figure 1.1: Pie Chart representing the contribution of different fluids to the total energy density. *Left*: Values estimated as of today from Planck measurements. *Right*: The values of density parameters estimated at the time of recombination

Now, we rewrite the Friedmann equations in terms of the density parameters we defined earlier. The first Friedmann equation becomes;

$$\frac{H(t)^2}{H_o^2} = \frac{\Omega_m}{a^3} + \frac{\Omega_r}{a^4} + \Omega_\Lambda + \frac{\Omega_k}{a^2} \tag{1.27}$$

It is much more clear from this equation that the evolution of the Universe is dependent on the content of the Universe at different times. The second Friedmann

equation can be rewritten as:

$$\frac{\ddot{a}}{a} = H^2(\Omega_\Lambda + -2\Omega_m - 4\Omega_r) \tag{1.28}$$

This describes whether the Universe expands or contracts. It should be noted that the only component that is driving the expansion is the dark energy when matter and radiation are trying to contract the Universe.

Friedmann equations allow us to describe the Universe with the scale factor and cosmic time. Similarly, we can use another measure to describe a particular moment in the history of the Universe, defined by the redshift ($z$). The expansion of the Universe causes the wavelength of the photon to be stretched; therefore, the photon loses energy and becomes redder as it travels. This can be expressed mathematically as;

$$z = \frac{\lambda_o - \lambda_e}{\lambda_e} = \frac{\lambda_o}{\lambda_e} - 1 \tag{1.29}$$

$$1 + z = \frac{\lambda_o}{\lambda_e} \tag{1.30}$$

where $\lambda_o$ is the wavelength of the photon observed at time $t_o$ and $\lambda_e$ is the wavelength of the photon emitted at time $t_e$. The ratio of the wavelengths is also equal to the ratio of the scale factor as observed today to the scale factor at the time the photon was emitted:

$$1 + z = \frac{\lambda_o}{\lambda_e} = \frac{a(t_o)}{a(t_e)} \tag{1.31}$$

## 1.2 Timeline of our Universe

In the previous sections, we discussed the equations that describe the evolution and content of the Universe. If we rewind the scale factor $a$, we can study the size of the Universe at various stages of its lifetime. This information along with the known laws of physics, such as the standard model of particles, allows us to paint a picture of the early Universe and its contents. The ΛCDM model is successful in describing the Universe as young as $10^{-32}s$. Beyond that, the known laws of physics may not work as we know it today. There have been many attempts to describe the early stages of the Universe, but these ideas and theories are beyond the capacity of mankind today to test or disprove it. When extrapolating the model beyond the known physics, i.e. $a = 0$, the Universe collapses into an initial singularity of infinite density, known as the *Big Bang*.

We do not clearly know what happens in the first $10^{-43}s$, denoted by *Planck time* and the *Planck scale*. Then, until the first $10^{-32}s$, it is proposed that the fundamental forces of the standard model are still unified except gravity according to the Grand Unified Theory (GUT, Georgi and Glashow 1974, De Boer 1994, Workman et al. 2022), and the Universe undergoes a rapid expansion due to cosmic inflation. It still remains the best candidate for explaining this period of time in the early Universe

and is yet to be proven or disproved. Now, we briefly look into various evolutionary phases of the Universe leading up to the present day and its possible future according to the ΛCDM model. The Figure 1.2 shows the timeline of the Universe with major epochs. The tabular summary of different epochs of the Universe with estimated time of the epoch, the redshift range, and a short description are provided in Table 1.1 and Table 1.2

- **Early Universe**:
  - **Planck Epoch**: The first $10^{-43}s$ marks the limit of known physics, defined by the Planck time and scale. It is proposed that the Universe was dominated by the quantum effects of gravity but we are yet to have a successful quantum theory of gravity
  - **Grand Unification Epoch (GUT Epoch)**: The fundamental forces of the Universe are still unified except gravity until $t = 10^{-36}s$.
  - **Inflationary Epoch**: The Universe expands enormously by at least a factor of $10^{26}$. The expansion cools the Universe, and the strong interaction decouples from the electroweak interaction. This is expected to be lasted till $t = 10^{-32}s$.
- **Radiation dominated era ($10^{-32}$ seconds - 47 Kyr)**: The energy density of the massless and near massless relativistic components like photons and neutrinos dominates both the matter and dark energy.
  - **Electroweak Epoch**: After the end of inflation, the inflaton field decays into particles, matter starts to exist. The weak and electromagnetic forces are still unified. The duration is believed to be from $10^{-32}s$ to $10^{-12}s$
  - **Quark Epoch**: The weak and electromagnetic forces separate as the Universe cools further down during time, $10^{-12}s$ to $10^{-6}s$. All the forces have separated and have taken the form as they have it today. The Universe is filled with a hot plasma of quarks and gluons, called the *quark-gluon plasma*.
  - **Hadron Epoch**: The temperature of the Universe has gone down enough by the first second for the quarks to form protons, neutrons, and other hadrons. The particle-antiparticle annihilation ends with mostly hadrons remaining.
  - **Lepton Epoch**: The Universe is mostly filled with radiation and hadrons, leptons, and anti-leptons. In the first $10s$, similarly to what happened with hadrons, there is an excess of leptons (mostly electrons) after the mutual annihilation process.
  - **Neutrino Decoupling**: At around the first second, the neutrinos and anti-neutrinos stop interacting with matter, and the Universe became transparent to neutrinos.

– **Photon Epoch**: After the annihilation of antiparticles and neutrino decoupling, the plasma is dominated by photons and the baryonic component is composed of neutrons, protons, and electrons. This marks the beginning of the *photon epoch*, which will last until the recombination. Neutrons (which have a mean life of $\tau_n = 886s$) will begin to decay into protons by the $\beta - decay$ process. The plasma is abundant with the charged particles interacts heavily with photons and make plasma very opaque. Dark matter will begin to fall into the potential wells created by the photons and forms the first seeds of the structures.

– **Primordial Nucleosynthesis**: In the first $3 - 20$ minutes, the Universe is cooled to the point (temperature drops to $10^9 K$ and, along with the decrease in pressure) that the conditions become perfect for the beginning of atomic fusion, known as primordial nucleosynthesis. This effectively confines a fraction of neutrons in helium (He) nuclei before their decay, and the rest of the neutrons decay into protons. Other elements and heavier nuclei are also produced in smaller amounts. The nucleosynthesis stops after 20 minutes and the baryonic component of the plasma is a mix of the ionised hydrogen (H) and He with traces of other elements. The photons are still coupled with the baryons because of the large cross section, $\sigma_T$ of the charged particles and photons. The photons still dominate the plasma.

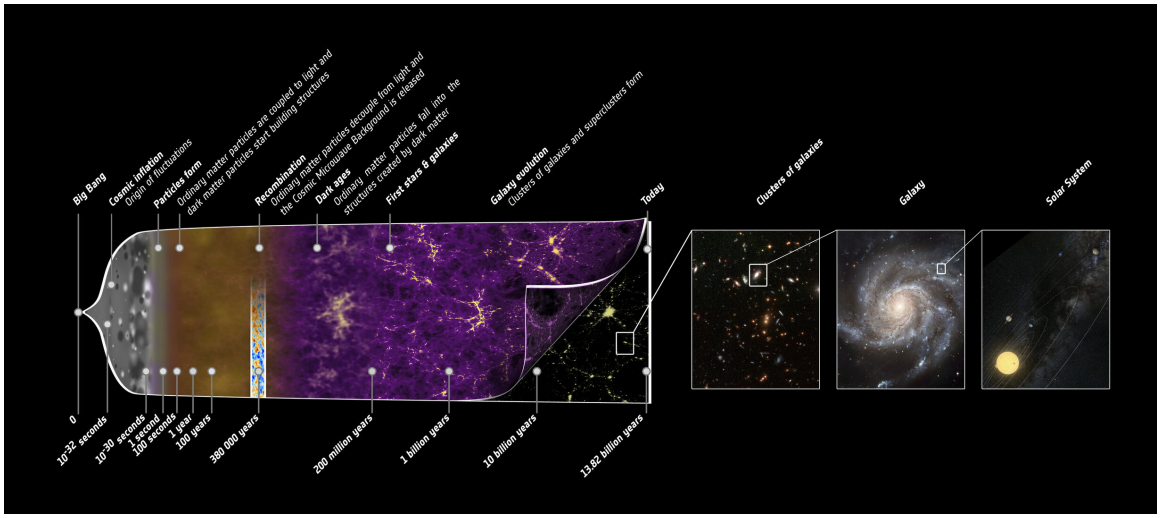



Figure 1.2: History of the Universe - depicting different evolutionary phases of the Universe. Image Credit: NASA

- **Matter dominated era (47 K yr - 9.8 Gyr)**: At around, $t = 47Kyr$ and $z \sim 3000$, the energy density of matter becomes equal to the energy density of radiation. This is known as the equivalence time. And after that the energy density of matter dominates both the energy density of radiation and dark energy, resulting in a decelerated expansion of the Universe. The baryonic matter and photons fall into the potential wells created by dark matter. The

Universe is still opaque, as the ionised baryonic matter is still coupled with the photons.

- **Recombination**: The photon epoch comes to an end as the temperature keeps dropping and, at a time around 380 000 years and redshift ($z \sim 1100$), the temperature drops below 3000 K. The photons will no longer be able to keep the H in plasma ionised and the protons and electrons combine to form neutral H. This process is known as *recombination*. The photons stop interacting with the baryonic matter because the neutral atom has a much smaller cross section with the photons. Photons decouple and freely propagate in the Universe. They stop falling into the potential wells and their contribution to structure formation ceases. These photons are currently observed as the background on the microwave bandwidth, known as cosmic microwave background (CMB), and their energy depends on the temperature of the plasma during recombination.
- **Dark Ages**: The Universe becomes transparent after recombination. Except for CMB photons and weak 21 cm emission from the plasma, there are no sources of light. This period is known as *dark ages*. The temperature continues to decrease, and the first structures start to form at the end of this era.
- **Stelliferous Era**: The formation of the first generation of stars marks the beginning of the Stelliferous era around $z \sim 20$. The earliest quasars and dwarf galaxies formed during this period around $z \sim 11$. These highly luminous sources reionise the neutral H in the Universe. This is known as *reionisation* and ends around $z \sim 6$. Hierarchical structure formation begins with small structures merging and growing under gravitational influence.
- **Galaxy Formation**: Baryonic matter along with dark matter distribution forms the Large Scale Structure (LSS), which forms an important cosmological probe, as we shall explain later in this chapter for studying the structure formation and evolution. The energy density of the matter decreases, but that of the dark energy remains constant.

- **Dark Energy dominated era (9.8 Gyr - 13.8 Gyr)**: The Dark Energy density dominates both the matter and radiation energy density, and the expansion of the Universe becomes accelerated. This continues without interruptions.

  - **Present**: The Universe is around 13.8 *Gyr* old. The surface of the last scattering marks the boundary of the observable Universe with CMB photons being the farthest observable. The Universe is very cold with an average temperature of $T = 2.7K$.
  - **Future**: The Universe will continue its accelerated expansion driven by dark energy. Most sources will move out of the observable boundary. Eventually, all the stars will run out of H and die, forming either white dwarfs

or black holes. The star formation in galaxies will end by around $10^{14}$ years, marking the transition from Stelliferous era to the Degenerate era. The Universe becomes extremely dark and cold.

| **Epoch/Era** | **Time** | **Redshift** | **Description** |
|---|---|---|---|
| | **Early Universe** | | |
| Planck Epoch | $t < 10^{-43}$ s | | Unknown Physics. All four fundamental forces may have been in a unified form. |
| Grand Unification Epoch | $10^{-43} - 10^{-36}$ s | | All forces except gravity remain unified. |
| Inflationary Epoch | $10^{-36} - 10^{-32}$ s | | Exponential expansion of spacetime. |
| | **Radiation-Dominated Era** | | |
| Electroweak Epoch | $10^{-32} - 10^{-12}$ s | | Electromagnetic and weak forces separate. Matter starts to exist. |
| Quark Epoch | $10^{-12} - 10^{-6}$ s | | Universe is filled with a hot quark-gluon plasma. |
| Hadron Epoch | $10^{-6} - 1$ s | | Quarks combine into form protons and neutrons. |
| Lepton Epoch | 1 s – 10 s | | Leptons (electrons, neutrinos) dominate interactions and Neutrino decouples at around the first second. |
| Photon Epoch | 10 s – $380,000$ yrs | | Universe is filled with hot plasma; photons scatter off electrons, preventing light from traveling freely. |

Table 1.1: Chronology of the Universe: Early Universe and Radiation-Dominated Era

| Epoch/Era | Time | Redshift | Description |
|---|---|---|---|
| **Matter-Dominated Era** | | | |
| Recombination | $380,000$ yrs | $z \approx 1100$ | Electrons combine with protons to form neutral atoms. Photons decouple and form the cosmic microwave background as we observe today. |
| Dark Ages | $380,000 - 200$ Myr | $1100 - 20$ | No stars yet; Universe is mostly neutral hydrogen. Tiny density perturbations continue to grow. |
| Reionization | $200 - 1000$ Myr | $20 - 6$ | The first stars and galaxies form, ionizing the neutral hydrogen. |
| Galaxy Formation | $1 - 6$ Gyr | $6 - 1$ | Large-scale structures with galaxies and clusters take shape. |
| **Dark Energy-Dominated Era** | | | |
| Accelerated Expansion | 9.8 Gyr – Present | $z < 1$ | Accelerated expansion of the Universe driven by dark energy. |
| Present Day | 13.8 Gyr | $z = 0$ | Universe keeps expanding with galaxies moving away from each other. |

Table 1.2: Chronology of the Universe: Matter and Dark Energy-Dominated Eras

## 1.3 Cosmological Parameters in ΛCDM

The standard ΛCDM model only requires six independent parameters to fully describe the cosmological model; as we shall explain briefly below and refer to Table 1.3 for an overview of these parameters with a short description. It should be noted that, in addition to the six independent parameters, the model has fixed and derived parameters. The physical parameters are much more technically parametrised to obtain the maximum information from the observations, see Planck Collaboration (2020) and Kosowsky et al. (2002). The six independent parameters are;

- **Physical Baryon Density, $\Omega_b h^2$:** In cosmology, all matter that interacts with electromagnetic radiation, can be clubbed together under the umbrella of baryonic matter. This includes gas clouds, dusts, stars, planets, black holes, etc. There used to be an observed discrepancy between the amount of baryonic matter directly detected in the present Universe and that measured from the CMB and primordial nucleosynthesis predictions. The observed baryonic matter only represented half of the predicted amount, widely known as the *missing baryon problem*. The problem is claimed to be resolved in recent years with missing baryons is in the form of hot gas outside the dark matter halos.

- **Physical Dark Matter Density, $\Omega_c h^2$:** The dark matter forms the abundant matter in the Universe, amounting to 85% of the total matter. Since, it does not interact with light, direct observations are not possible. It is only observed indirectly by its gravitational effects. Dark matter is not coupled with with the radiation at early Universe, thus the perturbations in dark matter can grow well before recombination. The dark matter requires to be cold instead of hot, i.e. non-relativistic at early Universe, otherwise perturbations below certain scale could be erased by diffusion. This would mean that the perturbations of galactic scale could be erased leaving only LSS to be present today. The cosmological CDM is perfectly compatible with the astrophysical CDM observed in clusters and galaxies using gravitational lensing, rotation curves etc, which solidifies its existence.

- **Age of the Universe, $t_o$:** This parameter tells us the time elapsed since the Big Bang. It has been measured by different observations; direct CMB observations by Planck interpreted with ΛCDM model yield a value $t_o = 13.8 \pm 0.020 Gyr$.

- **Scalar Spectral Index, $n_s$:** In the standard model of cosmology, the structures that we observe today have formed by gravitational accretion from initial density perturbations. These perturbations are expected to be formed during cosmic inflation (Guth 1981). According to inflation, perturbations are expected to have a power law spectrum like $k^{n_s-1}$, where k is the wave number representing the size of the perturbations. At, $n_s = 1$, the power spectrum is same for all scales, known as scale invariance. Different models of inflation predict a value slightly less than 1, compatible with both the observations of CMB and LSS.

- **Curvature Fluctuation Amplitude, $\Delta^2_R$:** The amplitude of the perturbations measured as the perturbations in the curvature and can be related to the the perturbations in density.

- **Reionisation Optical Depth, $\tau$:** As we have already explained, reionisation is the process by which the photons from the early stars and quasars reionised the neutral H. This increases the density of free electrons and therefore the probability for the photon to undergo a scattering process, thus reducing the transparency of the Universe. This affects the measurement of the small scale perturbations on the CMB. Although the evolution of this process is not fully understood yet, but it can be approximated by a single optical depth; and the process can be approximated at first order as an instantaneous event at $z_{reionsation}$.

These parameters can be used to compute other interesting physical parameters, such as the density parameters of different components of the Universe. The redshift and age of the Universe at different epochs like recombination and reionisation.

| Parameter | Symbol | Description |
|---|---|---|
| Physical Baryon Density | $\Omega_b h^2$ | Describes the baryon content of the Universe |
| Physical Dark Matter Density | $\Omega_c h^2$ | Describes the dark matter content of the Universe. |
| Age of the Universe | $t_o$ | Estimated to be $t_o = 13.8 Gyr$ with an uncertainty of 20 Myr. |
| Scalar Spectral Index | $n_S$ | Describes the variation of density perturbations with scale |
| Reionization Optical Depth | $\tau$ | Gives the extent of CMB photons scattered by free electrons from reionisation |
| Primordial Fluctuation Amplitude | $A_s$ | Normalised density perturbations |

Table 1.3: The six independent parameters of the ΛCDM cosmology along with their symbols and short description.

## 1.4 Cosmological Probes

There are various observational methods and tools to investigate and study the structure and evolution of our Universe. The cosmological probes like cosmic microwave background (CMB), large-scale structure (LSS), Type Ia supernovae, gravitational lensing, and galaxy clustering help us determine and constrain various cosmological parameters like the rate of expansion, density and composition of the Universe, etc. We will briefly explain two important cosmological probes, CMB and LSS, in the following sections.

### 1.4.1 Cosmic Microwave Background

The Cosmic Microwave Background (CMB) is the isotropic emission of photons from the surface of the last scattering during recombination. The discovery of the CMB is credited to Arno Penzias and Robert Woodrow Wilson (Penzias and Wilson 1965), as they observed constant background noise while testing a radio antenna for an experiment. Their discovery was later awarded the Nobel Prize in Physics in 1978. Since its discovery 60 years ago, various ground-based and satellite-based experiments have been designed to study the CMB spectrum, their angular distribution, and anisotropies.

The Universe was filled with a plasma mainly composed of H and He ions before the recombination epoch. As the temperature dropped during the recombination epoch, the photons could no longer keep the plasma ionised. This resulted in the formation of neutral atoms with protons and electrons combining together, decreasing the cross-section of ions with the photons and freeing them. The charged baryons and photons were in local thermodynamic equilibrium due to the constant interaction with each other. Thus, the photons follow a thermal distribution, called black body radiation, given by Planck's law:

$$B_v(T) = \frac{2hv^3}{c^2} \frac{1}{e^{\frac{hv}{k_B T}} - 1} \tag{1.32}$$

where $B_v(T)$ is the spectral radial density (rate of energy per solid angle per area normal to the propagation), $h$ is the Planck constant, $v$ is the frequency, $c$ is the speed of light in vacuum, $k_B$ is the Boltzmann constant, and $T$ is the temperature of the black body. This is in fact the most perfect blackbody that exists in the Universe. Although the black body nature of the CMB has become more evident ever since its discovery, it was only confirmed with greater accuracy by the Cosmic Background Explorer (COBE) in 1990 (Mather et al. 1994).

The black body spectral density given by Planck's law can be related to the number density of the photons N, per unit frequency, as:

$$\frac{dN}{dv} = \frac{4\pi}{c} \frac{B_v(T)}{hv} \tag{1.33}$$

This gives the photon distribution at the moment of the recombination and is affected by the expansion of the Universe. The expansion affects both the number

density and the frequency of the photons. We know that the number density evolves like $a^{-3}$ and the cosmological redshift changes the frequency as $a^{-1}$. The number density of photons observed in the current Universe is given by:

$$\frac{dN_o}{dv_o} = \frac{dN}{dv}|_{T_o=Ta} = \left[\frac{4\pi}{hc(va)}\right]\left[\frac{2h(va)^3}{c^2}\frac{1}{e^{\frac{h(va)}{k_B T a}} - 1}\right] \tag{1.34}$$

This means that what we observe currently as CMB photons is still a blackbody but with temperature corresponding to $T_o = a(t_e)T_e$, i.e. the expansion of the Universe only causes a shift in the apparent temperature of the CMB photons. The current temperature of the CMB black body is measured at $T_o = 2.735K$ with small fluctuations depending on the direction, of the order of $10^{-4}K$. The Figure 1.3 shows the CMB temperature map with anisotropies of the order $10^{-4}K$ as measured by the ESA's Planck mission and Figure 1.4 shows the energy spectrum of the CMB as measured by the COBE FIRAS experiment.

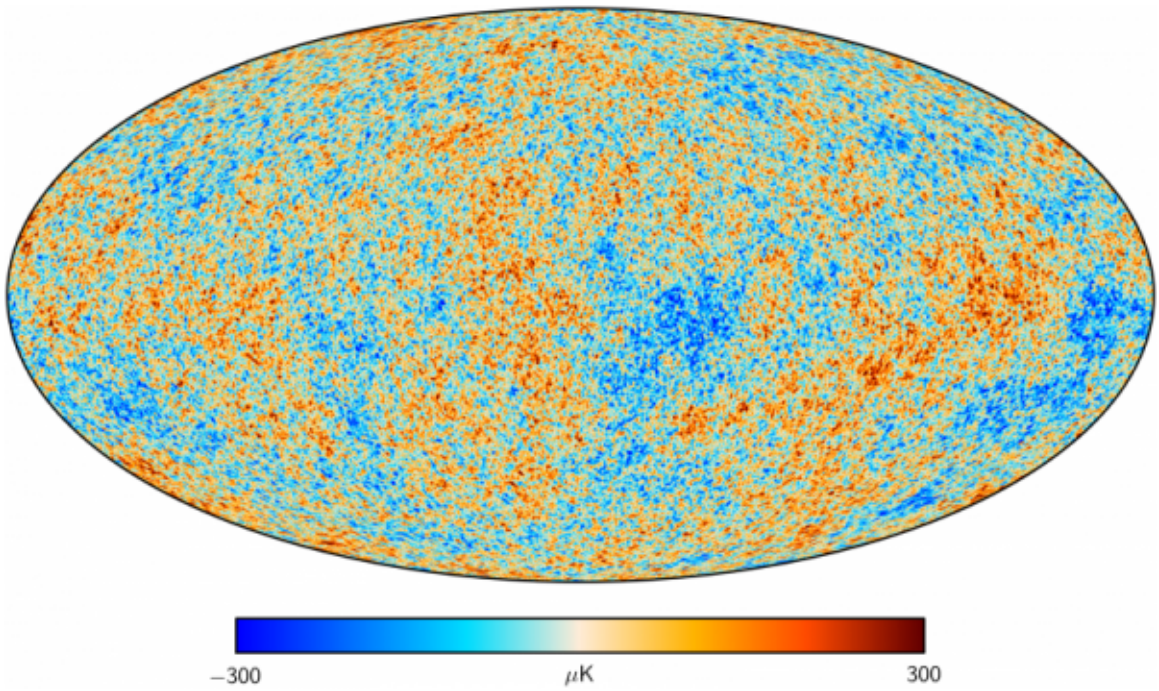



Figure 1.3: CMB Temperature map showing anisotropies obtained by ESA's Planck mission. Image Credit: ESA

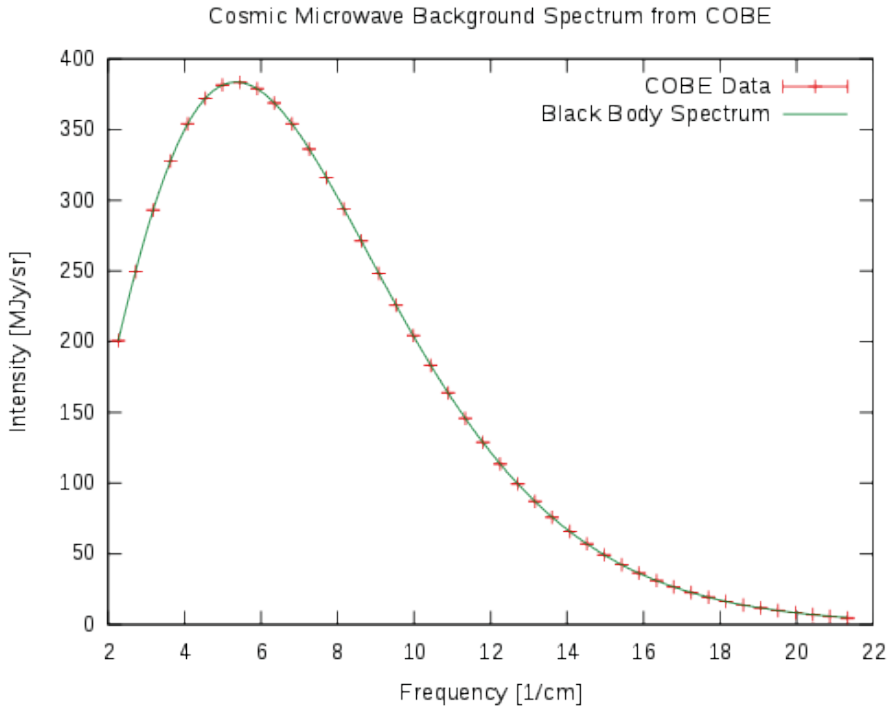



Figure 1.4: CMB energy spectrum as measured by COBE FIRAS instrument. Image Credit: NASA

### 1.4.2 Large Scale Structure

The LSS is the other crucial cosmological probe after the CMB and will be the focus of our thesis work. It is used to study the late Universe, test different cosmological models, and estimate some cosmological parameters. It refers to the distribution of matter in the observable Universe, and how it evolves and forms structures over time. Various phenomena at different scales dominate the evolution of the structures: the formation of galaxies and stars, AGNs, and feedback mechanisms can play a significant role in the evolution of their surroundings and make it complex to study at smaller scales. However, the evolution at larger scales is mainly driven by gravity and can be modelled with much greater accuracy. The LSS is observed at low redshift, i.e., a late Universe probe, compared to CMB, an early Universe probe. Thus, the information from both observations is expected to be compatible. We will now look briefly at the mathematical formalism for studying the formation and evolution of these structures.

#### Structure Formation and Evolution

To understand and formulate how structures form from the initial seeds of perturbations, we look at the dynamics of a self-gravitating fluid. We will consider a perfect fluid of density $\rho(\vec{r})$, pressure $p(\vec{r})$, and velocity $v(\vec{r})$. Let the position of the fluid be:

$$r(\vec{t}) = a(t)\vec{x} \tag{1.35}$$

where $\vec{r}$ and $\vec{x}$ are the proper and comoving cordinates of the fluid element respectively and the perfect fluid must obey the continuity, Euler, and Poisson equations as follows:

$$\frac{\partial \rho}{\partial t} + \vec{\nabla}_r(\rho v) = 0 \tag{1.36}$$

$$\frac{\partial \vec{v}}{\partial t} + (\vec{v}.\vec{\nabla}_r)\vec{v} + \frac{\vec{\nabla}_r p}{\rho} + \vec{\nabla}_r \Phi = 0 \tag{1.37}$$

$$\vec{\nabla}_r^2 \Phi - 4\pi G\rho = 0 \tag{1.38}$$

Since, our objective is to study the behaviour of the fluid as the Universe evolves. We need to work in comoving units and do the following coordinate transformations:

$$t \to t', \vec{x} = \frac{\vec{r}}{a(t)} \tag{1.39}$$

At $\vec{r} = constant$,

$$\frac{\partial}{\partial t} = \frac{dt'}{dt}\frac{\partial}{\partial t'} + \frac{dx^j}{dt}\frac{\partial}{\partial x^j} = \frac{\partial}{\partial t'} - \frac{\dot{a}}{a}\vec{x}.\vec{\nabla}_x \tag{1.40}$$

And at $t = constant$

$$\vec{\nabla}_r = \frac{1}{a}\vec{\nabla}_x \tag{1.41}$$

If we consider the fluid to be not exactly homogeneous and there are a few density fluctuations present at the linear order. then we can write the density of the fluid as:

$$\rho(\vec{x}, t) = \bar{\rho}(t)[1 + \delta(\vec{x}, t)] \tag{1.42}$$

where $\bar{\rho(t)}$ is the mean density of the background and $\delta(\vec{x}, t)$ is the fractional fluctuation of the fluid density at the comoving position $\vec{x}$, and the time $t$. The proper velocity of the fluid element can be written as:

$$\vec{v}(\vec{x}, t) = \frac{d\vec{r}(\vec{x}, t)}{dt} = \frac{da(t)}{dt}\vec{x} + a(t)\frac{d\vec{x}}{dt} \equiv \vec{v}_H + \vec{v}_p \tag{1.43}$$

Similarly, we can write pressure and potential with first-order linear expansion:

$$p(\vec{x}, t) = \bar{p}(t) + \delta p(\vec{x}, t) \tag{1.44}$$

$$\Phi(\vec{x}, t) = \bar{\Phi}(t) + \delta\Phi(\vec{x}, t) \tag{1.45}$$

We will now rewrite the continuity, Euler, and Poisson equations with the linear expanded terms and comoving coordinates as:

$$\frac{\partial}{\partial t}\delta(\vec{x}, t) + \frac{1}{a}\vec{\nabla}_x.\vec{v}_p = 0 \tag{1.46}$$

$$\frac{\partial \vec{v}_p}{\partial t} + \frac{\dot{a}}{a}\vec{v}_p + \frac{c_s^2}{a\rho}\vec{\nabla}_x\delta\rho + \frac{1}{a}\vec{\nabla}_x\delta\Phi = 0 \tag{1.47}$$

where $c_s = \sqrt{\frac{\delta p}{\delta \rho}}$, called the adiabatic speed of the sound.

$$\frac{1}{a^2}\vec{\nabla}_x^2 \delta\Phi(\vec{x},t) - 4\pi G\bar{\rho}\delta(\vec{x},t) = 0 \tag{1.48}$$

When combining the above three equations of the fluid, we get a second-order differential equation in the density contrast $\delta$.

$$\ddot{\delta} + 2\frac{\dot{a}}{a}\dot{\delta} = \frac{c_s^2}{a^2}\triangle_x \delta + 4\pi G\bar{\rho}\delta \tag{1.49}$$

This equation is known as the *gravitational instability equation*, describing the evolution of density fluctuations in a continuously expanding medium with pressure gradients and gravitational potential wells. In Fourier space, the equation becomes:

$$\ddot{\delta}_k + 2\frac{\dot{a}}{a}\dot{\delta}_k = -\frac{k^2 c_s^2}{a^2}\delta_k + 4\pi G\bar{\rho}\delta_k \tag{1.50}$$

$$\ddot{\delta}_k + 2H\dot{\delta}_k = \left[4\pi G\bar{\rho} - \frac{k^2 c_s^2}{a^2}\right]\delta_k \tag{1.51}$$

where $H = \frac{\dot{a}}{a}$ and it should be noted that perturbation at each scale is independent of each other.
We can further simplify the equation by introducing the proper wavenumber $k^{(p)} = \frac{k}{a(t)}$ and the Jeans proper wavenumber and wavelength, given by:

$$k_J^{(p)} = \sqrt{\frac{4\pi G\bar{\rho}}{c_s^2}} \tag{1.52}$$

$$\lambda_J = \frac{2\pi}{k_J^{(p)}} = \sqrt{\frac{\pi c_s^2}{G\bar{\rho}}} \tag{1.53}$$

The Eq. 1.51 becomes:

$$\ddot{\delta}_k + 2H\dot{\delta}_k = 4\pi G\bar{\rho}\delta_k\left[1 - \left(\frac{k^{(p)}}{k_J^{(p)}}\right)^2\right] \tag{1.54}$$

The Jeans proper wavenumber introduces a boundary above which the density perturbations will be gravitationally dominated ($k^{(p)} \gg k_J^{(p)}$) and below which it will be dominated by pressure ($k^{(p)} \ll k_J^{(p)}$). In the gravity-dominated regime, the perturbations corresponding to the longer wavelengths grow monotonically and plasma tends to collapse and forming structures, whereas in the pressure-dominated regime, the solutions to instability equations are of damped harmonic oscillations in nature, plasma tends to oscillate until it eventually stops. Let us briefly look into the evolution of these density fluctuations during different cosmological regimes.

- **Radiation-dominated Universe:** The most relevant scales fall under the pressure dominated regime in the radiation dominated Universe, the gravitational instability equation can be written as:

$$\ddot{\delta_k} + 2H\dot{\delta_k} + \frac{k^2 c_s^2}{a^2}\delta_k = 0 \tag{1.55}$$

  This equation describes a damped harmonic oscillator with period, $T = \frac{2\pi}{kc_s}$ and damping time, $\tau = \frac{1}{2H}$. The perturbations will oscillate until recombination.

- **Matter-dominated Universe:** After the photon-baryon decoupling during the recombination epoch, pressure becomes negligible, and the evolution perturbation is dominated by the gravity term. Assuming that the total energy density is dominated by matter and $\Omega_o = 1$, we know that $a \propto t^{2/3}, H \propto a^{-3/2}$, and $t = \frac{2}{3H}$. The instability equation can be written as:

$$\ddot{\delta}_k + \frac{4}{3t}\dot{\delta}_k = \frac{2}{3t^2}\delta_k \tag{1.56}$$

  The equation has power law solutions of the form $\delta_k \propto t^\alpha$ with $\alpha_+ = 2/3$ (growing mode) and $\alpha_- = -1$ (decaying mode). For $\alpha_+ = 2/3$, $\delta_k \propto t^{2/3} \propto a$, the perturbation will grow proportional to the scale factor, independent of the scale, as long as we are in the linear regime. For $\alpha_- = -1$, $\delta_k = t^{-1}$, the decaying solution is not physical and can be proven to be equivalent to a gauge change in the framework of GR.

- **Dark Energy-dominated Universe:** After the Λ-matter density equivalence, the expansion of the Universe is accelerated by dark energy. The pressure and gravity terms on the RHS of the instability equation can be neglected and the equation can be simplified into:

$$\ddot{\delta}_k + 2H\dot{\delta}_k = 0 \tag{1.57}$$

  where $H$ is the constant expansion rate characterising the de Sitter expansion phase. The solution to this equation is of the form $\delta_k \propto \textit{constant}$, i.e. the density perturbations are frozen in the Λ-dominated Universe. It should be kept in mind that this is only applicable to the linear regime, as for highly evolved structures ($\delta \gg 1$) their evolution will continue.

### Statistical Properties of LSS

In the previous section, we derived the gravitational instability equation governing the evolution of density perturbations and looked at how these fluctuations grew in different regimes. Now, we look at the statistical properties of these spatial fluctuations and discuss some notable methods by which we can study these perturbations from the observational data.

**Correlation Functions:** The distribution of matter in space can be studied by correlation functions. The correlation functions provide a description of the clustering properties of a set of points distributed in space. It can be a three-dimensional space, but we can also obtain useful information for two-dimensional distribution of points.

The joint probability $\delta^2 P_2$ of finding one point in a small volume $\delta V_1$ and another in volume $\delta V_2$, separated by a fixed vector $r_{12}$, when choosing the two volumes randomly, is given by

$$\delta^2 P_2 = n_V^2[1 + \xi(r_{12})]\delta V_1 \delta V_2 \tag{1.58}$$

where $n_V$ is the number of points per unit volume and the function $\xi(r)$ is the two-point correlation function. Assuming the cosmological principle, it can be said that $\xi(r)$ depends only on the magnitude of the vector $r$. If the distributions of the points are random, then the correlation between the points is $\xi(r) = 0$. If $\xi(r) > 0$, then the points are clustered and if $\xi(r) < 0$ then the points tend to avoid each other.

In an analogous way, we can define correlation functions of higher order, $N > 2$. It can be defined as the joint probability of finding $N$ points in $N$ disjoint volumes $\delta V_i$. However, this function will contain contributions from correlations of lower order than $N$. A more useful measure would be a reduced correlation function that is simply the terms of $\xi^N$ that do not depend on correlations of lower order. We can write the reduced three-point correlation function as:

$$\delta^3 P_3 = n_V^3[1 + \xi(r_{12}) + \xi(r_{23}) + \xi(r_{31}) + \zeta(r_{12}, r_{23}, r_{31})]\delta V_1 \delta V_2 \delta V_3 \tag{1.59}$$

where $\zeta \equiv \xi_3$ is the reduced three-point correlation function. For $N = 4$, the spatial correlation function, $\xi_{(4)}$ is defined as:

$$\begin{aligned}\delta^4 P_4 = n_V^4[1 + \xi(r_{12}) + \xi(r_{13}) + \xi(r_{14}) + \xi(r_{23}) + \xi(r_{24}) + \xi(r_{34}) + \\ \xi(r_{12})\xi(r_{34}) + \xi(r_{13})\xi(r_{24}) + \xi(r_{14})\xi(r_{23}) + \\ \zeta(r_{12}, r_{23}, r_{31}) + \zeta(r_{12}, r_{24}, r_{41}) + \\ \zeta(r_{13}, r_{34}, r_{41}) + \zeta(r_{23}, r_{34}, r_{42}) + \\ \eta(r_{12}, r_{13}, r_{14}, r_{23}, r_{24} r_{34})]\delta V_1 \delta V_2 \delta V 3 \delta V_4\end{aligned} \tag{1.60}$$

Similarly, we can define higher-order correlation functions.

**Power Spectrum:** In large scales, it is better to measure the Fourier transform of the two-point correlation function rather than the correlation function itself. This is because the Fourier transformed function provides insights into the underlying Physics which can be readily interpreted.

For a statistically homogeneous distribution, the two-point correlation function is the Fourier transform of the *power spectrum*. The Fourier transform of the overdensity can be written as

$$\delta(\vec{x}, t) = \int \delta_{\vec{k}}(t) exp[i\vec{k}.\vec{x}] \frac{d^3 k}{8\pi^3} \tag{1.61}$$

The matter power spectrum can be defined as the inverse Fourier transform of the correlation function:

$$P(\vec{k}) = \int \xi(\vec{r}) exp[-i\vec{k}.\vec{r}] d^3k \tag{1.62}$$

These equations satisfy the following relation:

$$\langle \delta_{\vec{k}}(t) \delta_{\vec{k'}}(t) \rangle = 8\pi^3 \delta_D(\vec{k} + \vec{k'}) P(\vec{k}) \tag{1.63}$$

where $\delta_D$ is the Dirac-delta function. It is possible to study different scales independently as different scales are not correlated. Invoking cosmological principle and assuming the initial matter power spectrum to be in the form of scale-invariant power spectrum, we can write down the initial matter power spectrum as:

$$P(\vec{k}) = Ak^{n_S - 1} \tag{1.64}$$

where $A$ is the initial amplitude and $n_S$ is the spectral index. To compute the matter power spectrum at a different cosmological time, we need to consider the evolution of the initial matter power spectrum, and some other effects like neutrino drag, and this can be mathematically achieved by multiplying the power spectrum by a scale dependent factor called *transfer funstion*, $T(k,t)$, which takes into account all these effects.

$$P(k,t) = Ak^{n_S - 1} T^2(k,t) \tag{1.65}$$

These methods are useful when the discrete distribution is uniform and isotropic. As we shall see, studying the multiscale anisotropic and inhomogeneous complex structure and morphology of LSS requires methods other than the n-point correlation functions and matter power spectrum.

## 1.5 Cosmic Web

The LSS of the Universe, formed by the distribution of matter, on Megaparsec scales is nonuniform and known to be arranged in a multiscale web-like pattern called a cosmic web (Bond et al. 1996). The weblike structure can be noticed in the spatial distribution of galaxies, initially suggested by the early nearby galaxy redshift surveys (Gregory et al. 1978; de Lapparent et al. 1986; Shectman et al. 1996) and later confirmed by many large galaxy redshift surveys like the Two-degree-Field Galaxy Redshift Survey (2dFGRS, Colless et al. 2003), the Six-degree-Field Galaxy Redshift Surveys (6dFGS, Jones et al. 2004), the Cosmic Evolution Survey (COSMOS, Scoville et al. 2007), the Sloan Digital Sky Survey (SDSS, Tegmark et al. 2004), the Galaxy and Mass Assembly (GAMA, Driver et al. 2011) and the Two Micron All Sky Survey Redshift Survey (2MASS, Huchra et al. 2012)

The emergence of such large non-linear structures from the initial perturbations can be intuitively understood by the Zel'dovich approximation (Zel'dovich 1970). The Zel'dovich approximation predicts a rich complex morphological structure of cosmic web (Doroshkevich et al. 1980; Pauls and Melott 1995). The cosmic web can be classified into four classes of substructures: dense and compact clusters or halos,

long elongated filaments, two-dimensional tenuous sheets or walls, and surrounding enormous nearly empty voids, all intricately connected to form a really complex network, showing different morphological properties and characteristics. It should be noted that this classification is not very rigid. The Figure 1.5 shows the cosmic web structure obtained from the Millenium simulation at redshift, $z = 0$.

- **Halos**: Halos form the most overdense regions of the cosmic web. Their evolution is driven by gravitational collapse. Though they were initially introduced to account for the invisible matter in galaxies to explain the rotation curves, the current models focus on their large scale distribution and the distribution of matter within them.

  Halos were considered approximately spherical dark matter concentrations in the beginning, but they are now understood as more or less ellipsoidal in shape, as they exhibit triaxiality, with three distinct axes of three different lengths (Kasun and Evrard 2005; Novikov et al. 2006; Hayashi et al. 2007; Prada et al. 2019). Galaxy clusters are formed inside the halos as a result of infalling matter.

- **Filaments**: Cosmic filaments form the most defining feature of the cosmic web. Filaments are one-dimensional tubular structures in the cosmic web. They are over dense regions, and appear to be 'highways', channelling matter and gas into the higher density nodes or halos (Knebe et al. 2004), possibly enabling star formation (Keres et al. 2005; Seth and Raychaudhury 2020). The evolution of filaments is also gravitationally driven as in the case of halos but at a slower pace.

  Filaments are observed to scale from a few Mpc to 100 Mpc, and are found to connect supercluster complexes such as the great attractor (Lynden-Bell et al. 1988), and the Vela supercluster (Kraan-Korteweg et al. 2017).

- **Sheets**: Like filaments, Sheets are also over dense regions but two dimensional, usually considered as a bundle or collection of filaments. However, they are much more difficult to find in the spatial mass distribution traced by galaxies as they are populated by low-luminous galaxies (Cautun et al. 2014).

  Some prominent flattened supercluster configurations, known as *Great Walls* have been recognised while observing the spatial structure outlined by the clusters. Some notable examples are the CfA great wall (Geller and Huchra 1989), the Sloan Great Wall (Gott et al. 2005), and the BOSS Great Wall (Lietzen et al. 2016).

- **Voids**: Voids are large underdense regions that are devoid of any galaxies, though they are highly irregular individually but can be considered spherical in shape on average. Their sizes range from 20-50 $Mpc/h$ and occupy the majority of space in the Universe. The evolution of voids is outward as matter collapses into surrounding structures. Recently, voids as standard spheres have been used to study and constrain cosmic expansion history, see Hamaus et al. (2020).

Along with the filaments, the void plays an important role in the spatial organisation of the cosmic web. Their existence was initially found in early galaxy redshift samples (Chincarini and Rood 1975; Gregory et al. 1978). More recent studies have been involved in mapping and cataloguing voids in the Local Universe (Pan et al. 2012; Sutter et al. 2012).

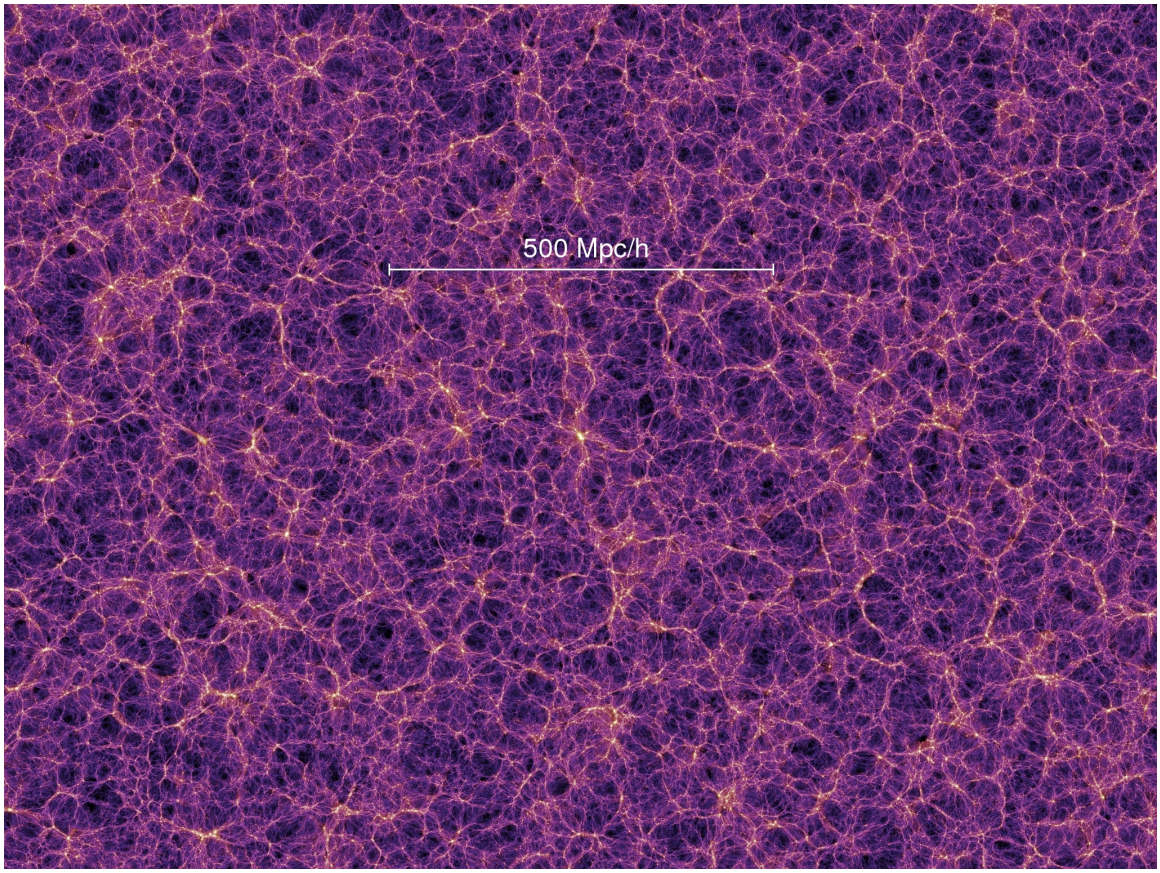



Figure 1.5: Cosmic Web. The image shows a projected density field for a 15Mpc/h thick slice at redshift z = 0 (t = 13.6Gyr) obtained by Millenium Simulation. Image Credit: Volker Springel, MPA.

# Chapter 2

# Simulations and Data Analysis Tools

In this chapter, we will discuss N-body simulations in the context of cosmology. These simulations play a huge role in cosmology, helping us to study the formation and evolution of the large-scale structures, anisotropies in the cosmic microwave background, and even testing various alternate theories of gravity. It is a difficult task to study large scale structures because of its multiscale complex network of structures with different morphological properties and the fact that the most of the matter in LSS is in the form of dark matter, which is non-luminous and interacts only through gravitation as we know of today. We need to look at the distribution of matter that traces the underlying LSS. We test various cosmological models with cosmological simulations, by performing various statistics on them and comparing them with the observational data from large-scale galaxy surveys. The observation of LSS is limited by various factors, such as galaxies being biased tracers of LSS, redshift, and instrumentation limitations, whereas the n-body simulations provide information on the 3D distribution of matter across different redshift. We will then introduce the cosmological simulation of our choice for this thesis work, called DEMNUni, which is a simulation of dark matter particles that account for extended cosmological scenarios with different neutrino masses and varying dark energy models. We will also briefly discuss the role of massive neutrinos and dark energy in structure formation and evolution.

Finally, we introduce the software tools and methods used to perform data analysis with these simulations. We will be looking at the snapshots of these simulations at different redshifts. We require specific tools to visualise this full-sky distribution of the matter which naturally exists on a spherical surface. We use a pixelisation scheme called HEALPix which allows us to do this task, be it the galaxy distribution map from high redshift surveys or dark matter halos distribution from simulations. We will be working with a Python package called *healpy*, based on the pixelation scheme provided by HEALPix.

## 2.1 N-body Simulations

As already mentioned, the observation of the large-scale distribution of matter is limited since most of the matter is in the form of non-baryonic dark matter and galaxy distributions are a biased tracer of the LSS. This is where N-body simulations play an important role. N-body simulations are the simulation of a dynamical system of particles which are evolved over time under the influence of various physical forces, usually gravity. Numerical simulations of different cosmological models and parameters are created and statistical methods are used to analyse these simulations, which is then compared to the observational data obtained from large-scale galaxy surveys such as the Sloan Digital Sky Survey (SDSS, Tegmark et al. 2004), Dark Energy Survey (DES, Sánchez 2016), Euclid (Euclid Collaboration 2024), etc.

There are a vast number of n-body simulations in the literature today. Simulations considering different physical processes, varying volume, resolution, and particles with various properties are developed to study different questions like the nature of dark matter and dark energy, structure formation and evolution, or galaxy evolution. Simulations can also be classified on the basis of different techniques used to evolve the particles. For example, there are collisionless n-body simulations with dark matter particles that only interact gravitationally mainly to study LSS and then there are hydrodynamical simulations with particles representing dark matter and gas to study galaxy formation and AGNs. For a detailed review on various techniques used for N-body simulations and different large dark matter simulations, see Dolag et al. (2008); Angulo and Hahn (2022).

Numerical simulations have played an important role in the last few decades in testing the ΛCDM model. Advances in algorithms and computational power during this period have made the simulations more realistic, vast, and precise. This has enabled to test various competing models to be analysed and compared with the observational data. ΛCDM model has so far been compatible with the observations of LSS and even with simulations accounting for non-linear physics, see Springel et al. (2006); Vogelsberger et al. (2020))

Simulations are currently widely used in cosmology for various purposes, like for the interpretation of observations in terms of cosmological parameters and underlying physics, for the development and testing of analytic models for structure formation, for generating input data to train for data-driven methods, creating mock universes for current and upcoming large-scale surveys for testing statistical methods and error analysis, and to study different aspects of cosmological models and physical processes.

The large volume of n-body simulations is defined by the simulation box, which represents a finite volume of the Universe. The box represents a finite volume of the Universe. The simulation box is defined by the parameters. Some notable box parameters that define the simulation box are the box size that defines the dimensions of the box ($L$), the number of particles in the box ($n_p$), the mass of the particles ($m_p$), and the resolution determined by the softening length ($\epsilon$), which prevents numerical artefacts in gravitational interactions. The Figure 2.1 shows an example of the simulation of structure formation in a simulation box.

For this work, we will be using simulations called DEMNUni Castorina et al.

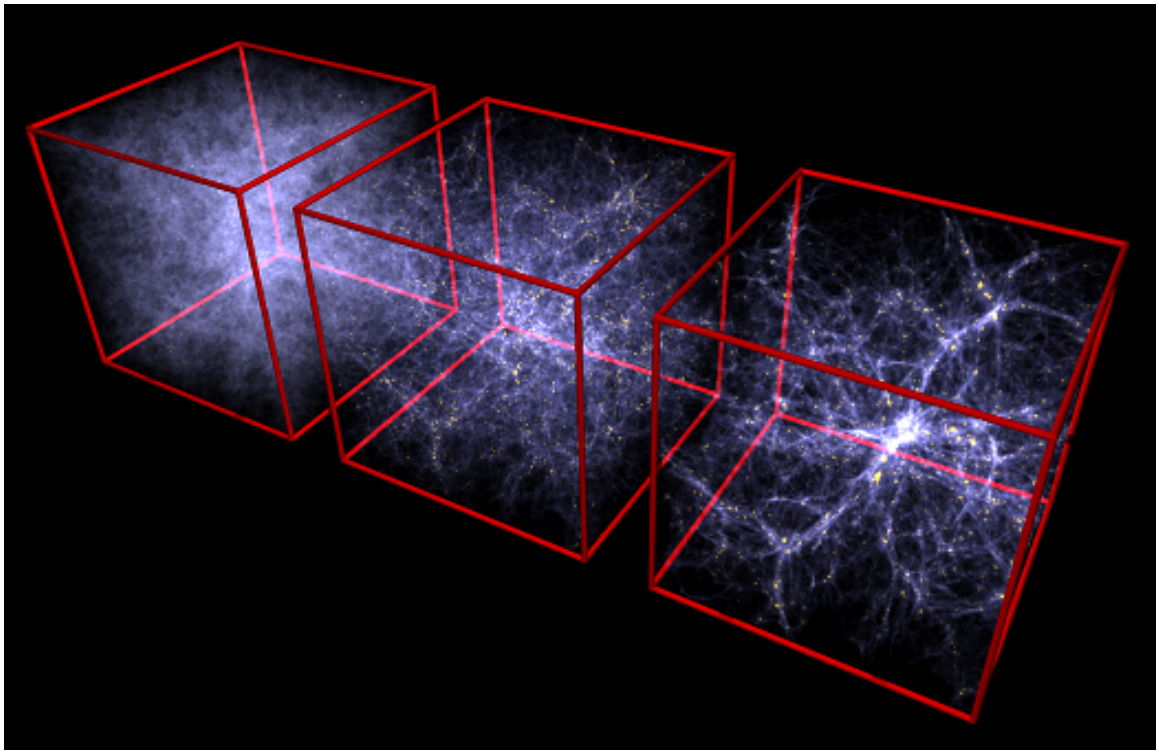


Figure 2.1: An example of simulation of structure formation in a simulation box. From left to right shows the simulation of the gaseous component of the Universe at age, $t = 0.9$ billion years old, 3.3 billion years old and the present moment respectively. Image credit: Volker Springel; Max-Planck-Institute for Astrophysics

(2015), which consider cosmological realisations of different but fixed neutrino masses and different models of varying dark energy.

### 2.1.1 DEMNUni

DEMNUni stands for the *Dark Energy and Massive Neutrino Universe* simulations (Castorina et al. 2015). It is a state-of-the-art large set of cosmological N-body simulations in terms of volume and mass resolution. The DEMNUni simulations have been developed for testing different cosmological probes like galaxy distribution, lensing, and their cross-correlations in the presence of massive neutrinos and a dynamical dark energy component. The simulation was originally conceived to study how neutrinos with their non-zero mass affect the CMB, LSS and the distribution of matter in the Universe.

The three neutrinos of the standard model of particle physics, are the electron neutrino ($\nu_e$), the muon neutrino ($\nu_\mu$) and the tau neutrino ($\nu_\tau$) and are considered massless in the ΛCDM cosmological model. Though, the evidence for neutrino oscillations presented by the Super-Kamiokande Collaboration in 1998 (Fukuda et al. (1998)), suggesting that at least two out of the three are massive neutrinos. Studying the effects of neutrino masses on cosmological observables is of great relevance for

two distinct reasons. First, the absolute neutrino mass scale remains unknown, and cosmology can play a key role in its determination, since gravity is sensitive to the total neutrino mass $\Sigma m_\nu$, rather than to the mass splitting. Second, an accurate description of the effects of massive neutrinos on LSS is important to avoid systematic errors in the determination of cosmological parameters.

Cosmologies with massive neutrinos have been extensively studied in the literature, especially the linear perturbation theory in the presence of massive neutrinos is well explored and understood, and widely used to put constraints on the neutrino mass, $\Sigma m_\nu$. With increasing efforts to precisely measure and estimate cosmological parameters, an accurate description of non-linear corrections is required. This can be achieved by directly analysing N-body simulations, which take into account the presence of massive neutrinos. This turns out to be very expensive from a computational point of view and really challenging if a good mass resolution and a large box size are required simultaneously to generate mock catalogues for present and future galaxy surveys.

The present mass constraints from lepton flavour oscillations have fixed the lower limit of total neutrino mass, $\Sigma m_\nu \equiv m_{\nu_e} + m_{\nu_\mu} + m_{\nu_\tau} \geq 0.06$ eV (see Lesgourgues and Pastor 2006, 2012, Lesgourgues et al. 2013, and Lesgourgues and Pastor 2014 ) . This implies that the neutrinos free-stream with larger thermal velocities after becoming non-relativistic, suppressing the growth of neutrino density perturbations at smaller scales than the free-streaming length, given by

$$\lambda_{fs}(z, m_\nu) \simeq 8.1 H_0 \frac{(1+z)}{H(z)} \left(\frac{1eV}{m_\nu}\right) \frac{Mpc}{h} \tag{2.1}$$

where $m_\nu$ is the mass of the single neutrino species. This results in gravitational backreaction effects, altering the evolution of density of cold dark matter and baryons, and the suppression of total matter power spectrum at scales much smaller than free-streaming length ($\lambda \ll \lambda_{fs}$). The upper limit of total neutrino mass, $\Sigma m_v < 0.12$ eV comes from cosmological observations (see Planck Collaboration 2020). The present mass constraints suggest that the neutrinos become non-relativistic after the epoch of recombination, and therefore modifying the radiation density contribution. This late transition from the relativistic to non-relativistic regimes postpones the matter-radiation equivalence for a given value of the matter density, $\Omega_m h^2$, where $\Omega_m$ is the ratio between the matter density and the critical density at $z = 0$. The effects of massive neutrinos on the CMB anisotropies and lensing signals have been studied with DEMNUni simulations; see Carbone et al. (2016).

The cosmic voids which form the extended underdense regions of the cosmic web are the largest observable structures in the universe. Their enormous size and underdense nature make them particularly suitable for investigating dark energy and neutrino effects. There are various cosmic void statistics, such as the void size function and the void-galaxy cross-correlation to test the different cosmological models. The dark energy effects on the LSS, particularly on cosmic voids have been studied recently with extended DEMNUni suite of simulations considering several dynamical dark energy equations of state (EoS) along with different total neutrino masses, see Verza et al. (2019); Kreisch et al. (2019) and Verza et al. (2023)

Now, let us look into the DEMNUni simulations in detail. DEMNUni simulations are performed using the specifically modified tree particle mesh-smoothed particle hydrodynamics (TreePM-SPH) code GADGET-3 which is an improved version of the codes described in Springel (2005), modified to incorporate the presence of massive neutrinos. The modified GADGET-3 (Viel et al. 2010) code follows the evolution of the Cold Dark Matter (CDM) and neutrino particles, considering them as two separate collisionless fluids.

The set of simulations has a starting redshift of $z_{in} = 99$ and is characterised by a large comoving volume of $8h^{-1}Gpc^3$ to include the very large scale perturbation modes and a good mass resolution with $2048^3$ dark matter particles and $2048^3$ neutrino particles to investigate the effects of small-scale nonlinearities i.e. the nonlinear evolution of density perturbations and neutrino free-streaming.

The dark matter halo catalogues are built from the DEMNUni simulation outputs by identifying halos using a friends-of-friends (FoF) algorithm (Davis et al. 1985) applied to dark matter particles. The code is only applied to CDM particles with linking length set to 0.2 times the inter-particle distance. A minimum number of 32 particles are assumed to identify a structure, fixing the minimum halo mass to $M_{FoF} \sim 2.5 \times 10^{12} h^{-1} M_{\odot}$. These halo catalogues can be further processed with the SUBFIND algorithm (Springel et al. 2001; Dolag et al. 2009) to produce subhalo catalogues. This procedure can cause some of the initial halos to be split into multiple substructures.

The Figure 2.2 shows the full sky distribution of dark matter halos obtained from DEMNUni simulations with ΛCDM cosmology and zero total neutrino mass, $\Sigma m_{\nu} = 0$.

As mentioned earlier, DEMNUni simulations have been used to study massive neutrino effects on CMB anisotropies and lensing, and LSS, especially studying the cosmic voids and statistics (Kreisch et al. 2019, Schuster et al. 2019). In this thesis work, we focus on the cosmic filaments, by reconstructing filaments in DEMNUni simulations using the filament finding algorithm called SCMS, which we shall explain in the Chapter 3.

## 2.2 HEALPix

HEALPix stands for the Hierarchical Equal Area isoLatitude Pixelisation of a sphere (Górski and Hivon 2011). As the name suggests, it is a pixelation scheme that allows a division of a spherical surface, where each pixel has the same area on the sphere. HEALPix was originally developed for the constructing full-sky maps of microwave sky at a very high resolution of a few arcminutes from the multifrequency data obtained from the satellite missions like NASA's Wilkinson Microwave Anisotropy Probe (WMAP, Bennett et al. 2003), and ESA's Planck mission (Planck Collaboration 2014). The fundamental requirements in the development of this tool were to create a mathematical structure which allows a suitable discretization of functions on a sphere and to facilitate fast and accurate statistical and astrophysical analysis of massive full-sky data sets.

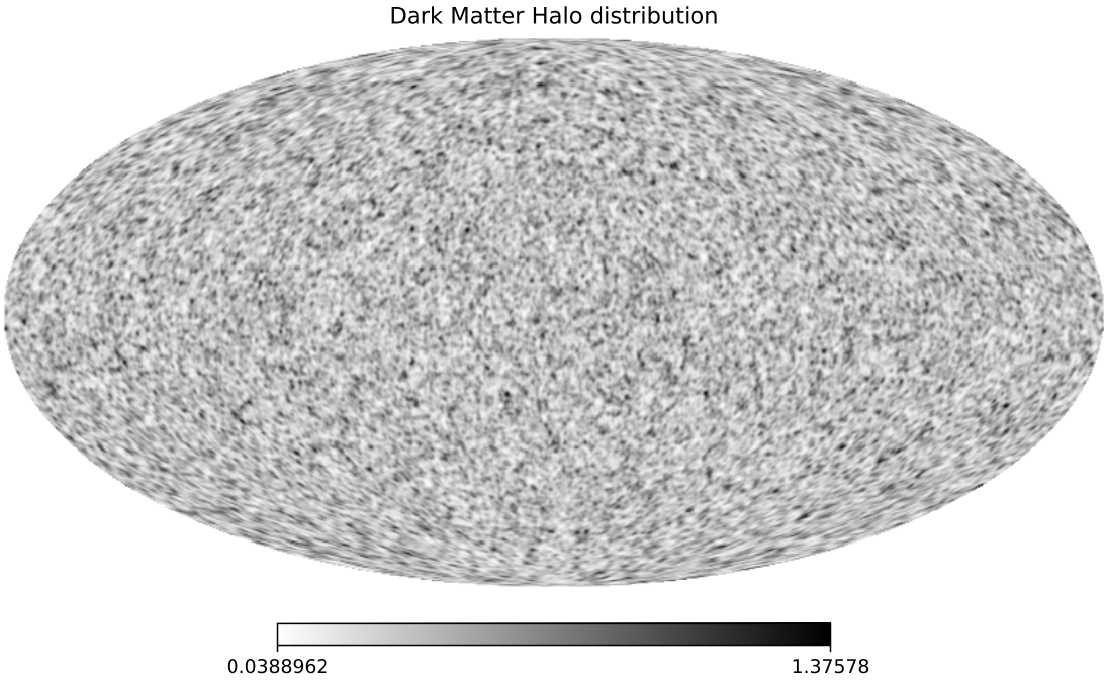



Figure 2.2: Mollweide projection of dark matter halos distribution map at $z = 1.007$ obtained from DEMNUni simulations with ΛCDM cosmology and the total neutrino mass, $\Sigma m_\nu = 0$. The map has been smoothed with a Gaussian kernel with fwhm = 1 degree for visualisation purpose.

This can be achieved because it has the following three essential properties:

- The sphere is hierarchically tessellated into curvilinear quadrilaterals with the lowest resolution comprised of 12 base pixels. The resolution is increased by dividing each base pixel into four new ones, and so on. See Figure 2.3, which depicts the increasing resolution on the sphere.
- At a given resolution, all pixels have equal areas.
- The pixels are distributed along lines of constant latitude. This is essential for all harmonic analysis applications that involve spherical harmonics. This iso-latitude distribution of sampling points improves the speed of computation of integrals over individual spherical harmonics because it scales as $N^{1/2}$ with the total number of pixels, as compared to noniso-latitude sampling distributions which scale as N.

HEALPix software is available in various languages like C, C++, Fortran90, IDL, Java, and Python with an extensive library of specific tools and supporting routines. The *healpy* is a Python package to handle pixelated data on the sphere based on the HEALPix pixelisation scheme and bundles the HEALPix C++ library (Zonca et al. 2020). healpy provides utilities in spherical harmonic transforms allowing for fast

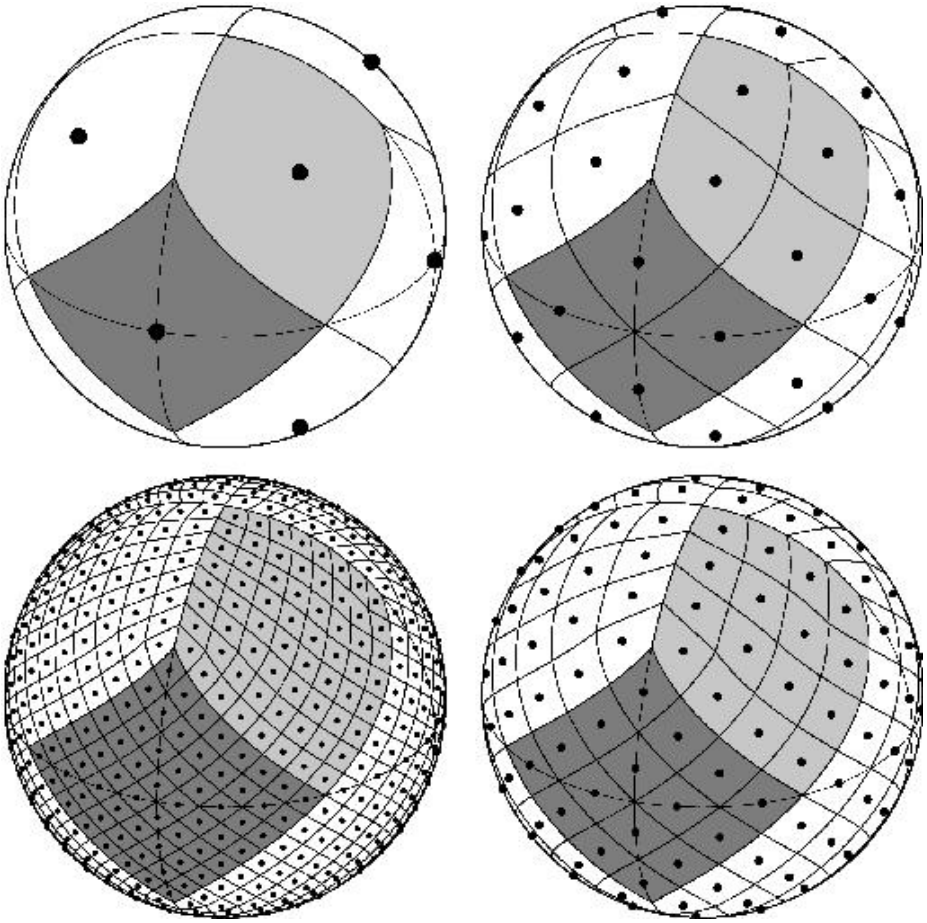


Figure 2.3: HEALPix Pixelisation scheme on a sphere with increasing resolution. The sphere is partitioned into 12, 48, 192 and 768 pixels, respectively. Image Credit: HEALPix

simulation and analysis of full-sky CMB maps, pixel manipulation such as changing resolutions, masking and pixel queries, and visualization of maps in various projections. The resolution of the map in *healpy* is defined by a parameter called *nside*. The total number of pixels, $N_{pix}$ is given by;

$$N_{pix} = 12 \times nside^2 \tag{2.2}$$

and the pixel area can be computed by;

$$A_{pix} = \frac{4\pi}{N_{pix}} = \frac{\pi}{3 \cdot nside^2} \tag{2.3}$$

The corresponding angular resolution is approximately;

$$\theta_{pix} \approx \frac{58.6}{nside} \tag{2.4}$$

The Figure 2.4 shows the Quasar distribution map in mollweide projection from the SDSS DR16 catalogue visualised using the *healpy* package.

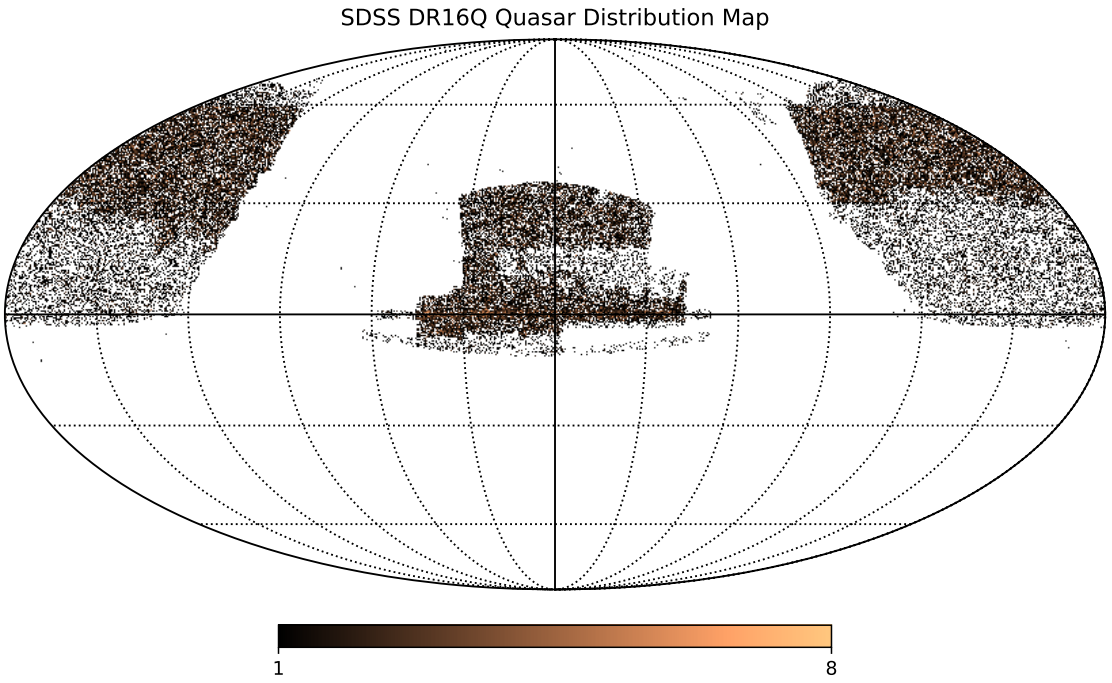



Figure 2.4: Mollweide projection of Quasar distribution obtained from the SDSS DR16 Quasar Catalog, visualised using healpy package.

# Chapter 3

# Filament Reconstruction Algorithms and SCMS

In this chapter, we briefly introduce some widely used and existing filament-finding algorithms based on various mathematical frameworks. The filaments and other cosmic web structures are reconstructed from given tracers such as the maps of distribution of galaxies and clusters. And we use the simulated data to test the reconstruction techniques and model the physical properties of the cosmic web structures. It is possible to use dark matter halos directly as a tracer of the large-scale structure in this case. Most of the algorithms require the tracer maps to be smoothed to obtain a continuous distribution from the discrete tracer field. Smoothing reduces noise in the data, enhances large-scale structure by filtering finer smaller discrete structures, and also facilitates the calculation of the first- and second-order derivatives, required for Hessian-based methods. The smoothing is performed by a kernel of our choice for the regularisation of the tracer field, which is chosen to be Gaussian in our case as it preserves LSS, is isotropic, and computationally faster. The kernel could be of any type, such as Boxcar, exponential, etc. The smoothing scale is determined by the fwhm of the kernel. Thus, optimisation of the smoothing scale is of utmost importance in these methods and needs to be studied in detail, as we shall see in this chapter and the following.

In particular, we explain the mathematical formalism and implementation of the Subspace Constrained Mean Shift (SCMS) algorithm that we adopt for the detection and reconstruction of filaments in the DEMNUni simulations of dark matter halos, and the uncertainty estimation in the reconstructed filaments. We extend the algorithm by a novel method of two-step filament reconstruction using two different kernel scales to smooth the density maps to improve the reliability and robustness of the SCMS algorithm, as we shall explain in this chapter later. We also discuss some statistical methods for verifying the accuracy of our reconstructed filaments.

## 3.1 Filament Finders

As we already know, matter in the Universe is distributed in a weblike pattern called the cosmic web, and the cosmic filaments form the most defining features of this web. Filaments delimit cosmic voids and connect halos and clusters.

Cosmic filaments have been extensively studied in the literature from different perspectives (see Colberg et al. 2005, Aragón-Calvo et al. 2010, Cautun et al. 2014). Unlike halos, for which there is a well-known universality in their structure (Navarro et al. 1997, Merritt et al. 2006) and recent studies suggesting a similar universality in the properties of voids (Pan et al. 2012, Hamaus et al. 2014, Nadathur et al. 2015), filaments lack such universality in properties regarding their evolution and structural features, such as length and width (see Galárraga-Espinosa et al. 2020, Aragón-Calvo et al. 2010). Since there is a lack of such universality, it is usually necessary to use realistic simulations for cosmological inferences, like the sensitivity of filamentary structure to the nature of dark matter and dark energy or the presence of massive neutrinos. And this is a difficult problem to address due to the complexity and multiscale structure of filaments, and the corresponding disagreement between the results of different filament finding algorithms (Libeskind et al. 2017).

The lack of consensus in the definition of filaments in the literature, along with the need to handle different input data types and to study different questions, has led to the development of various algorithms based on different mathematical and physical approaches to reconstruct and trace the position of filaments in large galaxy survey data and simulations.

We will now look at some widely used filament-finding algorithms based on various techniques. A detailed classification and comparative study of different algorithms that trace the cosmic web and its components can be found in Libeskind et al. (2017). Recently, some efforts have been made to understand the importance of optimised smoothing in recovering the filament properties by reducing any bias generated by the filament finder. (Dhawalikar and Paranjape 2024).

- **Graph & Percolation methods**: The connectivity of the elongated supercluster structures in the distribution of matter was initially studied by percolation analysis (Zeldovich et al. 1982, Zel'dovich and Shandarin 1989, Shandarin 2004), which helps in quantifying the single large scale structure made from small isolated groups. A similar graph-based method, the minimum spanning tree (MST) of galaxy distribution was extensively studied (Barrow et al. 1985) with the aim of developing an objective measure of filamentarity. For example, MST algorithms by Alpaslan et al. (2014), or Pereyra et al. (2020).

- **Hessian based methods**: There is a class of algorithms that utilises the geometric information included in the Hessian of the density, tidal or velocity shear fields, which describes different aspects of matter distribution and dynamics of the cosmic web like mapping the mass distribution or mass flow around or within the filaments. Most Hessian-based algorithms are defined on single scale smoothing of the field, there are algorithms which use multiscale smoothing and

look at the structures from a scale space perspective, where the smoothing scale defines an extra dimension. Examples: MMF (Aragón-Calvo et al. 2007) and NEXUS (Cautun et al. 2013).

- **Topological Methods**: Algorithms based on topological methods focus on analysing the cosmic web by studying the connectivity and topological properties of the underlying field compared to Hessian-based algorithms that look at the local geometric structure of density, velocity, or tidal field. An example of this class of methods is the Spineweb procedure (Pereyra et al. 2020), which is capable of tracing the structures of the cosmic web purely on topological basis. Spineweb does this by identifying the central axis of the filaments and the core plane of the walls with the boundaries between the watershed basins of the density field. It is based on Morse theory (Morse 1934) which describes spatial connectivity of the density field based on the location and identity of its singularities - maxima, minima, and saddle points. Their connectivity is defined by the gradient of the density field. Examples: DisPerSE (Discrete Persistent Structures Extractor, Sousbie 2011) which is a widely used algorithm these days. They identify filaments and voids on both observational and simulated data. Skeleton analysis (Novikov et al. 2006)

- **Stochastic Methods**: Stochastic methods use the statistical evaluation of stochastic geometric concepts. The examples include filament finding algorithms based on the Bayesian sampling of parametrised and well-defined spatial (marked) point processes that model distinct geometric configurations. A major advantage of this method is that the algorithm can be applied directly to the galaxy point distribution instead of computing a continuous density field. Stoica et al. (2005) and Tempel et al. (2016) use the Bisous model as an object point process of aligned and connected cylinders to locate and catalogue filaments in galaxy surveys.

- **ML based methods**: In addition to these methods, recently there have been some developments on algorithms based on machine learning classification such as random forests (Buncher et al. 2021) and convolutional neural networks (CNNs) like in Aragon-Calvo (2019), where a CNN is trained using a segmentation method on a N-body simulation to perform semantic segmentation of filaments and walls in LSS.

Many of these filament finding algorithms have been further modified to make use of different tracer maps and reconstruct filaments on both n-body simulations and large-scale galaxy surveys and study different cosmological questions. For this work, we will be using the SCMS algorithm as proposed in Chen et al. (2015) based on the ridge formalism.

| Algorithm | Class | Identified Feature |
|---|---|---|
| MST | Graph & Percolation | Filaments |
| Bisous | Stochastic | Filaments |
| NEXUS | Hessian | Cosmic Web |
| MMF | Hessian | Cosmic Web except nodes |
| Spineweb | Topology | Cosmic Web except nodes |
| DisPerSE | Topology | Cosmic Web except nodes |
| SCMS | Hessian | Filaments |

Table 3.1: Tabular summary of some popular algorithms and their class, and the identified structure.

## 3.2 SCMS Algorithm

SCMS stands for the Subspace Constrained Mean Shift algorithm. The SCMS algorithm is based on a modified gradient ascent method that models filaments as one-dimensional maxima called ridges in the galaxy density distribution. The algorithm is well understood mathematically (as discussed in Eberly 1996 and Genovese et al. 2012) and has good properties such as convergence and allows us to estimate the uncertainty for consistent detection, which is why SCMS is the algorithm of our choice for this study. Now, let us look at the underlying mathematical formalism of the algorithm.

### 3.2.1 Ridge Formalism

If we consider a galaxy or dark matter halo distribution with $n$ galaxies or halos with locations $x_1, x_2, x_3, ....., x_n$, which are $d$-dimension points. For typical galaxy surveys, $d = 2$, if the distribution is constrained to the redshift shell or $d = 3$, a three-dimensional distribution with redshift being the third dimension. We consider galaxy positions as random variables drawn from an unknown density function $p$. Let $g(x) = \nabla(p(x))$ and $H(x)$ be the gradient and the Hessian of $p(x)$ respectively. Let $v_1(x), v_2(x), ..., v_d(x)$ be the eigenvectors of the Hessian matrix, with the associated eigenvalues $\Lambda_1(x) \geq \Lambda_2(x) \geq .... \geq \Lambda_d(x)$. We define $V(x)$ to be the matrix of all eigenvectors orthogonal to the first, i.e. $[v_2(x), ...., v_d(x)]$, and the ridge, R, is defined as

$$R = Ridge(p) = x : G(x) = 0, \Lambda_2(x) < 0 \tag{3.1}$$

where

$$G(x) = V(x)V(x)^T g(x) \tag{3.2}$$

is the projected gradient. So, in short, ridges are the local maxima with a projected gradient of 0 in the subspace spanned by the eigenvectors $v_2(x), ...,$
$v_d(x)$. The ridges are one-dimensional generally smooth curve-like structures with a density higher than that of its neighboring regions, and thus have the properties of filaments. Compared with the other filament detection methods, density ridges use information from both the gradient and the Hessian matrix of density, unlike MMF, NEXUS and NEXUS+, which only use second derivatives, and DisPerSE models, which use only the first derivatives. Another advantage of the density ridge model is that there exists a well-established statistical theory for the consistent estimation of the density ridge (Chen et al. 2015; Genovese et al. 2012).

Although the definition of the ridges aligns with the intuitive idea of the filaments, i.e., they are one-dimensional overdense structures. It is possible to slightly modify the ridge formalism to define the other structures of the Cosmic web as following;

- **Halos**: If the subspsace $V(x)$ is the whole space and $\lambda_1 < 0$, we get the 0-dimensional overdense structures which correspond to the nodes or halos.
- **Voids**; If the subspsace $V(x)$ is the whole space and $\lambda_3 > 0$, we get the 0-dimensional underdense structures, which correspond to the cosmic void centres.
- **Sheets**: If the subspsace $V(x)$ is only spanned by $v_3$ and $\lambda_3 < 0$, we get the 2-dimensional overdense structures, which correspond to the sheets or walls.

All these definitions always include structures of lower dimensionality i.e., the definition of sheets includes filaments or the definition of filaments includes halos.

### 3.2.2 Filament Detection

The SCMS algorithm is an iterative algorithm and the way we implemented for filament reconstruction mainly has two steps as follows: estimation of the underlying density function of a distribution of galaxy positions and then defining a uniform mesh grid, and moving the grid points towards the filamentary structures in iterative steps.

We will consider galaxies within different thin redshift slices, i.e. 2-dimensional distributions of points projected on the sphere, to obtain information about the filaments at different redshifts. An important advantage of this approach is that it will reduce errors caused by distance estimation at the expense of losing sensitivity to filaments along the line of sight. Detailed step-by-step implementation of the SCMS algorithm is given below (Chen et al. 2015, Carrón-Duque 2022)

*Input*: A collection of points on the sphere given by $\{x_i\} \equiv \{(\theta_i, \phi_i)\}$, which represents the galaxy distribution or dark matter halo distribution within a redshift slice.

1. Estimate a density function from the given distribution defined over the sphere, $d(\theta, \phi)$ by smoothing the distribution of points with a certain kernel of choice.

2. Calculate the gradient and Hessian of the density, given by $g(\theta, \phi) \equiv \bigtriangledown d(\theta, \phi)$ and $H(\theta, \phi)$ respectively.

3. Diagonalise the Hessian at every point.

4. Find the eigenvectors $v_1(\theta, \phi), v_2(\theta, \phi)$ with the smallest eigenvalue corresponding to $v_2$

5. Project the gradient $g$ onto the eigenvector $v_2$. Let $G(\theta, \phi)$ be this projection.

6. Define a uniform distribution of points $\{y_k\}$ around a sphere for detecting filaments. Iterate until convergence. At every step n, move all points $y_k^n$ in the direction of projection at that point, i.e., $y_k^{n+1} = y_k^n + c \cdot G(y_k^n)$, where c is the normalization of the projection.

7. The points have reached the filaments if they stopped moving between the iterations.

*Output*: A collection of points $y_k^n$ on the sphere that overlap the filaments.

The Figure 3.1 depicts the implementation of the SCMS algorithm on a distribution of dark matter halos obtained from DEMNUNi simulations. However, it should be noted that this method has been used to reconstruct filaments in 3D distribution by transforming redshift into distance estimates and using the definition of ridges with $d = 3$. But we will be working in 2D distribution by selecting all the objects within a small redshift range, so they are a slice of distribution of galaxies or halos within a thin redshift shell to avoid any artifacts such as redshift-space distortions, finger-of-god effect or Kaiser effect (see, Hikage and Yamamoto 2015) .

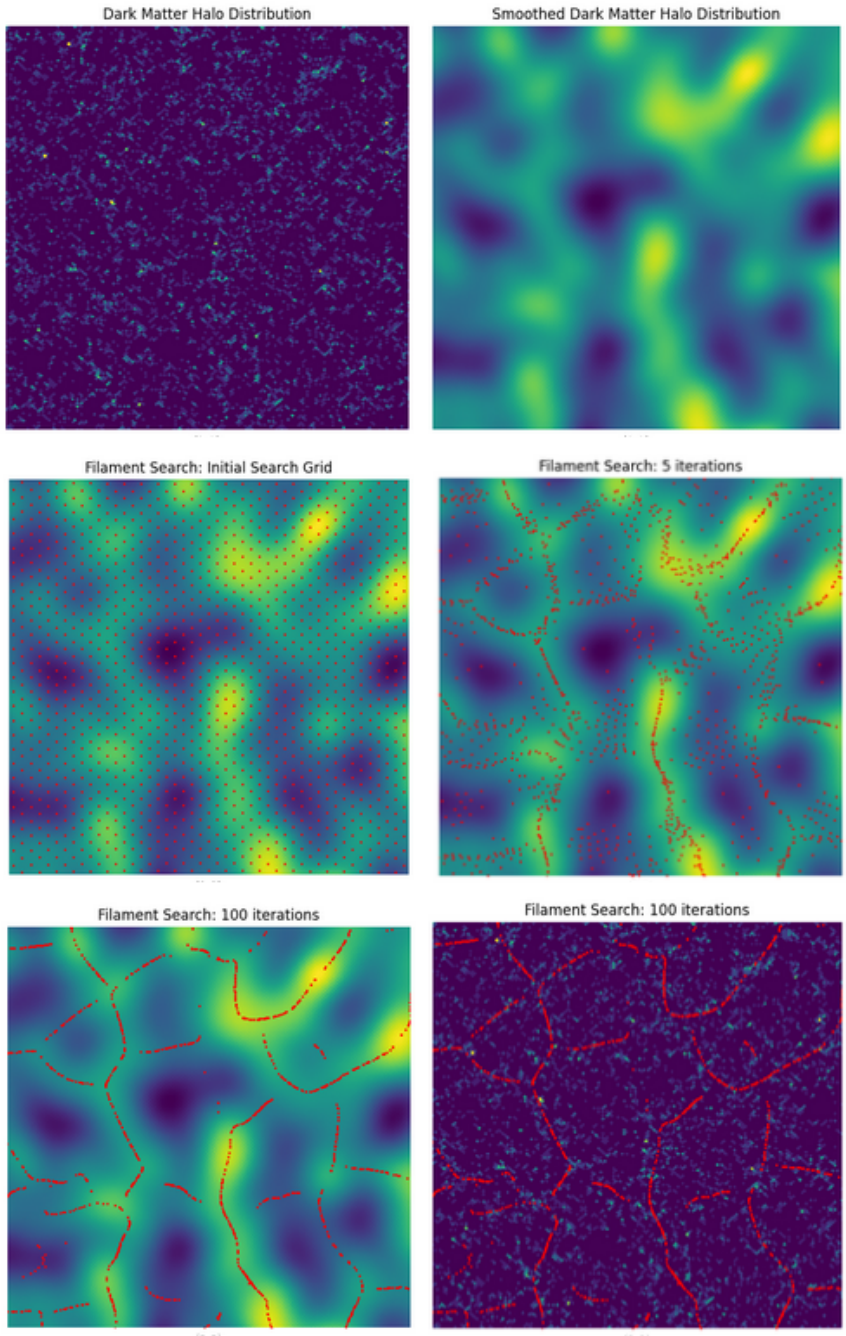



Figure 3.1: Pictorial representation of SCMS algorithm reconstructing filaments in dark matter halo distribution. *Top left*: Initial Dark Matter halo distribution in gnomview projection. *Top right*: Smoothed halo distribution map. *Center left*: Smoothed density map overlayed with initial uniform filament search grid points by center of the pixel of a map whose resolution is given by $nside = 128$. *Center right*: Grid points after 5 iterations of SCMS algorithm. *Bottom left*: Reconstructed filaments after 100 iterations. *Bottom right*: Reconstructed Filaments overlayed on initial halo distribution map. All patches are of the size $1000 \times 1000$ pixels with the resolution of 1 arc minute per pixel and centred at $RA = 0$, $Dec = 0$.

### 3.2.3 Kernel Size Effects

The filament reconstruction using the SCMS algorithm explained in the previous section depends on the Gaussian kernel scale used for smoothing and density estimation. To analyse the effects of various smoothing scales on reconstruction, we initially choose three arbitrary angular smoothing scales: 2, 3 and 4 degrees (which are also guided by the values used in the literature) and compare the reconstructed filaments using the stacking method, which we will explain later in this chapter. It should be noted that the physical scales corresponding to these angular scales vary a lot between different redshift slices (as we shall see later in Chapter 4, refer to Table 4.3). And therefore we look at filament reconstruction using a fixed physical scale as a kernel across different redshift slices to reconstruct filaments.

The angular scale $\theta$ in radians subtended by a constant physical scale ($D_{ps}$) at any given redshift $z$, is given by the following relation:

$$D_{ps} = D_A(z) \cdot \theta \tag{3.3}$$

$$\theta = \frac{D_{ps}}{D_A(z)} \tag{3.4}$$

where $D_A(z)$ is the angular diameter distance at redshift z and is defined as

$$D_A(z) = \frac{D_c(z)}{1+z} \tag{3.5}$$

where $D_c(z) = c\int_0^z \frac{dz'}{H(z')}$ is the comoving distance at $z$.

We find that the filament reconstruction by the SCMS algorithm is poor when using small physical scales for smoothing. This can be attributed to the fact that the magnitude of projection by which the grid points move with each step of the iteration varies significantly when smoothed with lower physical scales and this cannot be resolved by normalisation of the magnitude with a single value, as we shall see in the Chapter 4. To overcome this issue, we develop a novel method, modifying the existing SCMS algorithm by using two kernels of different scales, which we shall call the *Two-step Filament Reconstruction* method.

## 3.3 Two-step Filament Reconstruction

Using a single physical scale optimal for smoothing all the redshift slices comes with its own challenges. We need to optimally choose a physical scale that will not oversmooth the lower redshift slices and undersmooth the higher redshift slices. We find that the convergence of the points is highly affected by a highly varying distribution of magnitudes of projection at each pixel position, which determines how much a grid point moves with each iteration, while using the SCMS method using a smaller physical scale, as we shall see in section 4.3. To overcome this problem of nonconvergence of points in SCMS, we propose a novel method of filament reconstruction using the SCMS Algorithm in two steps; we shall call it the two-step filament reconstruction (2sFR) method.

We smooth the matter distribution map with a kernel of larger physical scale of choice, perform SCMS for $n$ iterations (typically between 10-25), and save the positions of the grid points after $n$ iterations. We then smooth the original map with the smaller physical scale of choice, and perform SCMS again but using the output of the first set of iterations as the initial grid points for the second set of iterations. The idea behind this method is to aggressively smooth the map with a larger kernel scale to properly initiate filament search and facilitate the convergence, and then a less aggressive smoothing could be able to retain the smaller structures in the field. Let us take a look at how we perform this method computationally.

**2sFR Algorithm**

*Input*: Distribution of galaxies or dark matter halos in a particular redshift slice

1. Perform SCMS for $n$ iterations on the matter distribution map smoothed using a larger kernel of choice $k_1$.
2. Save the final positions of the grid points after the first set of iterations, say $y_{k_1}$.
3. Perform SCMS again on the original map that is now smoothed with a smaller kernel of choice $k_2$, but using the output of the first set of iterations, $y_{k_1}$ as the initial grid points to search for filaments.
4. The points will have reached the filaments once they have stopped moving between the iterations.

*Output*: Filament positions, $y_{k_2}$ on a map smoothed with a smaller kernel $k_2$.

The Figure 3.2 shows the 2sFR method performed on a redshift slice in gnomonic projection.

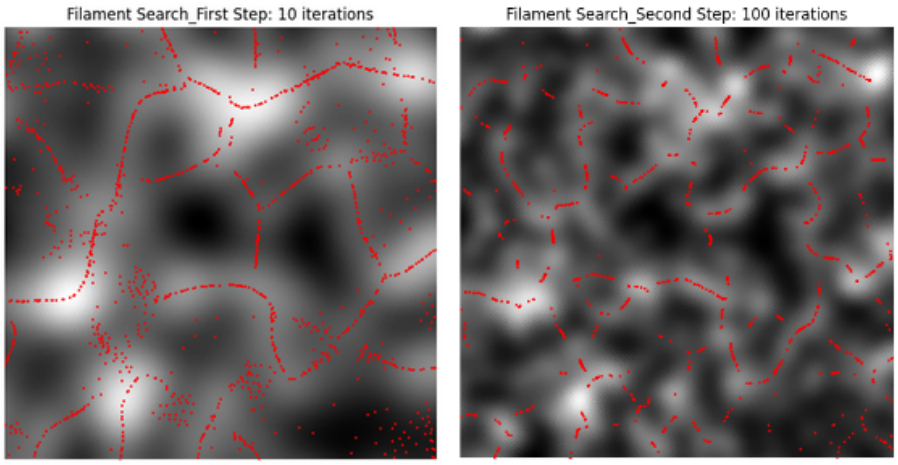



Figure 3.2: The 2-step filament reconstruction method. *Left*: Grid points after 10 iterations on a map of halos smoothed with the larger kernel $k_1$. *Right*: Grid points after 100 iterations on a map smoothed with the smaller kernel. All patches are of the size $1000 \times 1000$ pixels with the resolution of 1 arc minute per pixel and centred at $RA = 0$, $Dec = 0$.

## 3.4 Uncertainty in filament Reconstruction

Filament reconstruction by the SCMS algorithm detailed in the previous sections does not provide any robustness of the result. This can be overcome by measuring the uncertainty of each reconstruction using the uncertainty algorithm provided in Chen et al. (2015).

The concept behind this method is to generate different realizations of the distribution of galaxies or halos using the bootstrapping method. We take a random resample of the galaxies to create a new modified realization of the galaxy distribution. Then, we compare the filaments reconstructed in the original distribution with the filaments reconstructed in the new realization of galaxies or halos. This process is repeated several times to compute the robustness of each reconstructed filament. This algorithm is executed computationally as follows:

**SCMS Uncertainty Algorithm**

*Input*: Original distribution of galaxies or dark matter halos

1. The *true filaments* are reconstructed on the original data using SCMS; let $y_j$ be a point on the real filament.

2. A new set of resampled galaxies is generated by bootstrapping the original distribution of galaxies. We keep the number of galaxies or halos the same as in the original distribution.

3. SCMS is performed on this new realization of the galaxies and filaments are

reconstructed; let $y_j^n$ be the points on the filaments reconstructed on this bootstrap sample.

4. For each point on the true filament, we compute the minimum distance to the closest filament in the bootstrapped sample, $\rho_n(y_j) = min(d(y_j, y_j^n))$. The smaller the distances, the more consistent the reconstruction of the filaments.

5. Steps 2 to 4 are repeated for a total of N = 10 times. For each point on the true filament, we compute a minimum distance for each bootstrap sample; then we compute the quadratic mean of all simulations: $\rho(y_j) = \sqrt{\frac{1}{N}\Sigma_{n=1}^{N}\rho_n(y_j)^2}$

*Output*: An error estimate i.e., the average minimum distance for every point in the true filaments, $\rho(y_j)$.

The estimated value of the uncertainty can be interpreted as the average variability of the location of a given point in a filament as the distribution of galaxies changes slightly with keeping the same underlying distribution. Increasing the number of bootstrap samples will improve the uncertainty estimation. We limit the number of bootstrap samples to 10 considering the fact that the process is computationally very expensive and time consuming.

It should be noted that the uncertainty estimation in the case of 2sFR method follows the same algorithm as above but instead of using SCMS we perform the 2sFR method to reconstruct filaments in the original and bootstrap samples.

## 3.5 Stacking of Filaments

The stacking technique is commonly used to measure and study the typical contribution of a group of objects when the individual contribution of each object in the group is dominated by noise. Thus, averaging the contribution from different objects can reduce the noise and retrieve a signal with a higher signal-to-noise ratio.

Various stacking methods have been widely used in the field of Galaxy Cluster research (Dutson et al. 2013; Morandi et al. 2015) and more recently to study cosmic filaments (de Graaff et al. 2019; Hadzhiyska et al. 2024). In addition to the difficulty of accurately detecting the filaments, the main challenge is raised by the anisotropic nature of the filaments: they have a preferred direction along their axis. Therefore, it is necessary for all filaments to be stacked along their preferred axis to preserve and study this anisotropy. This is also why we cannot use other standard tools like the angular power spectrum or n-point correlation functions,as they will give no information on the anisotropy as they are an average over all directions. We will stack the density contrast field traced by the filaments to assess the quality of the filament reconstruction, we should be able to see a filamentary structure surrounded by underdense voids when stacking, if the filament reconstruction is correctly implemented.

We adopt the stacking method used in Carrón-Duque (2022) which was used to study the galaxy overdensity traced by cosmic filaments. Let us consider a spherical

map M (full sky dark matter halos distribution in our case). For a single filament and a position in the sky $\alpha \in S^2$. We have a signal measured at $\alpha$ that can be expressed as the sum of the filament and other effects:

$$M(\alpha) = M_{filament}(\alpha) + M_{noise}(\alpha) \tag{3.6}$$

where the noise factor is dominant. When we add the contributions of many filaments, say $X_{filament}$, aligned in the same way, $X_{filament}$ will grow linearly approximately, losing information on individual filaments is traded with a significant increase in the signal-to-noise ratio $SNR_{filament}$ of the average signal of the filament. Whereas $X_{noise}$ is independent of the filaments, we will reduce the average noise contribution by adding all the independent random realisations of the noise.

To computationally implement the stacking, we randomly select a set of points on the reconstructed filament network. For every point, we have its location, the direction of the filament, and, optionally, the uncertainty of the detection. We then create a rectangular grid of dimensions 200 Mpc x 40 Mpc centered at every point such that the short side of the box is parallel to the filament axis. In this way, we will focus on the radial effects of the filaments. The analysis of the filaments along the axis is limited by the fact that they bend and intersect with each other and, therefore, cannot be extended further than this.

The axis of a filament has two possible directions; we choose the direction in which the density of the halos in the original tracer map increases and stack it at the lower part of the grid. The rectangular grid in Mpc around each point is translated into spherical coordinates, with angular size depending on the redshift ( based on Planck $\Lambda$CDM values). We then compute the value of the map, M at each point of the grid to obtain a rectangular patch for every filament point considered. Then, we average the contributions of all patches, using a weight: the inverse of the uncertainty of the detection. A mask can be included in the analysis to avoid bad pixels or edges in the case of an incomplete sky. This process can be mathematically written as

$$S(\alpha) = \frac{\Sigma_j w_j \cdot r_j(M \cdot m)(\alpha)}{\Sigma_j w_j \cdot r_j(m)(\alpha)} \tag{3.7}$$

where $S$ is the stacked map at a point $\alpha$, $w_j$ is the weight of filament $j$, $M$ and $m$ are the map and mask, respectively, and $r_j$ is a spherical rotation operator defined so that the filament $j$ is rotated to fixed coordinates (say, $RA = 0$, $Dec = 0$, and $angle = 0$). This rotation allows all the filaments to align in the same way. We then have the stacked map, which is the average of filaments and noise factor independent of them:

$$S(\alpha) =< M(\alpha) >_{filaments} = \overline{M_{filaments}(\alpha)} + < M_{noise}(\alpha) >_{filaments} \tag{3.8}$$

From the above expression, it is evident that on averaging the filament signal to noise ratio ($SNR_{filament}$) will increase as $< M_noise >$ decreases drastically while the filament signal remain consistent.

We will then obtain a 200 Mpc x 40 Mpc grid of the average filament signal, where the filament axis lies along the centre of the image parallel to the shortest direction.

We also compute the average along the shortest direction to estimate the average filament profile as a function of the radial distance to the central axis.

## 3.6 Distance between Halos and Filaments

The filaments trace the overdense regions in the halos distribution map. Therefore, the distance between the halos and the reconstructed filaments has a correlation by construction. We can use this correlation as a figure of merit to check the accuracy of the filament reconstruction and quantify the differences between the two methods of filament reconstruction explained earlier. This is computationally achieved by;

**Halos to Filament Distance Algorithm**

*Input*: Original distribution of dark matter halos.

1. Perform SCMS or 2sFR method to reconstruct filaments on dark matter halos distribution map.
2. Let $x_i$ be the position of a halo and $y_j$ be the position on the filament closest to this halo. Compute the distance between these two positions, $d_k = d(x_i, y_j)$.
3. Repeat step 2 for every halo on the map.

*Output*: An array of distances between a halo to its nearest filament for every halo on the map.

# Chapter 4

# Filament Reconstruction with SCMS

We have explained different filament reconstruction algorithms and our motivation to use the Subspace Constrained Mean Shift (SCMS) algorithm for detecting and reconstructing cosmic filaments in the DEMNUni simulations along with various statistical methods to check the validity of the cosmic filament reconstruction in the previous chapter.
In this chapter, we discuss the main results of this thesis work. Firstly, we look at the filament reconstruction using the SCMS algorithm with Gaussian kernels of different angular scales at different redshifts. We will then discuss filament reconstruction using a fixed physical scale as a Gaussian kernel for smoothing. This is followed by the results of the novel two-step filament reconstruction method, which was developed to improve the filament reconstruction in density maps smoothed with small physical scales. We qualitatively check the accuracy of the filament reconstruction by looking at the reconstructed filaments at different locations in the full-sky map and by studying the underlying halo density traced by the filament points. We verify the accuracy quantitatively by estimating the uncertainty of the reconstructed filaments. We analyse the results by studying the underlying density of the halos traced by the filaments by stacking the density contrast maps at the position of the reconstructed filaments. This will give us information on the properties of the filaments and its surroundings on average. We also estimate the distance of halos from the reconstructed filaments at each redshift slice.

## 4.1 Input Dataset

For this thesis work, we perform filament reconstruction on a DEMNUni simulation with ΛCDM cosmology assuming zero neutrino mass. The cosmological and box parameters for the simulation set we used are as follows: $\Omega_{CDM} = 0.27$, $\Omega_b = 0.05$, $\Omega_\nu = 0.00$, $\Omega_\Lambda = 0.68$, $h = 0.67 \frac{km}{s \cdot Mpc}$ and *Box size* $L = 2000$ *Mpc/h*.

For studying smoothing scale dependence on the filament reconstruction, we take a

sample of three redshift slices of dark matter halos that correspond to low ($0.0968 < z < 0.1015$), intermediate ($1.0078 < z < 1.0159$), and high ($2.0702 < z < 2.0842$) redshifts from the full 3D N-body DEMNUni simulations. The redshift slices are sliced in a way that they are of 20 Mpc wide in comoving coordinates at every slice and kept thin to avoid projection effects. Let us call this sample A. Table 4.1 provides a summary of sample A with the lower and upper redshift limits of the redshift slice, redshift width or interval of the slice, the mean redshift of dark matter halos in the slice, and finally the number of halos per slice, denoted by $z_{low}$, $z_{high}$, $\Delta z$, $z_{mean}$, and $n_{halos}$ respectively. This is a sub-sample taken from a larger sample given below, to perform a faster analysis for studying the smoothing scale effects at different redshift ranges, and also makes it computationally cheaper to do so.

| Redshift Slice | $z_{low}$ | $z_{high}$ | $\Delta z$ | $z_{mean}$ | $n_{halos}$ |
|---|---|---|---|---|---|
| Low | 0.0968 | 0.1015 | 0.0047 | 0.0992 | 27228 |
| Intermediate | 1.0078 | 1.0159 | 0.0081 | 1.0119 | 1261773 |
| High | 2.0702 | 2.0842 | 0.014 | 2.077 | 1323572 |

Table 4.1: An overview of Sample A with lower and upper redshift bound, redshift width, mean redshift of halos, and the number of halos for low, intermediate and high redshift slices.

Similarly, a larger sample, Sample B of 11 redshift slices, is taken to understand and study how the use of a constant physical scale to smooth affects our filament reconstruction and the width of the recovered filaments as estimated by stacking the halo density traced by the filaments. Table 4.2 provides an overview of the 11 redshift slices in sample B and the distribution of halos with respect to the redshift is shown in Figure 4.1. Sample A is a subsample taken from sample B, as shown in the figure.

It should be noted that the number of halos is less at low redshift slices than at higher redshift slices, which seems counter-intuitive at first. However, this could be due to a variety of reasons, such as halo mergers or volume effects.

| Redshift Slice | $z_{low}$ | $z_{high}$ | $\Delta z$ | $z_{mean}$ | $n_{halos}$ |
|---|---|---|---|---|---|
| Slice 1 | 0.0968 | 0.1015 | 0.0047 | 0.0992 | 27228 |
| Slice 2 | 0.2189 | 0.2239 | 0.0050 | 0.2214 | 124365 |
| Slice 3 | 0.3498 | 0.3552 | 0.0054 | 0.3525 | 288085 |
| Slice 4 | 0.4915 | 0.4974 | 0.0059 | 0.4945 | 506379 |
| Slice 5 | 0.6464 | 0.6529 | 0.0065 | 0.6496 | 762007 |
| Slice 6 | 0.8173 | 0.8245 | 0.0072 | 0.8209 | 1038783 |
| Slice 7 | 1.0078 | 1.0159 | 0.0081 | 1.0119 | 1261773 |
| Slice 8 | 1.2222 | 1.2314 | 0.0092 | 1.2268 | 1468448 |
| Slice 9 | 1.4659 | 1.4763 | 0.0104 | 1.4711 | 1610109 |
| Slice 10 | 1.7457 | 1.7577 | 0.012 | 1.7517 | 1569741 |
| Slice 11 | 2.0702 | 2.0842 | 0.014 | 2.077 | 1323572 |

Table 4.2: An overview of Sample B, same as in Table 4.1

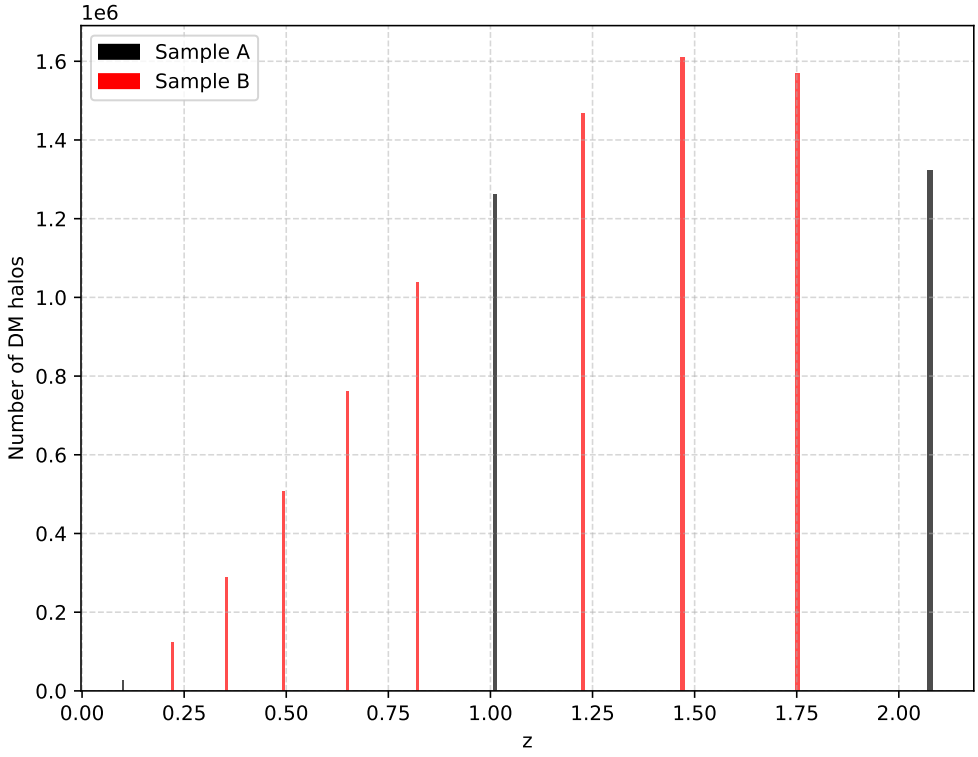



Figure 4.1: Sample B: Number of the dark matter halos as a function of redshift in the 11 redshift slices. The bars in black represents the redshift slices that are also in the Sample A.

## 4.2 Filament Detection and Reconstruction

We detect and reconstruct filaments in the three redshift slices in sample A using the SCMS algorithm, as described in Section 3.2. We perform SCMS on a map of the DM halos distribution with a resolution of $6.87'$ given by $nside = 512$, and reconstruct the filaments using grid points which are the centre of the pixels on a map with a smaller resolution of $13.7'$ given by $nside = 256$. A lower value, $nside = 128$ can also be taken to save computational time and faster convergence of the points. The first objective is to investigate how the smoothing kernel sizes affect the filament reconstruction in different redshift slices.

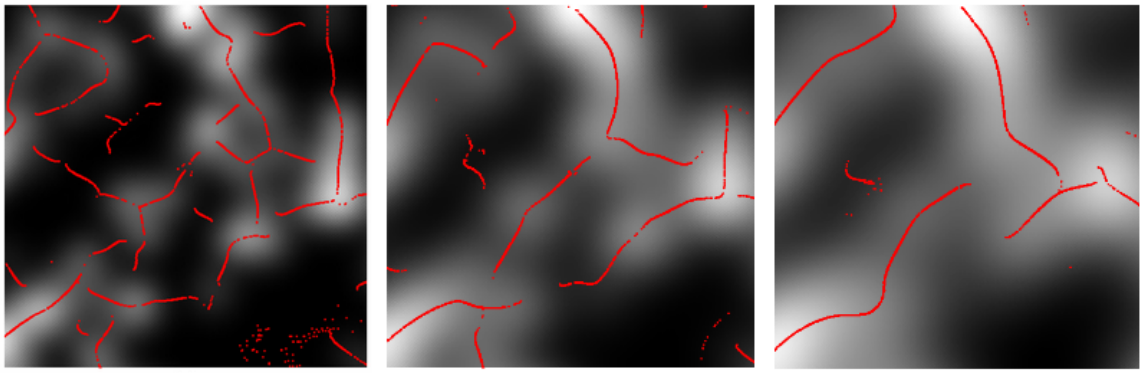
### 4.2.1 Kernel Size Effects

To understand and study how the kernel size used for smoothing affects filament reconstruction, we arbitrarily take three Gaussian kernels with fwhm corresponding to angular scales of 2, 3, and 4 degrees. The Table 4.3 shows the physical size corresponding to the angular kernel scale in each redshift slice.

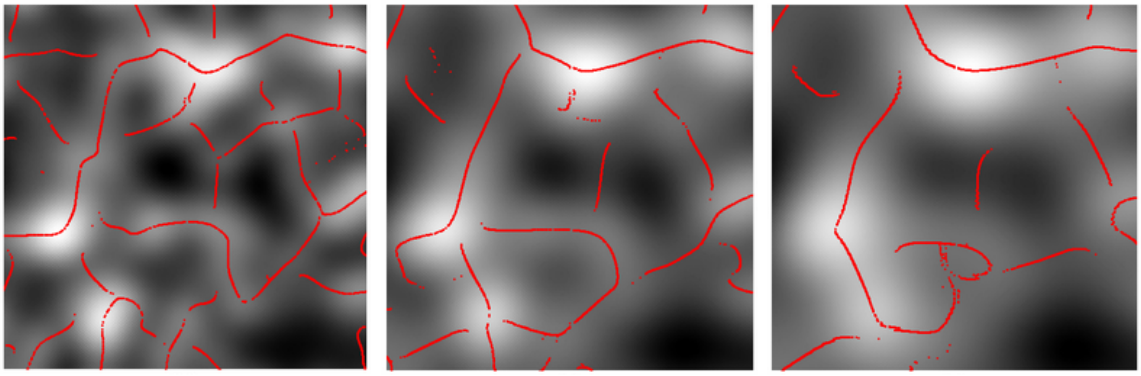
| Redshift | Angular Scale (Degrees) | Physical Scale (Mpc) |
|---|---|---|
| low | 2 | 13.60 |
| | 3 | 20.41 |
| | 4 | 27.21 |
| intermediate | 2 | 59.40 |
| | 3 | 89.10 |
| | 4 | 118.80 |
| high | 2 | 61.51 |
| | 3 | 92.26 |
| | 4 | 123.01 |

Table 4.3: Physical scales corresponding to the smoothing kernels in angular scales at each redshift for the slices in sample A.

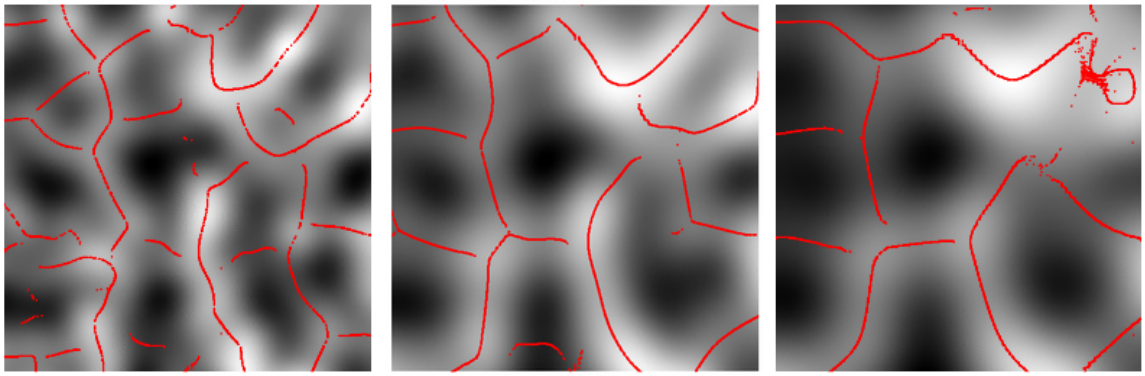
In Figure 4.2, we can observe how the filament reconstruction performs in different redshift slices for different angular scales. In the case of low redshift slice, we notice that for the map smoothed with the smallest kernel scale of 2 degrees, there are many grid points that have not yet converged to the overdense regions even after 100 iterations; this could be due to the lower number of halos in the low redshift slice. We also notice that the reconstructed filaments vary significantly with the increasing angular scales for smoothing. The number of filaments decreases; the filaments become longer and less curved with increasing kernel scale.



(a) Low redshift slice



(b) Intemediate redshift slice



(c) High redshift slice

Figure 4.2: Filament reconstruction visualised in gnomonic projection after 100 iterations in low, intermediate and high redshift slices (from top to down) in sample A. Each row represents a redshift slice smoothed with Gaussian kernels of angular scale 2 , 3, and 4 degrees (from left to right). All patches are of the size 1000 $\times$ 1000 pixels with the resolution of 1 arc minute per pixel and centred at $RA = 0$, $Dec = 0$.

Similarly, in the case of intermediate and high redshift slices, we observe a similar effect on the shape and length of filaments as we did in the low redshift slice. The number of filaments decreases, and the filaments are longer and less bent in shape.

But we do not see as many points that have not reached the overdense regions as we did in the case of the low redshift slice. This could possibly be due to the fact that there are far fewer halos in the low redshift slice compared to the other two.

If we look at each column of the Figure 4.2, we can visually analyse how the filament reconstruction performs when using the same arbitrary angular smoothing scale across different redshift ranges. We can see that the filament reconstruction improves as we go to high redshift slices, thanks to increasing number of halos per pixel leading to better estimation of density and projections.

We also observe that the map starts to get oversmoothed beyond a scale of 3 degrees. We can also notice how we lose information on smaller structures as we increase the smoothing scale and the reconstructed filaments become less fragmented, longer, and curvier.

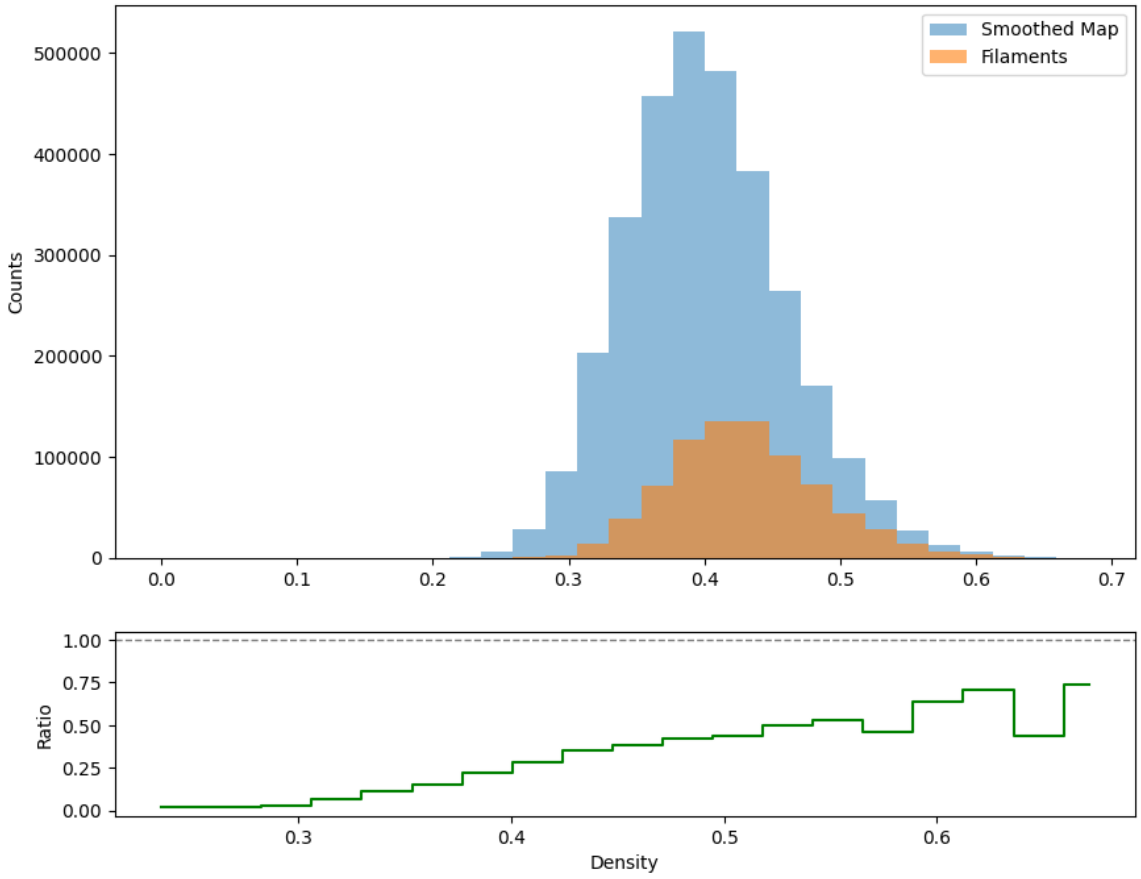



Figure 4.3: Histogram plot showing the underlying halo density traced by filaments compared to that of the whole map and their ratio for the intermediate redshift slice smoothed with a 3 degree Gaussian kernel.

Apart from visual confirmation, we check if our implementation of SCMS algorithm is correctly reconstructing the filaments by looking at the halo density at the positions of the points tracing the filaments compared to the overall smoothed halo density distribution of the input halo map. In Figure 4.3, the histogram plot shows

that the distribution of the halo density at the position of points tracing the filament on the intermediate redshift slice with a Gaussian kernel with FWHM of 3 degrees is off-centered or skewed compared to the density distribution of the smoothed input map. And if we look at the ratio of the counts of filaments to that of the overall smoothed map across the fixed bins, we can see an increasing trend as we move towards the overdense regions. Both indicate that our reconstructed filaments trace the overdense regions in the map compared to nearly empty voids. If the implementation of the algorithm was not correct, we would have expected the distribution of density traced by the filaments to be not skewed, would have looked like a smaller subsample of the normal distribution of the overall density, and the ratio would not have shown this consistent increasing trend.

### 4.2.2 Uncertainty

The uncertainty in the filament reconstruction is estimated using the SCMS uncertainty algorithm explained in Section 3.4. The uncertainty in degrees is estimated in the reconstruction of the filaments in a single redshift slice for the three different kernel scales. The Figure 4.4 shows the normalised distribution of the estimated uncertainty in the filaments reconstructed in the intermediate redshift slice. It is evident from the plot that the uncertainty increases with the increasing smoothing scale, indicating that over-smoothing not only affects reconstruction of smaller filaments but also increases the uncertainty in the filament detection as well.

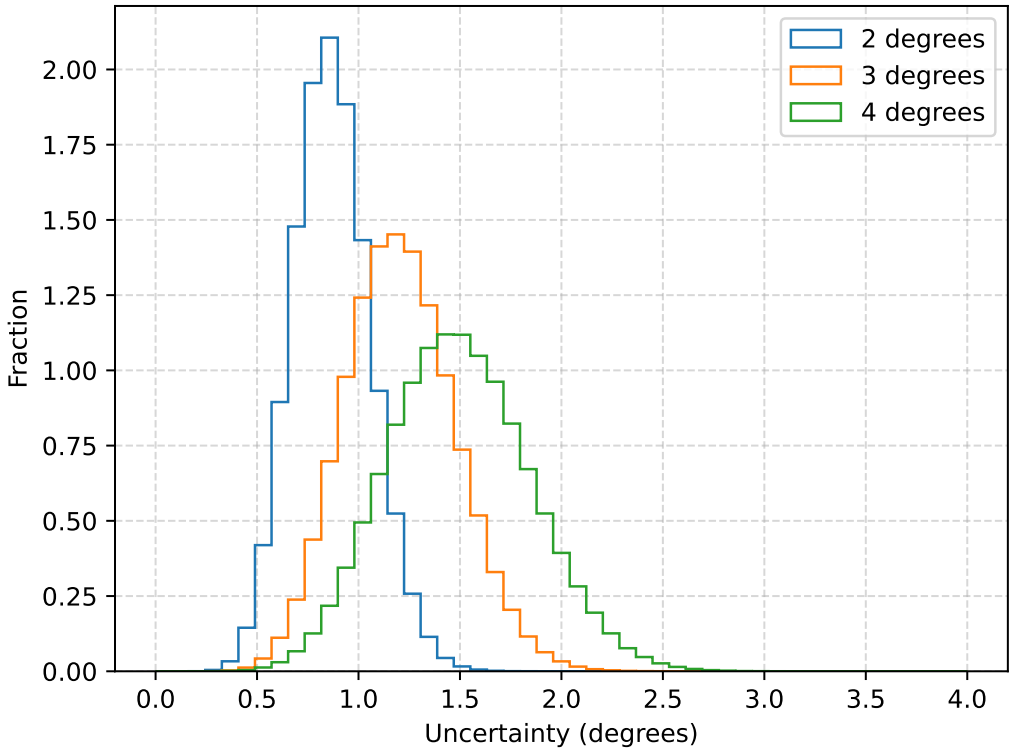



Figure 4.4: The normalised distribution of estimated uncertainty in filament reconstruction for the intermediate redshift slice in Sample A when smoothing the input halo density map with a Gaussian kernel of FWHM = 2, 3, and 4 degrees.

And in Figure 4.5, the histogram plot shows the normalised distribution of uncertainties in degrees estimated for the reconstructed filaments in all redshift slices of Sample A when smoothed with a Gaussian kernel of FWHM of angular scale of 3 degrees. The uncertainty is bigger in the low redshift slice compared to the intermediate and high redshift slices. This could again indicate to the point that a lower number of halos in the low redshift slice affects the consistent filament reconstruction.

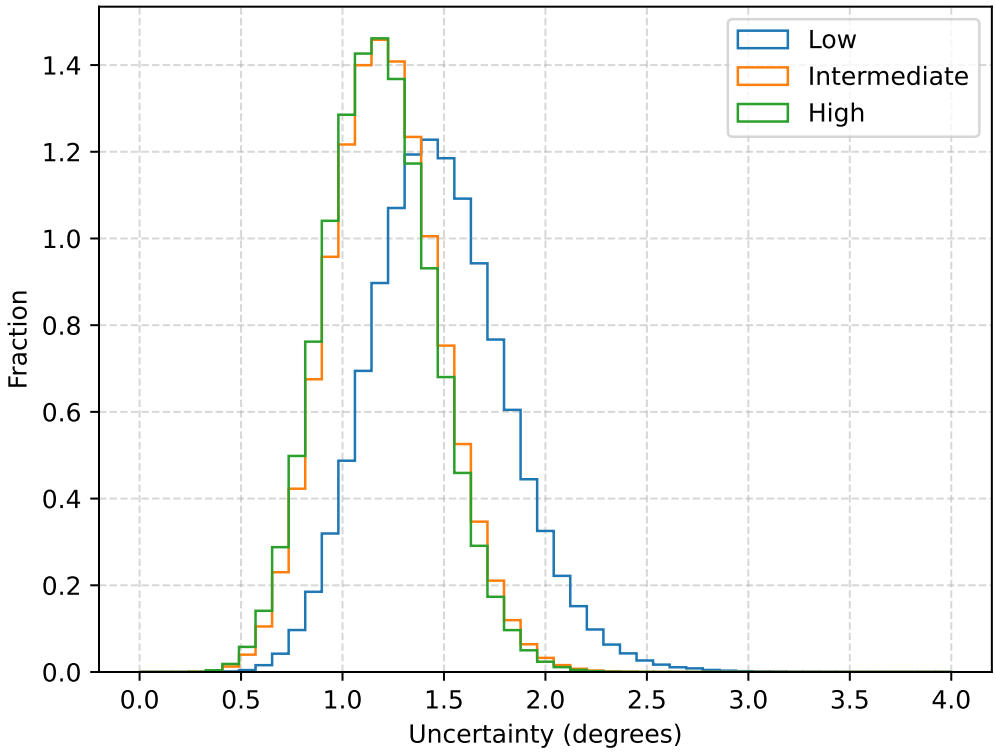



Figure 4.5: The normalised distribution of estimated uncertainty in filament reconstruction for all the redshift slices in Sample A when smoothing the input halo density map with a Gaussian kernel of FWHM = 3 degrees.

### 4.2.3 Stacking of Filaments

We stack dark matter halo density contrast ($\delta$) traced by the filaments using the stacking method explained in Section 3.5. The density contrast is defined as the relative density compared to the mean background density. It can be expressed mathematically as:

$$\delta = \frac{\rho - \langle\rho\rangle}{\langle\rho\rangle} \tag{4.1}$$

where $\rho$ is the density at a point and $\langle\rho\rangle$ is the mean background density. We perform stacking for the case of the intermediate redshift slice smoothed with kernel sizes 2, 3 and 4 degrees and recover the 200 Mpc $\times$ 40 Mpc 2D stacked grid and the stacked profile which is the average density contrast along the direction orthogonal to the filament axis.

We observe a gradient along the direction of the filament with a peak at the bottom centre of the grid; see Figure 4.6. This is expected as we are stacking filaments along the direction of increasing density, i.e. the filament's axis, and when we rotate the stacked grid, the density maxima will be at the bottom. We can also observe the underdense voids immediately on both sides of the filaments. The filaments appear wider, and the voids become less prominent with increasing kernel size.

In Figure 4.7, we have the stacked profile of filaments reconstructed in the intermediate redshift slice smoothed with the three angular scales, we can see the underdense voids and overdense filaments much more clearly in the profile and how increasing kernel size affects the filament width and the underdense regions. The oversmoothing of the input halo density map will result in poor detection of the position of the filament, as the centre of the filament becomes less prominent to be identified. This is clearly reflected in both the stacked grid and profile. The maximum of the profile is dispersed about the centre of the grid as the kernel size increases, and the profile looks more random wiggly.

Similarly, we do the stacking of filaments reconstructed in different redshift slices smoothed with angular scale of 3 degrees for all the slices in sample A. The results are shown in Figure 4.8 and Figure 4.9. From the stacked grid and the stacked profile for maps smoothed with Gaussian kernel of a given angular scale of 3 degrees, we observe that the filaments get wider at high redshift slices, and the peak value in the stacked profile starkly decreases with the increasing redshift. Both could be indicating towards the possible evolution of cosmic filaments with redshift. These maps were smoothed with a fixed angular scale, i.e., they are of different physical scales at these redshift slices. We also observed that varying angular scales for smoothing have an effect on the filament width. Thus, it would be interesting to see if there is such an evolution in filament width if we use a fixed physical scale across different redshift slices.

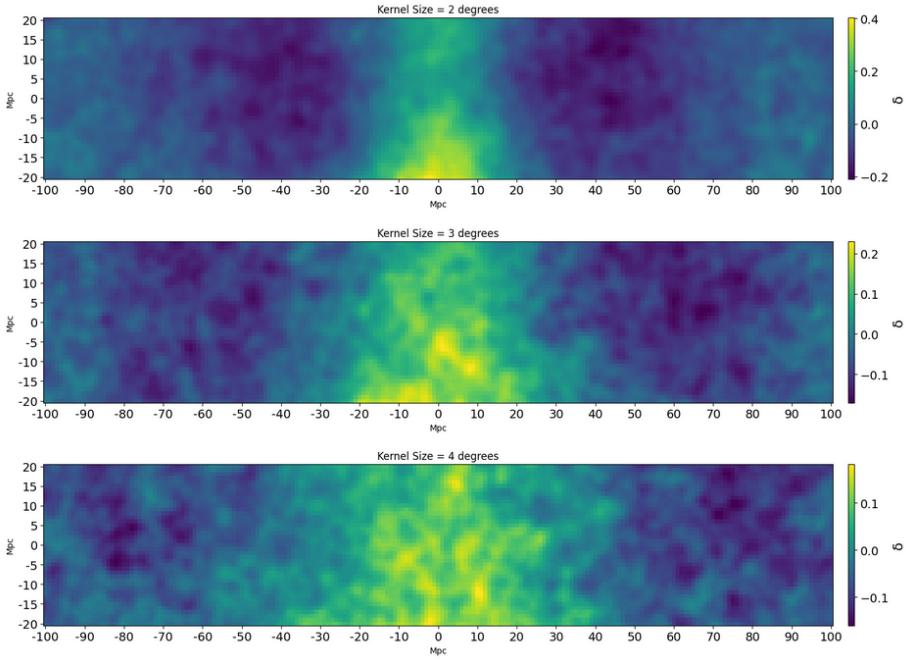



Figure 4.6: Density contrast of halos stacked at the position of filaments reconstructed on the intermediate redshift slice of Sample A smoothed with 2, 3, and 4 degrees from top to bottom respectively.

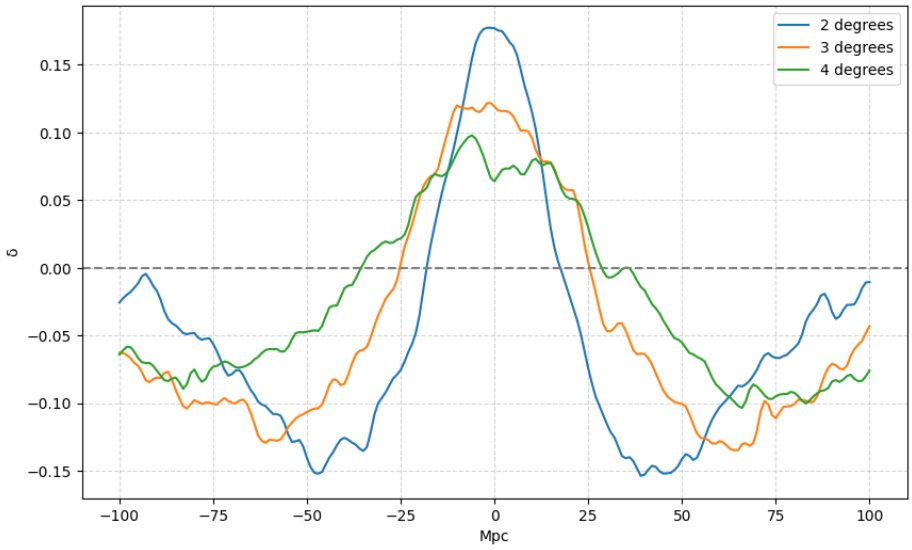



Figure 4.7: Stacked profile of filaments reconstructed on the intermediate redshift slice of Sample A smoothed with a Gaussian kernel of 2, 3, and 4 degrees.

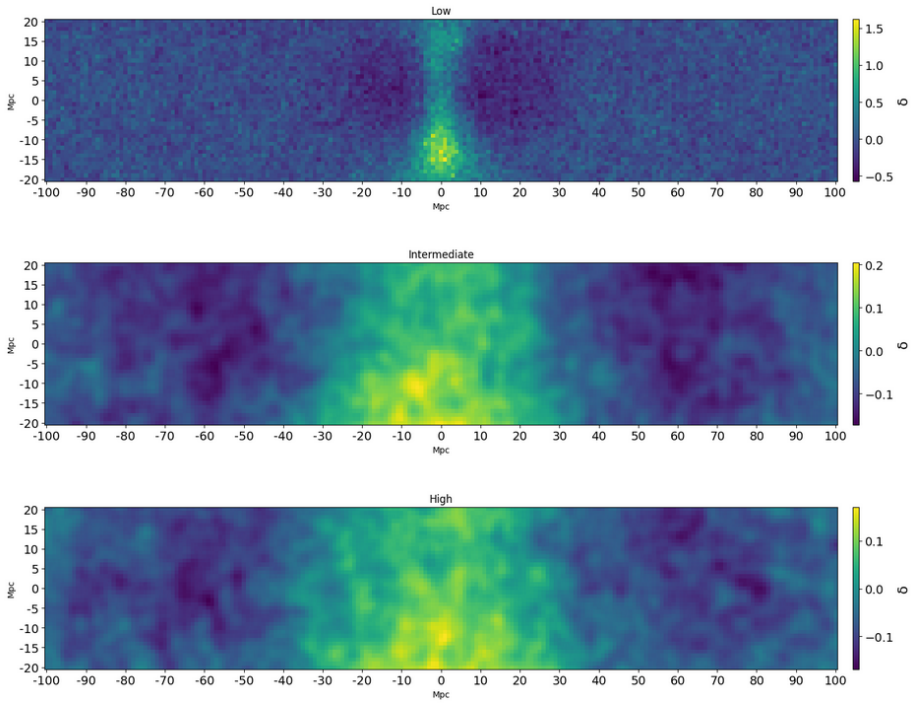



Figure 4.8: Density contrast of halos stacked at the position of filaments reconstructed on low, intermediate, and high redshift slices smoothed with a Gaussian kernel of FWHM = 3 degrees, from top to bottom respectively.

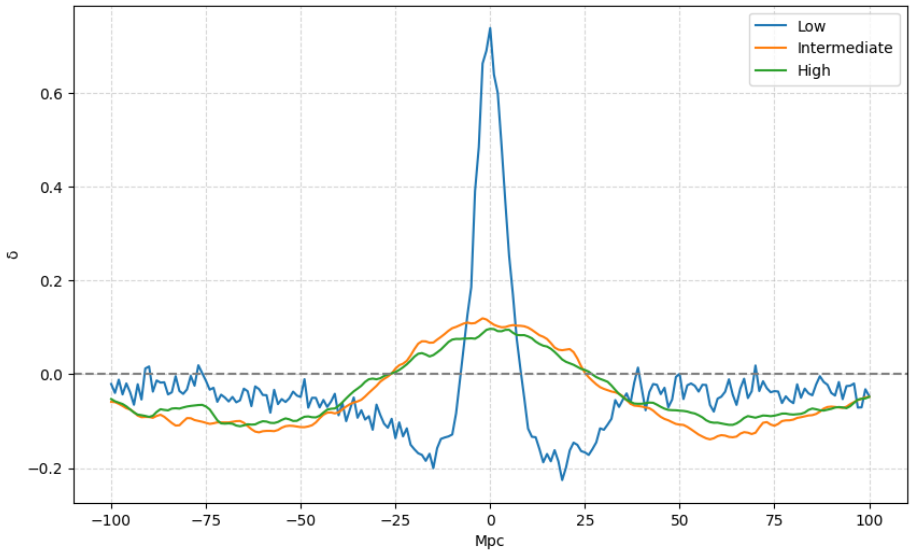



Figure 4.9: Stacked profile of filaments reconstructed on low, intermediate, and high redshift slices smoothed with a Gaussian kernel of FWHM = 3 degrees from top to bottom respectively.

### 4.2.4 Distance between DM Halos and Filaments

Filaments trace the overdense regions in the dark matter halo distribution map. Thus, we expect the distance of halos from the filaments to indicate this correlation. We can use this as a figure of merit to check the accuracy of the filaments reconstructed with the SCMS method. We estimate the distance between the halos and the reconstructed filament using the method explained in Section 3.6. In Figure 4.10 and Figure 4.11, we plot the histogram of the normalised distribution of distances in degrees from DM halos to filaments in the case of intermediate redshift slice smoothed with different kernel sizes and for all redshift slices in sample A with the same kernel size of 3 degrees, respectively.

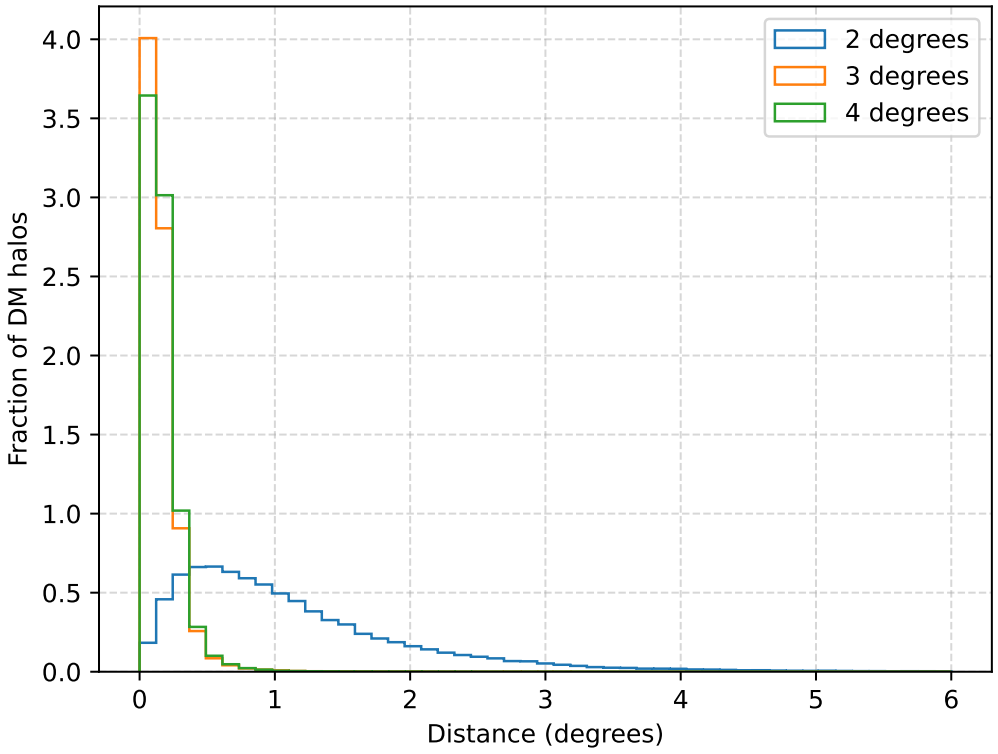



Figure 4.10: The normalised distribution of distances between halos and reconstructed filaments on the intermediate redshift slice smoothed with Gaussian kernel with FWHM = 2, 3, and 4 degrees.

We observe that the halos are dispersed from the centre of the filament in both cases. Although we expect some dispersion to be there naturally, there can be other factors adding to this, like error in estimating reconstructing filaments closer to the centre of the filament, which can be quantified to some extent with the uncertainty estimation for each case and the physical width of the filament. The higher dispersion in the low redshift slice in Figure 4.11 could be attributed to the sparse distribution

of halos and the higher dispersion in the case of the same intermediate slice smoothed with a lower kernel size in Figure 4.10 could be indicating that the 2 degree kernel is not the optimum kernel size for the intermediate slice to regularise the field.

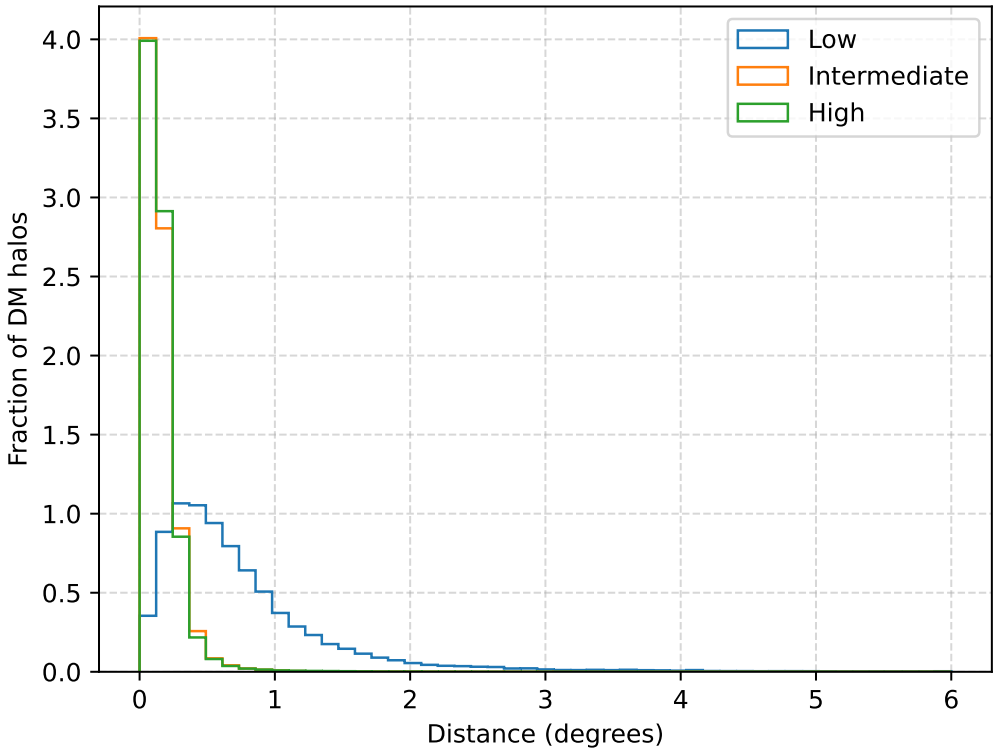



Figure 4.11: The normalised distribution of distances between halos and reconstructed filaments on the all redshift slices in sample A smoothed with a Gaussian kernel with FWHM = 3 degrees

## 4.3 Physical Scale as a choice of Kernel

In the previous section, we saw how increasing the kernel size in angular scales affected the filament reconstruction. Now, we investigate whether using a constant physical scale across all different redshift slices affects the reconstructed filaments and the stacked profile of the underlying density traced by the filaments as we saw for a fixed angular scale, filament width seemed to evolve with redshift as the stacked profile became narrower with decreasing redshift. We start with a fixed physical scale in the range of 5-10 Mpc for smoothing, as these scales are smaller than the fwhm of the stacked profile obtained from the previous case.

We notice that the filament reconstruction is very poor with these smoothing scales at low redshift slices; as the kernel size in angular scale falls much below 1 degree in size. It seems that the angular scale less than 1 degree at low redshift is not sufficient enough to regularise the field; therefore, the estimated projection vector magnitudes which determines how much a grid point moves with each iteration of the SCMS varies significantly ranging from $0 - 4$ radians or $0 - 200$ degrees, see Figure 4.12, resulting in the non-convergence of the grid points even after hundreds of iterations.

The Figure 4.12 shows how big the projection vectors are estimated at every point on the map of low redshift slice smoothed with a Gaussian kernel of FWHM $= 4\ Mpc$ $= 0.75$ degree. The value ranges between 0 and 4 radians, which is not ideal for filament reconstruction. For point convergence, as we have observed in the cases where filament reconstruction works well, we typically have most of the magnitudes of the projection vectors of the order of 1 degree (0.02 radians); see Figure 4.13.

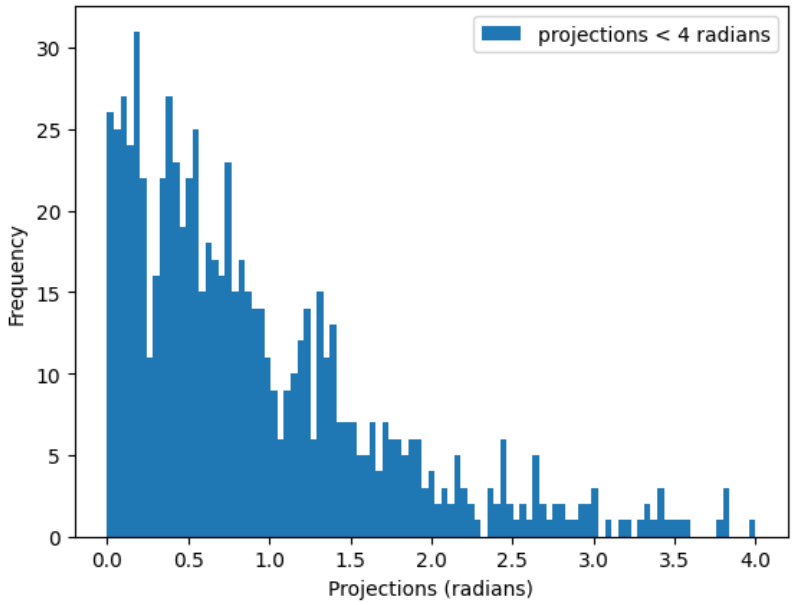



Figure 4.12: Histogram plot showing the normalised distribution of magnitudes of the estimated projections lesser than 4 radians ($\sim$ 230 degree) for the low redshift slice smoothed with a Gaussian kernel of FWHM $= 4\ Mpc = 0.75$ degree.

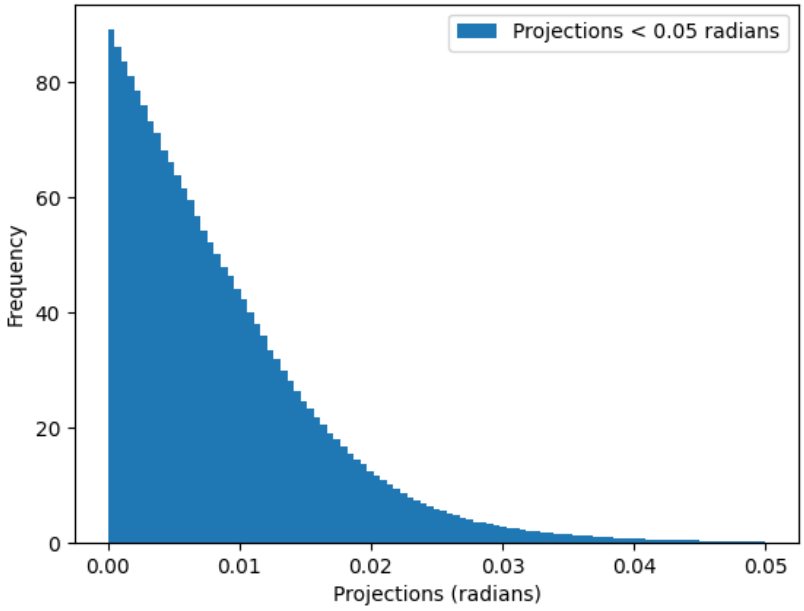



Figure 4.13: Histogram plot showing the normalised distribution of magnitudes of the estimated projections lesser than 0.05 radians ($\sim$ 2.9 degree) for the low redshift slice smoothed with a Gaussian kernel of FWHM = 13.60 $Mpc$ = 2 degree.

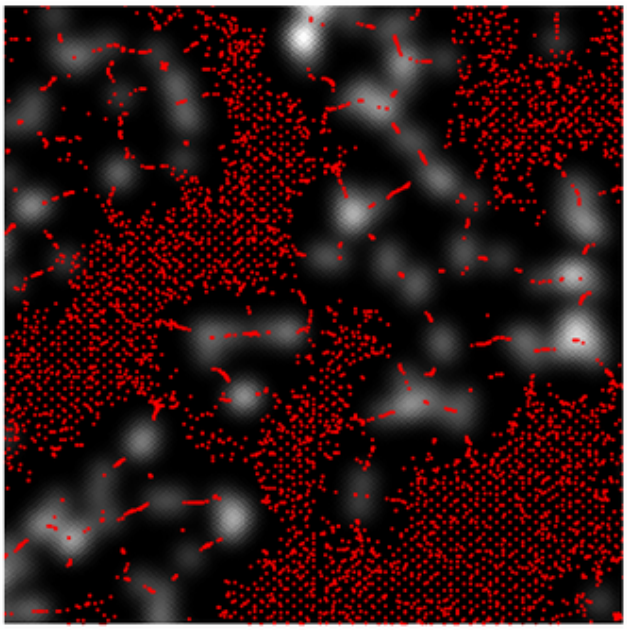


Figure 4.14: Non-convergence of the grid points after 200 iterations on the low redshift slice for smoothing kernel scale of 0.75 degrees. All patches is 1000 $\times$ 1000 pixels with the resolution of 1 arc minute per pixel and centred at RA = 0, Dec = 0

In Figure 4.14, it is clearly visible that most of the grid points do not converge in the overdense regions, for the low redshift slice smoothed with a Gaussian kernel of FWHM = 4 $Mpc$ = 0.75 degree even after 100 iterations of the SCMS algorithm.

We will see how this issue of non-convergence of grid points on maps smoothed with low physical scale can be resolved with the Two-step Filament reconstruction (2sFR) method we propose in the next section.

## 4.4 Two-Step Filament Reconstruction Method

We perform the Two-step Filament reconstruction (2sFR) method on the low redshift slice. In the first step, we perform the SCMS algorithm for 25 iterations on a map smoothed with a larger Gaussian kernel with FWHM = 20 $Mpc = 3$ *degrees* to initiate the convergence of the points towards the overdense regions. Then we perform the SCMS for 100 iterations method again with the output of the previous step as the initial positions of the grid points on the original map smoothed with the smaller Gaussian kernel with FWHM = 5 $Mpc = 0.75$ *degrees*. We see that most of the grid points have converged at the overdense regions of map after a total number of 125 iterations of 2sFR method, see Figure 4.15, except for a few outlier points which have not converged even after the first step, this could be due to the lesser number of halos in the low redshift slice, affecting the filament reconstruction in general. This is a massive improvement compared to the filament construction using the SCMS with a single smaller Gaussian kernel with FWHM = 0.75 $Mpc$, as we saw earlier in Figure 4.14.

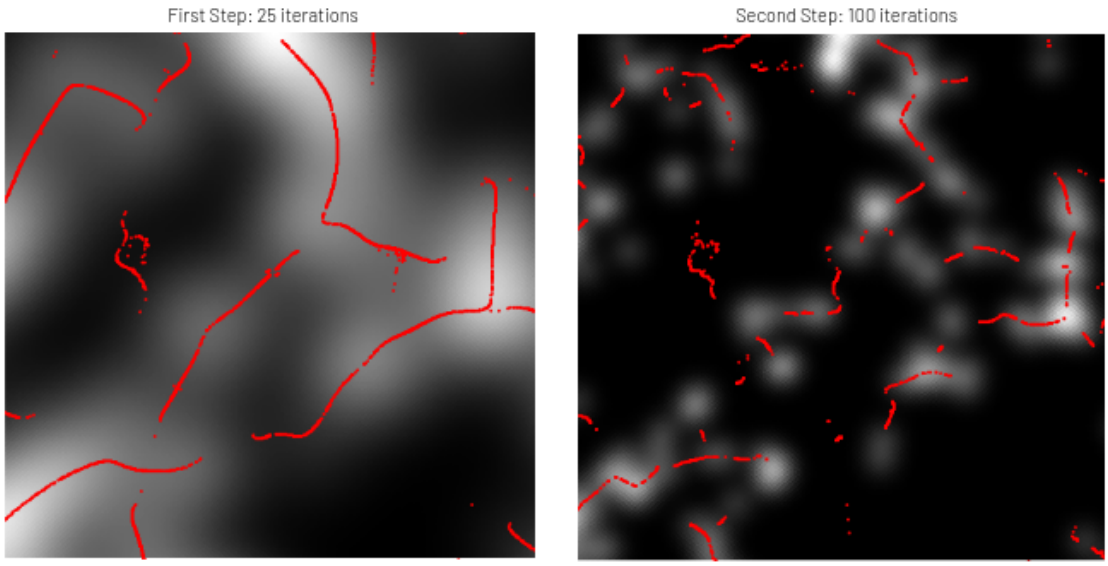



Figure 4.15: Filament reconstruction on the low redshift slice using the 2sFR method. *Left*: The position of the grid points after the first step of 25 iterations with larger smoothing kernel with FWHM = 20 $Mpc = 3$ *degrees*. *Right*: The position of the grid points after the second step of 100 iterations with smaller smoothing kernel scale with FWHM = 5 $Mpc = 0.75$ *degrees*. All patches are of the size 1000 × 1000 pixels with the resolution of 1 arc minute per pixel and centred at RA = 0, Dec = 0

For studying the filament reconstruction using the Gaussian kernel of fixed physical scale across all the redshift slices in sample B using the two-step filament reconstruction method or 2sFR-method. We take a larger Gaussian kernel with FWHM = 60 $Mpc$ for the first step of 25 iterations and a Gaussian kernel with FWHM = 30 $Mpc$ for the second step of 100 iterations. The lower scale is chosen as such so that

we can reconstruct filaments using the original SCMS method at all redshift slices without any issue of convergence, so that we can perform a direct comparison between the two methods. The Table 4.4 shows the corresponding angular scales $\theta_{30Mpc}$ and $\theta_{60Mpc}$ at each redshift slice in sample B for the 30 and 60 $Mpc$ physical scales, respectively.

| Redshift Slices | $\theta_{30Mpc}$ (degrees) | $\theta_{60Mpc}$ (degrees) |
|---|---|---|
| Slice 1 | 4.41 | 8.82 |
| Slice 2 | 2.26 | 4.52 |
| Slice 3 | 1.63 | 3.26 |
| Slice 4 | 1.33 | 2.66 |
| Slice 5 | 1.17 | 2.34 |
| Slice 6 | 1.07 | 2.13 |
| Slice 7 | 1.01 | 2.02 |
| Slice 8 | 0.97 | 1.95 |
| Slice 9 | 0.96 | 1.92 |
| Slice 10 | 0.96 | 1.92 |
| Slice 11 | 0.97 | 1.95 |

Table 4.4: Angular size (degrees) subtended by 30 Mpc and 60 Mpc physical scale at each redshift slices in sample B.

We compare the reconstructed filaments using the SCMS algorithm in the original form (Method A) for 200 iterations and the 2sFR method (Method B) for 125 iterations in all redshift slices in sample B smoothed with a Gaussian kernel with FWHM = 30 $Mpc$ physical scale by visually checking the filament reconstruction, analysing the estimated uncertainty, stacking profile, and the distance between halos and the reconstructed filaments for each method.

In Figure 4.16, we look at the filaments reconstructed in redshift slices, Slice 1, Slice 7, and Slice 11 of sample B from top to bottom, respectively, using Method A (on the left) and Method B (on the right). We can clearly notice that the filament reconstruction is better using Method B, as more points are converged at the overdense regions, to more finer structures compared to Method A. We also notice that there are more points in the underdense regions in the case of Method A which have not converged even after 200 iterations of the SCMS algorithm.

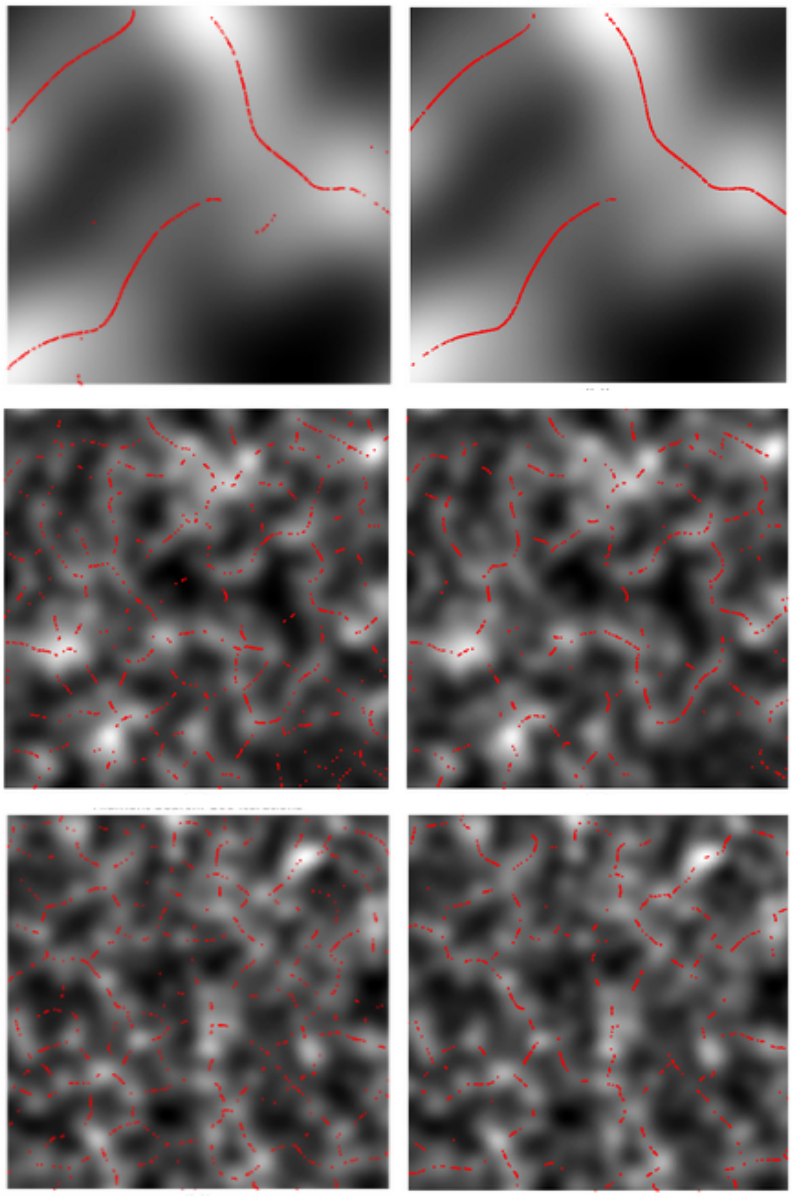
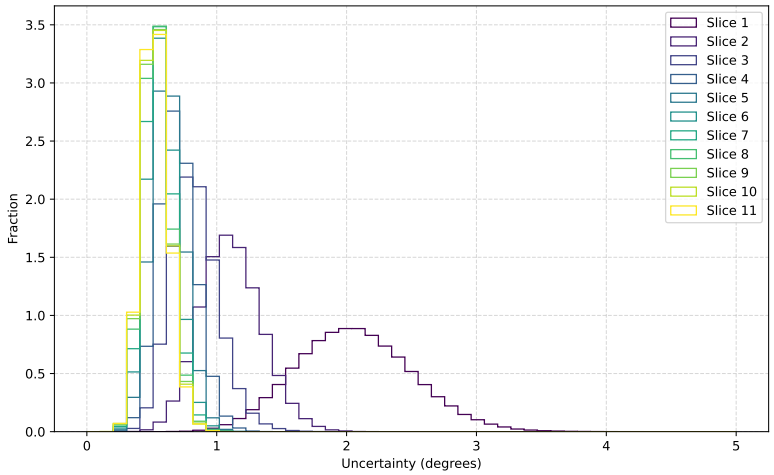


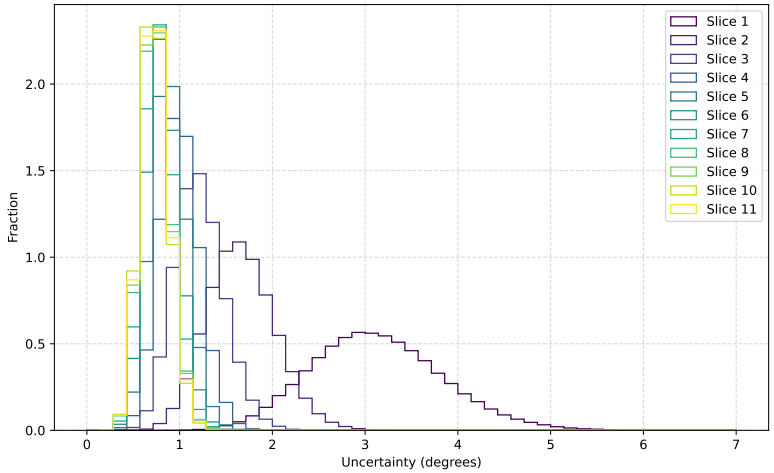
Figure 4.16: Filament reconstruction on the redshift slices, Slice 1, Slice 7, and Slice 11 of sample B from top to down respectively using the original SCMS method and the 2sFR method. *Left*: Method A. *Right*: Method B. All patches are of the size $1000 \times 1000$ pixels with the resolution of 1 arc minute per pixel and centred at RA $= 0$, Dec $= 0$

### Uncertainty

Figure 4.17 shows the histogram plot of the normalised distribution of uncertainties in degrees estimated for the filament reconstruction using the Method A and Method B in all the redshift slices in sample B.




(a) Method A




(b) Method B

Figure 4.17: The normalised distribution of uncertainties in degrees for all redshift slices in sample B.

The estimated uncertainties are larger for Method B, which is evident in the case of all the redshift slices. This could be because there are more points that have not converged and lie in the underdense regions of the map in the case of

Method A, as evident from Figure 4.17, this leads to false detection of filaments and reduces the uncertainty estimated by reducing the distance between true filaments and simulated filaments in the bootstrap sample of the original map while performing the SCMS uncertainty algorithm. In Figure 4.18, we can clearly see this difference in the estimated uncertainties for both methods with the mean uncertainty estimated as a function of the redshift.

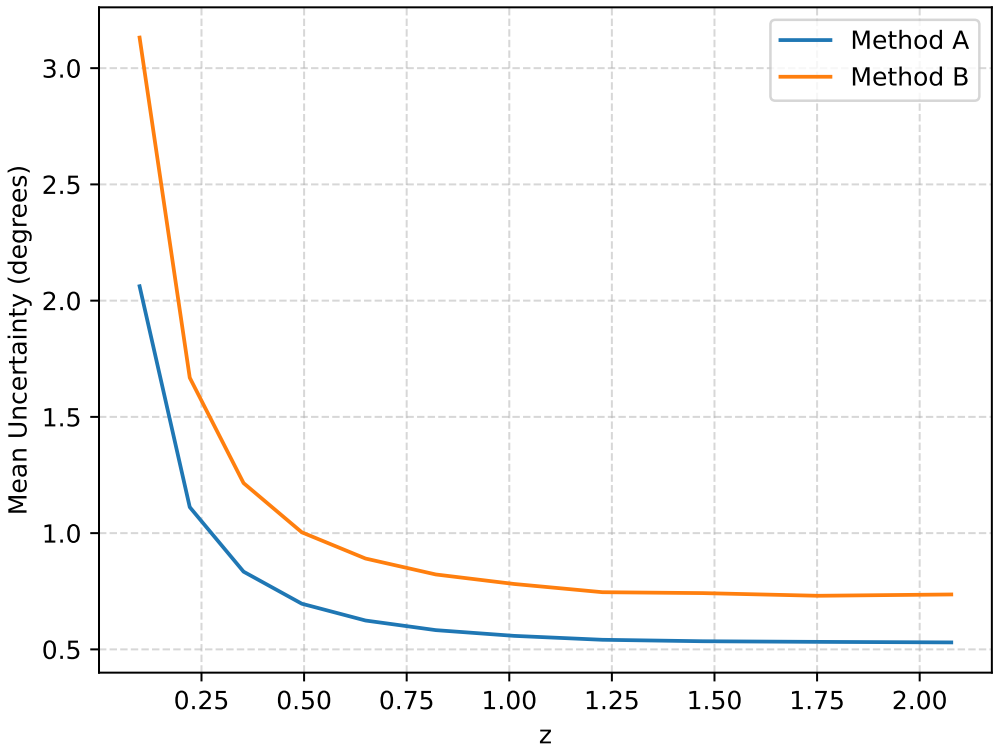
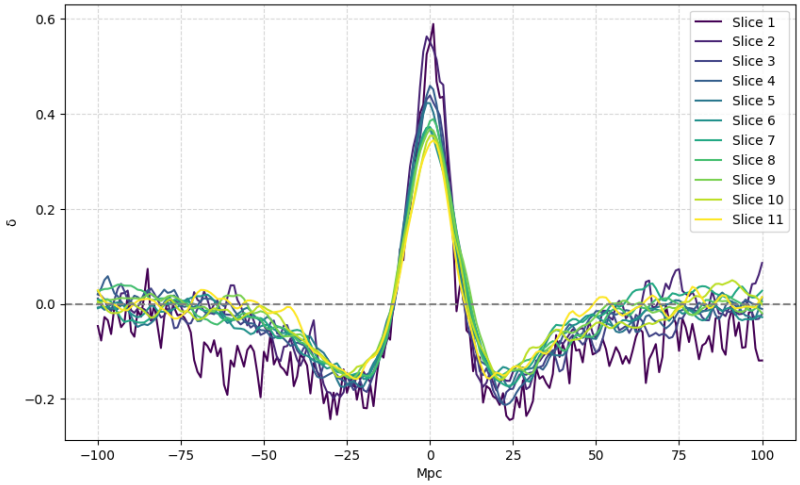



Figure 4.18: Mean uncertainty estimated as function of the redshift for Method A and Method B.

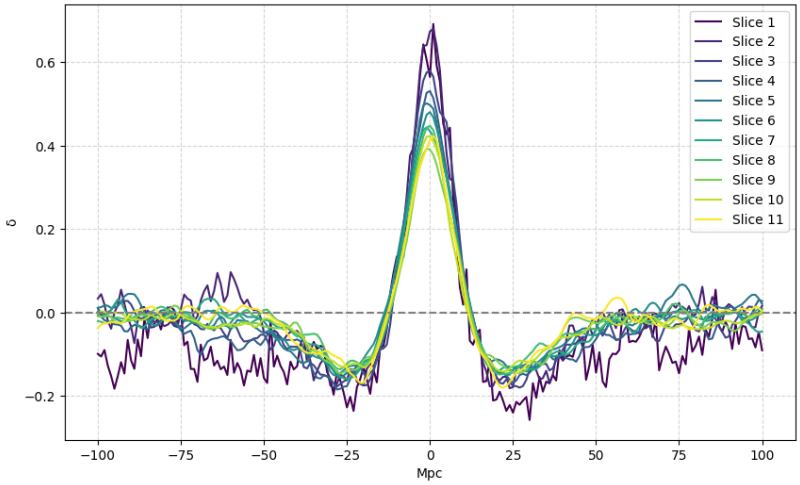
### Stacking Profile

From the filament profiles derived by stacking the halo density contrast traced by the filaments in all the redshift slices of sample B for both methods, see Figure 4.19, we can say that the filament width remains largely unaffected by the choice of a constant physical scale compared to the use of a constant angular scale for smoothing across the redshift range. This could be either suggesting that there is no evolution in filament width for maps smoothed with Gaussian kernels of constant physical scale or that we are still oversmoothing all the redshift slices at 30 $Mpc$ scale. However, the stacking profile peak is greater for all the slices for method B compared to method A, indicating that the 2sFR method is better at the convergence of grid points closer to the centre of the overdense regions. The Figure 4.19 also shows that their is an

evolution of the amplitude of the stacked profile density contrast of dark matter halos at the position of filaments with redshift for both methods. This is expected as ΛCDM cosmology predicts the filaments to be more dense at low redshift compared to higher redshift.




(a) Method A

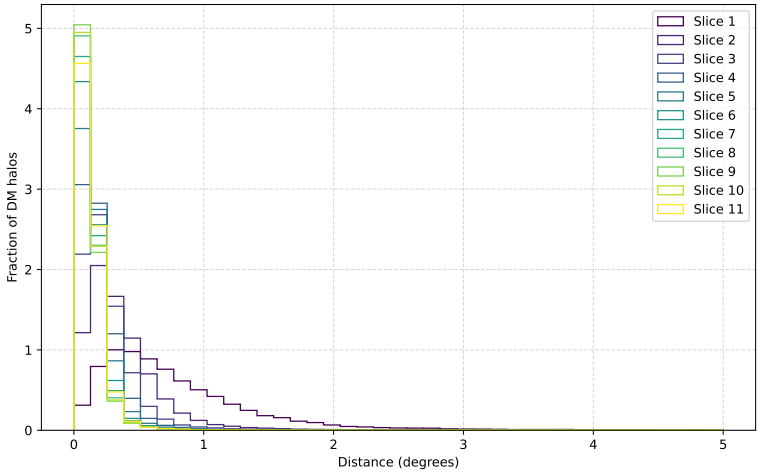



(b) Method B

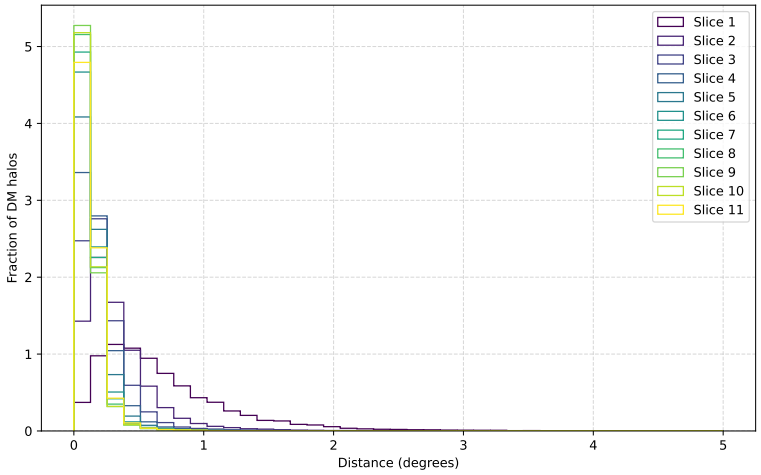
Figure 4.19: Stacking Profile of density contrast of dark matter halos at the positions of filaments for all the slices in Sample B.

### Distance between DM Halos and Filaments

The normalised distribution of the distance between DM halos to the closest reconstructed filaments looks very similar at first for both methods for all the redshift slices in sample B, see Figure 4.20. A closer look would tell us that there are more DM halos closer to the reconstructed filaments in Method B than in Method A. Thus, affirming what we already inferred from the reconstruction of filaments visually and the stacking profile results, the 2sFR method is better at reconstructing filaments at overdense regions than the original SCMS method.




(a) Method A




(b) Method B

Figure 4.20: Histogram of normalised distribution of the distance between DM halos to the closest reconstructed filament for all the redshift slices in sample B.

This is more clear in Figure 4.21, which shows the mean distance estimated between the DM halos to the closest reconstructed filament as a function of the redshift for Methods A and B. We can see that Method B is slightly better than Method A at all redshift slices smoothed with a Gaussian kernel of fixed physical scale of 30 $Mpc$.

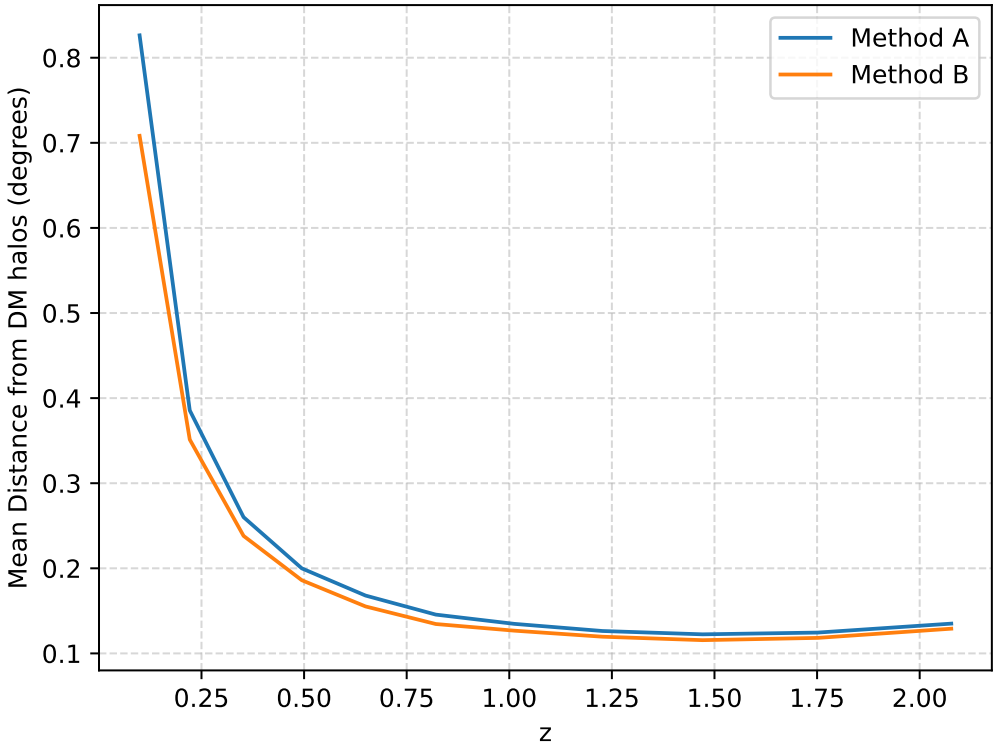
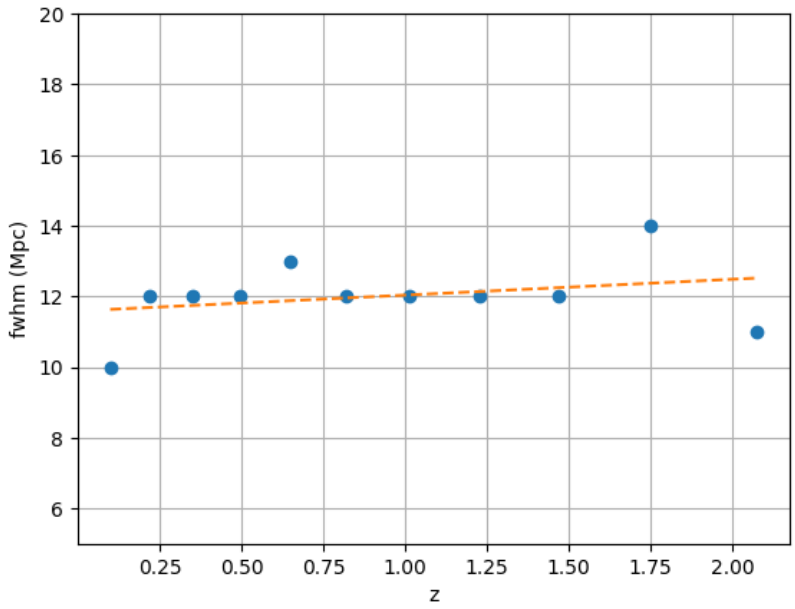


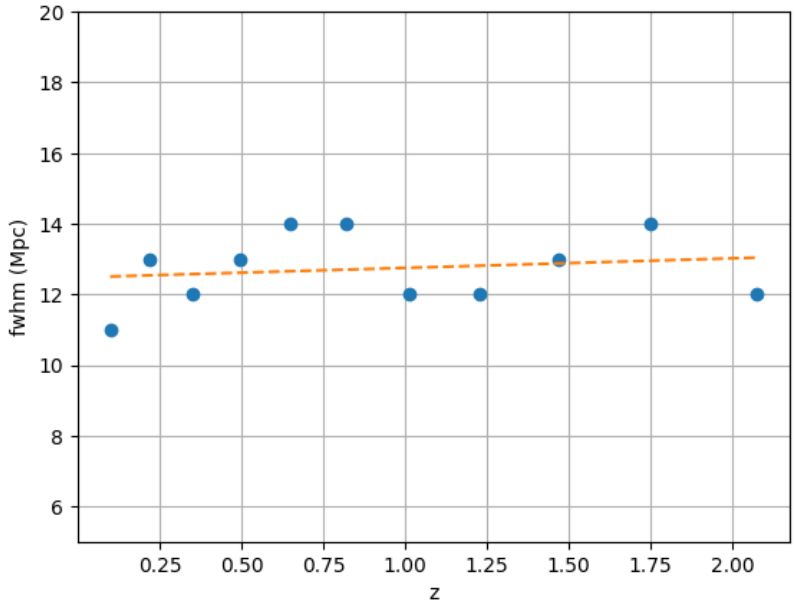

Figure 4.21: Mean distance estimates between the DM halos to the closest reconstructed filament as a function of the redshift for Method A and Method B.

## 4.5 Filament Width for constant physical scale

The fwhm of the stacked profile can be used as a proxy to study the width of the filaments on average at each redshift slice. We estimate the fwhm of the stacking profiles of all redshift slices in sample B for both methods, as seen in Figure 4.22. The FWHM of the stacking profile is consistent across different redshift slices for the constant physical smoothing kernel for both methods. This is in contrast to what we observed earlier with the case of constant angular scale across different redshift slices, suggesting that it could be an effect of varying smoothing at different redshift or with the constant physical scale of 30 $Mpc$ is not the optimum scale for smoothing for all the redshift slices in the sample.




(a) Method A




(b) Method B

Figure 4.22: FWHM of stacking profiles of reconstructed filaments at different redshift slices.

The average fwhm is higher for Method B than for Method A. This could also be a result of the better filament reconstruction with the 2sFR method, as we might be stacking a few points which have not converged to the overdense regions yet in the case of the original SCMS method. Plotting the FWHM as a function of redshift, we noticed that it seems to follow a linear trend. Thus, a linear fit was applied to both

distributions. The parameters obtained for both fits are as follows:

- Method A: slope = 0.45, y-intercept = 11.58
- Method B: slope = 0.27, y-intercept = 12.47

This also suggests that filament width is possibly being dominated by uncertainty in filament reconstruction due to oversmoothing with the physical scale of 30 $Mpc$ in all redshift slices in sample B.

# Chapter 5

# Conclusions and Future Scope

Current and upcoming large-scale galaxy surveys will provide huge amounts of data, so it is of great relevance to develop new and accurate methods to constrain the cosmological model. The state-of-the-art cosmological simulations provide us with a plethora of information about the evolution of the Universe. They can be used to test the impact of varying parameters of the models on the different observables and therefore to test the effectiveness of the data analysis methods in detecting those.

In this work, we explore the LSS of the Universe, especially the cosmic filaments. We study cosmic filament reconstruction using the SCMS algorithm based on the ridge formalism as proposed in Chen et al. (2015) in the DEMNUni simulations, which we choose with the larger aim in mind of studying the effects of neutrino mass in the cosmic filaments in the future. The first goal was to correctly implement the algorithm. For this, we reconstruct filaments in 2D maps of the distribution of dark matter (DM) halos at different redshift slices between the range $0.0968 < z < 2.0842$ extracted from the DEMNUni N-body simulations with $\Lambda$CDM model and zero neutrino mass. We initially verified the implementation of the SCMS algorithm by comparing the distribution of the underlying halo density at the positions traced by the reconstructed filaments with that of the entire input map. In the case of the SCMS algorithm, a critical step is represented by the smoothing process of the input tracer map. We looked at how the filament reconstruction is affected by the kernel of our choice used for smoothing, which is a free parameter of the SCMS algorithm. Throughout the thesis, we apply Gaussian smoothing kernels with varying FWHM. We checked the accuracy of the filament reconstruction using several methods. First, by estimating the uncertainty of reconstructed filaments for each redshift slice using bootstrap samples of the original distribution of dark matter halos. Second, by stacking the halo density contrast at the position of the reconstructed filaments, and finally by computing the distance between halos from the reconstructed filaments. We observed a broadening of the stacked profiles of the halo density contrast traced by the filaments when we increased the Gaussian smoothing kernel sizes (in angular scales) for low, intermediate, and high redshift slices in sample A, implying that we possibly oversmooth the dark-matter halos distribution map with increasing the kernel size, which is causing the density to be spread out compared to being concentrated at the centre of the filament. Upon using the same angular scale as a smoothing kernel for the

above three redshift slices, we observed the broadening of the stacking profile with increasing redshift, suggesting a possible evolution of filaments through the cosmic history as halos move towards the centre of the filaments, making them narrower.

The next step was to test the filament reconstruction using the same physical scale as a smoothing kernel for all the redshift slices in sample B. But this came with its own challenges, as we noticed that the SCMS algorithm fails to reconstruct filaments in density maps smoothed with small physical scales. This poses a regularisation problem; the estimated magnitude of the projection vectors in the SCMS algorithm, which determines how much a single grid point moves with each step of the iteration, varies significantly. The magnitude of these projection vectors ranges very widely from $0 - 10^2$ degrees, which should typically be in the range of $0 - 0.02$ radians for good convergence; this raises a difficulty in normalizing the projection vectors. This has led to the development of the two-step Filament Reconstruction method (2sFR method), a novel method of filament reconstruction using the SCMS algorithm but with two different smoothing kernel scales. We apply the SCMS algorithm on a density map smoothed with a larger physical scale for $10 - 25$ iterations and then move to the same density map smoothed with the desired smaller physical scale and perform the iterative steps of the SCMS using the output of the first $10 - 25$ steps as the initial grid point for the next 100 iterations to convergence.

For comparison with the original SCMS algorithm, we choose the 30 *Mpc* and 60 *Mpc* as our small and large constant physical scales as two kernel scales for the 2sFR method for smoothing across all redshift slices. The lower scale is chosen in such a way that the SCMS algorithm can also work on density maps smoothed with that kernel scale, so that we can compare the results of both methods. The grid points converge to the overdense regions better with the 2sFR method compared to the original SCMS method. Although the uncertainty was found to be larger in the case of 2sFR method, this could be because there are more points that have not converged and lie in the underdense regions of the map in the case of the original SCMS method, which leads to false detection of filaments and reduces the uncertainty estimated by reducing the distance between true filaments and simulated filaments in the bootstrap sample of the original map while performing the SCMS uncertainty algorithm. The peak of the stacked profile shows that the 2sFR method reconstructs filaments in regions closer to the centre of the overdense regions, and the halos are closer to the filaments reconstructed with the 2sFR method than the SCMS method. We estimate the fwhm of the stacked profile for each redshift slice and use it as a proxy to study the filament width as a function of the redshift. Unlike the results we get with a constant angular scale for smoothing, in the case of a fixed physical scale, we find that the filament width does not vary significantly. It stays almost constant, possibly indicating that we are not using the optimal kernel size to smooth the density maps. However, the peak of the stacked profile does change, increasing at lower redshift, implying that halos are closer to the centre of the filament at lower redshift, as predicted by the $\Lambda$CDM model.

The thesis work has made a few steps forward in the direction of filament reconstruction using the SCMS algorithm and has opened many more avenues to explore in the future. From the perspective of the algorithm, there can be further improve-

ments made in the direction of optimising the smoothing kernel size to be used across redshift slices and the number of the grid points used to search for filaments can be optimised for both faster computation and better reconstruction of filaments.

As mentioned above, a future application of this work is to study the effects of neutrino mass on the evolution of cosmic filaments. The necessity of an optimised algorithm for correctly measuring the cosmic filament profile and morphological properties is of utmost importance for the cosmological inference. Another possibility is to use the modified ridge formalism to extend the SCMS method to detect other structures of the cosmic web.

# Bibliography


Shadab Alam, Marie Aubert, Santiago Avila, and et al. Completed sdss-iv extended baryon oscillation spectroscopic survey: Cosmological implications from two decades of spectroscopic surveys at the apache point observatory. *Physical Review D*, 103(8), April 2021. ISSN 2470-0029. doi:10.1103/physrevd.103.083533. URL http://dx.doi.org/10.1103/PhysRevD.103.083533.

Mehmet Alpaslan, Aaron S. G. Robotham, Simon Driver, Peder Norberg, Ivan Baldry, Amanda E. Bauer, Joss Bland-Hawthorn, Michael Brown, Michelle Cluver, Matthew Colless, Caroline Foster, Andrew Hopkins, Eelco Van Kampen, Lee Kelvin, Maritza A. Lara-Lopez, Jochen Liske, Angel R. Lopez-Sanchez, Jon Loveday, Tamsyn McNaught-Roberts, Alexander Merson, and Kevin Pimbblet. Galaxy And Mass Assembly (GAMA): the large-scale structure of galaxies and comparison to mock universes. MNRAS, 438(1):177–194, February 2014. doi:10.1093/mnras/stt2136.

Raul E. Angulo and Oliver Hahn. Large-scale dark matter simulations. *Living Reviews in Computational Astrophysics*, 8(1):1, December 2022. doi:10.1007/s41115-021-00013-z.

M. A. Aragon-Calvo. Classifying the large-scale structure of the universe with deep neural networks. MNRAS, 484(4):5771–5784, April 2019. doi:10.1093/mnras/stz393.

M. A. Aragón-Calvo, B. J. T. Jones, R. van de Weygaert, and J. M. van der Hulst. The multiscale morphology filter: identifying and extracting spatial patterns in the galaxy distribution. A&A, 474(1):315–338, October 2007. doi:10.1051/0004-6361:20077880.

Miguel A. Aragón-Calvo, Rien Van De Weygaert, and Bernard J. T. Jones. Multiscale phenomenology of the cosmic web: Multiscale phenomenology of the cosmic web. *Monthly Notices of the Royal Astronomical Society*, 408(4):2163–2187, August 2010. ISSN 0035-8711. doi:10.1111/j.1365-2966.2010.17263.x. URL http://dx.doi.org/10.1111/j.1365-2966.2010.17263.x.

J. D. Barrow, S. P. Bhavsar, and D. H. Sonoda. Minimal spanning trees, filaments and galaxy clustering. MNRAS, 216:17–35, September 1985. doi:10.1093/mnras/216.1.17.

C. L. Bennett, R. S. Hill, G. Hinshaw, M. R. Nolta, N. Odegard, L. Page, D. N. Spergel, J. L. Weiland, E. L. Wright, M. Halpern, N. Jarosik, A. Kogut, M. Limon, S. S. Meyer, G. S. Tucker, and E. Wollack. First-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Foreground Emission. ApJS, 148(1): 97–117, September 2003. doi:10.1086/377252.

J. Richard Bond, Lev Kofman, and Dmitry Pogosyan. How filaments of galaxies are woven into the cosmic web. Nature, 380(6575):603–606, April 1996. doi:10.1038/380603a0.

Brandon Buncher, Awshesh Nath Sharma, and Matias Carrasco Kind. Survey2survey: a deep learning generative model approach for cross-survey image mapping. *Monthly Notices of the Royal Astronomical Society*, 503(1):777–796, February 2021. ISSN 1365-2966. doi:10.1093/mnras/stab294. URL http://dx.doi.org/10.1093/mnras/stab294.

Carmelita Carbone, Margarita Petkova, and Klaus Dolag. Demnuni: Isw, rees-sciama, and weak-lensing in the presence of massive neutrinos. *Journal of Cosmology and Astroparticle Physics*, 2016(07):034–034, July 2016. ISSN 1475-7516. doi:10.1088/1475-7516/2016/07/034. URL http://dx.doi.org/10.1088/1475-7516/2016/07/034.

Javier Carrón-Duque. *Stochastic Geometry Methods in Cosmology: Applications to Cosmic Microwave Background and Large Scale Structure Data*. PhD thesis, Sapienza University of Rome, 2022.

Emanuele Castorina, Carmelita Carbone, Julien Bel, Emiliano Sefusatti, and Klaus Dolag. Demnuni: the clustering of large-scale structures in the presence of massive neutrinos. *Journal of Cosmology and Astroparticle Physics*, 2015(07):043–043, July 2015. ISSN 1475-7516. doi:10.1088/1475-7516/2015/07/043. URL http://dx.doi.org/10.1088/1475-7516/2015/07/043.

Marius Cautun, Rien van de Weygaert, and Bernard J. T. Jones. NEXUS: tracing the cosmic web connection. MNRAS, 429(2):1286–1308, February 2013. doi:10.1093/mnras/sts416.

Marius Cautun, Rien van de Weygaert, Bernard J. T. Jones, and Carlos S. Frenk. Evolution of the cosmic web. *Monthly Notices of the Royal Astronomical Society*, 441(4):2923–2973, May 2014. ISSN 0035-8711. doi:10.1093/mnras/stu768. URL http://dx.doi.org/10.1093/mnras/stu768.

Yen-Chi Chen, Shirley Ho, Peter E. Freeman, Christopher R. Genovese, and Larry Wasserman. Cosmic web reconstruction through density ridges: method and algorithm. *Monthly Notices of the Royal Astronomical Society*, 454(1):1140–1156, September 2015. ISSN 1365-2966. doi:10.1093/mnras/stv1996. URL http://dx.doi.org/10.1093/mnras/stv1996.

G. Chincarini and H. J. Rood. Size of the Coma cluster. Nature, 257(5524):294–295, September 1975. doi:10.1038/257294a0.

J. M. Colberg, K. S. Krughoff, and A. J. Connolly. Intercluster filaments in a cdm universe. *Monthly Notices of the Royal Astronomical Society*, 359(1):272–282, May 2005. ISSN 1365-2966. doi:10.1111/j.1365-2966.2005.08897.x. URL http://dx.doi.org/10.1111/j.1365-2966.2005.08897.x.

Matthew Colless, Bruce A. Peterson, Carole Jackson, John A. Peacock, Shaun Cole, Peder Norberg, Ivan K. Baldry, Carlton M. Baugh, Joss Bland-Hawthorn, Terry Bridges, Russell Cannon, Chris Collins, Warrick Couch, Nicholas Cross, Gavin Dalton, Roberto De Propris, Simon P. Driver, George Efstathiou, Richard S. Ellis, Carlos S. Frenk, Karl Glazebrook, Ofer Lahav, Ian Lewis, Stuart Lumsden, Steve Maddox, Darren Madgwick, Will Sutherland, and Keith Taylor. The 2dF Galaxy Redshift Survey: Final Data Release. *arXiv e-prints*, art. astro-ph/0306581, June 2003. doi:10.48550/arXiv.astro-ph/0306581.

M. Davis, G. Efstathiou, C. S. Frenk, and S. D. M. White. The evolution of large-scale structure in a universe dominated by cold dark matter. ApJ, 292:371–394, May 1985. doi:10.1086/163168.

W De Boer. Grand unified theories and supersymmetry in particle physics and cosmology. *Progress in Particle and Nuclear Physics*, 33:201–301, January 1994. ISSN 0146-6410. doi:10.1016/0146-6410(94)90045-0. URL http://dx.doi.org/10.1016/0146-6410(94)90045-0.

Anna de Graaff, Yan-Chuan Cai, Catherine Heymans, and John A. Peacock. Probing the missing baryons with the sunyaev-zel'dovich effect from filaments. *Astronomy amp; Astrophysics*, 624:A48, April 2019. ISSN 1432-0746. doi:10.1051/0004-6361/201935159. URL http://dx.doi.org/10.1051/0004-6361/201935159.

V. de Lapparent, M. J. Geller, and J. P. Huchra. A Slice of the Universe. ApJ, 302: L1, March 1986. doi:10.1086/184625.

Saee Dhawalikar and Aseem Paranjape. Towards unbiased recovery of cosmic filament properties: the role of spine curvature and optimized smoothing, 2024. URL https://arxiv.org/abs/2402.18669.

Scott Dodelson. *Modern Cosmology*. Academic Press, Amsterdam, 2003. ISBN 978-0-12-219141-1.

K. Dolag, S. Borgani, S. Schindler, A. Diaferio, and A. M. Bykov. Simulation Techniques for Cosmological Simulations. Space Sci. Rev., 134(1-4):229–268, February 2008. doi:10.1007/s11214-008-9316-5.

K. Dolag, S. Borgani, G. Murante, and V. Springel. Substructures in hydrodynamical cluster simulations. MNRAS, 399(2):497–514, October 2009. doi:10.1111/j.1365-2966.2009.15034.x.

A. G. Doroshkevich, E. V. Kotok, A. N. Poliudov, S. F. Shandarin, Iu. S. Sigov, and I. D. Novikov. Two-dimensional simulation of the gravitational system dynamics and formation of the large-scale structure of the universe. MNRAS, 192:321–337, August 1980. doi:10.1093/mnras/192.2.321.

S. P. Driver, D. T. Hill, L. S. Kelvin, A. S. G. Robotham, J. Liske, P. Norberg, I. K. Baldry, S. P. Bamford, A. M. Hopkins, J. Loveday, J. A. Peacock, E. Andrae, J. Bland-Hawthorn, S. Brough, M. J. I. Brown, E. Cameron, J. H. Y. Ching, M. Colless, C. J. Conselice, S. M. Croom, N. J. G. Cross, R. de Propris, S. Dye, M. J. Drinkwater, S. Ellis, Alister W. Graham, M. W. Grootes, M. Gunawardhana, D. H. Jones, E. van Kampen, C. Maraston, R. C. Nichol, H. R. Parkinson, S. Phillipps, K. Pimbblet, C. C. Popescu, M. Prescott, I. G. Roseboom, E. M. Sadler, A. E. Sansom, R. G. Sharp, D. J. B. Smith, E. Taylor, D. Thomas, R. J. Tuffs, D. Wijesinghe, L. Dunne, C. S. Frenk, M. J. Jarvis, B. F. Madore, M. J. Meyer, M. Seibert, L. Staveley-Smith, W. J. Sutherland, and S. J. Warren. Galaxy and Mass Assembly (GAMA): survey diagnostics and core data release. MNRAS, 413(2):971–995, May 2011. doi:10.1111/j.1365-2966.2010.18188.x.

K. L. Dutson, R. J. White, A. C. Edge, J. A. Hinton, and M. T. Hogan. A stacked analysis of brightest cluster galaxies observed with the Fermi Large Area Telescope. MNRAS, 429(3):2069–2079, March 2013. doi:10.1093/mnras/sts477.

David H. Eberly. *Ridges in Image and Data Analysis*, volume 7 of *Computational Imaging and Vision*. Springer, 1996. ISBN 978-90-481-4761-8. doi:10.1007/978-94-015-8765-5. URL https://doi.org/10.1007/978-94-015-8765-5.

G. F. R. Ellis and E. R. Harrison. Cosmological Principles I. Symmetry Principles. *Comments on Astrophysics and Space Physics*, 6:23–27, January 1974.

Euclid Collaboration. Euclid. I. Overview of the Euclid mission. *arXiv e-prints*, art. arXiv:2405.13491, May 2024. doi:10.48550/arXiv.2405.13491.

A. Friedman. Über die krümmung des raumes. *Zeitschrift für Physik*, 10:377–386, 1922. doi:10.1007/BF01332580.

Y. Fukuda, T. Hayakawa, E. Ichihara, and et al. Evidence for oscillation of atmospheric neutrinos. *Physical Review Letters*, 81(8):1562–1567, August 1998. ISSN 1079-7114. doi:10.1103/physrevlett.81.1562. URL http://dx.doi.org/10.1103/PhysRevLett.81.1562.

Daniela Galárraga-Espinosa, Nabila Aghanim, Mathieu Langer, Céline Gouin, and Nicola Malavasi. Populations of filaments from the distribution of galaxies in numerical simulations. *Astronomy & Astrophysics*, 641:A173, September 2020. ISSN 1432-0746. doi:10.1051/0004-6361/202037986. URL http://dx.doi.org/10.1051/0004-6361/202037986.

Margaret J. Geller and John P. Huchra. Mapping the universe. *Science*, 246:897–903, 1989. doi:10.1126/science.246.4932.897.

Christopher R. Genovese, Marco Perone-Pacifico, Isabella Verdinelli, and Larry Wasserman. Nonparametric ridge estimation. *arXiv e-prints*, art. arXiv:1212.5156, December 2012. doi:10.48550/arXiv.1212.5156.

Howard Georgi and S. L. Glashow. Unity of All Elementary-Particle Forces. Phys. Rev. Lett., 32(8):438–441, February 1974. doi:10.1103/PhysRevLett.32.438.

K. M. Górski and Eric Hivon. HEALPix: Hierarchical Equal Area isoLatitude Pixelization of a sphere. Astrophysics Source Code Library, record ascl:1107.018, July 2011.

J. Richard Gott, III, Mario Jurić, David Schlegel, Fiona Hoyle, Michael Vogeley, Max Tegmark, Neta Bahcall, and Jon Brinkmann. A Map of the Universe. ApJ, 624(2): 463–484, May 2005. doi:10.1086/428890.

S. A. Gregory, L. A. Thompson, and W. G. Tifft. The Perseus/Pisces Supercluster. In *Bulletin of the American Astronomical Society*, volume 10, page 622, September 1978.

Alan H. Guth. Inflationary universe: A possible solution to the horizon and flatness problems. Phys. Rev. D, 23(2):347–356, January 1981. doi:10.1103/PhysRevD.23.347.

Boryana Hadzhiyska, Simone Ferraro, and Rongpu Zhou. Tracing cosmic gas in filaments and halos: Low-redshift insights from the kinematic sunyaev-zel'dovich effect, 2024. URL https://arxiv.org/abs/2412.03631.

Nico Hamaus, P.M. Sutter, and Benjamin D. Wandelt. Universal density profile for cosmic voids. *Physical Review Letters*, 112(25), June 2014. ISSN 1079-7114. doi:10.1103/physrevlett.112.251302. URL http://dx.doi.org/10.1103/PhysRevLett.112.251302.

Nico Hamaus, Alice Pisani, Jin-Ah Choi, Guilhem Lavaux, Benjamin D. Wandelt, and Jochen Weller. Precision cosmology with voids in the final BOSS data. J. Cosmology Astropart. Phys., 2020(12):023, December 2020. doi:10.1088/1475-7516/2020/12/023.

S. W. Hawking and G. F. R. Ellis. *The Large Scale Structure of Space-Time*. Cambridge Monographs on Mathematical Physics. Cambridge University Press, 1973.

E. Hayashi, J. F. Navarro, and V. Springel. The shape of the gravitational potential in cold dark matter haloes. *Monthly Notices of the Royal Astronomical Society*, 377 (1):50–62, May 2007. ISSN 1365-2966. doi:10.1111/j.1365-2966.2007.11599.x. URL http://dx.doi.org/10.1111/j.1365-2966.2007.11599.x.

Chiaki Hikage and Kazuhiro Yamamoto. Fingers-of-god effect of infalling satellite galaxies. *Monthly Notices of the Royal Astronomical Society: Letters*, 455(1):L77–L81, 11 2015. ISSN 1745-3925. doi:10.1093/mnrasl/slv153. URL https://doi.org/10.1093/mnrasl/slv153.

John P. Huchra, Lucas M. Macri, Karen L. Masters, Thomas H. Jarrett, Perry Berlind, Michael Calkins, Aidan C. Crook, Roc Cutri, Pirin Erdoğdu, Emilio Falco, Teddy George, Conrad M. Hutcheson, Ofer Lahav, Jeff Mader, Jessica D. Mink, Nathalie Martimbeau, Stephen Schneider, Michael Skrutskie, Susan Tokarz, and Michael Westover. The 2MASS Redshift Survey—Description and Data Release. ApJS, 199(2):26, April 2012. doi:10.1088/0067-0049/199/2/26.

D. Heath Jones, Will Saunders, Matthew Colless, Mike A. Read, Quentin A. Parker, Fred G. Watson, Lachlan A. Campbell, Daniel Burkey, Thomas Mauch, Lesa Moore, Malcolm Hartley, Paul Cass, Dionne James, Ken Russell, Kristin Fiegert, John Dawe, John Huchra, Tom Jarrett, Ofer Lahav, John Lucey, Gary A. Mamon, Dominique Proust, Elaine M. Sadler, and Ken-ichi Wakamatsu. The 6dF Galaxy Survey: samples, observational techniques and the first data release. MNRAS, 355 (3):747–763, December 2004. doi:10.1111/j.1365-2966.2004.08353.x.

S. F. Kasun and A. E. Evrard. Shapes and Alignments of Galaxy Cluster Halos. ApJ, 629(2):781–790, August 2005. doi:10.1086/430811.

D. Keres, N. Katz, D. H. Weinberg, and R. Dave. How do galaxies get their gas? *Monthly Notices of the Royal Astronomical Society*, 363(1):2–28, October 2005. ISSN 1365-2966. doi:10.1111/j.1365-2966.2005.09451.x. URL http://dx.doi.org/10.1111/j.1365-2966.2005.09451.x.

Yigon Kim, Chan-Gyung Park, Hyerim Noh, and Jai-chan Hwang. Cmass galaxy sample and the ontological status of the cosmological principle. *Astronomy amp; Astrophysics*, 660:A139, April 2022. ISSN 1432-0746. doi:10.1051/0004-6361/202141909. URL http://dx.doi.org/10.1051/0004-6361/202141909.

Alexander Knebe, Stuart P. D. Gill, and Brad K. Gibson. Interactions of Satellite Galaxies in Cosmological Dark Matter Halos. PASA, 21(2):216–221, January 2004. doi:10.1071/AS04018.

Arthur Kosowsky, Milos Milosavljevic, and Raul Jimenez. Efficient cosmological parameter estimation from microwave background anisotropies. Phys. Rev. D, 66 (6):063007, September 2002. doi:10.1103/PhysRevD.66.063007.

Renée C. Kraan-Korteweg, Michelle E. Cluver, Maciej Bilicki, Thomas H. Jarrett, Matthew Colless, Ahmed Elagali, Hans Böhringer, and Gayoung Chon. Discovery of a supercluster in the Zone of Avoidance in Vela. MNRAS, 466(1):L29–L33, March 2017. doi:10.1093/mnrasl/slw229.

Christina D Kreisch, Alice Pisani, Carmelita Carbone, Jia Liu, Adam J Hawken, Elena Massara, David N Spergel, and Benjamin D Wandelt. Massive neutrinos leave fingerprints on cosmic voids. *Monthly Notices of the Royal Astronomical Society*, 488(3):4413–4426, July 2019. ISSN 1365-2966. doi:10.1093/mnras/stz1944. URL http://dx.doi.org/10.1093/mnras/stz1944.

Pavan Kumar Aluri, Paolo Cea, Pravabati Chingangbam, Ming-Chung Chu, Roger G Clowes, Damien Hutsemékers, Joby P Kochappan, Alexia M Lopez, Lang Liu, Niels C M Martens, C J A P Martins, Konstantinos Migkas, Eoin Ó Colgáin, Pratyush Pranav, Lior Shamir, Ashok K Singal, M M Sheikh-Jabbari, Jenny Wagner, Shao-Jiang Wang, David L Wiltshire, Shek Yeung, Lu Yin, and Wen Zhao. Is the observable universe consistent with the cosmological principle? *Classical and Quantum Gravity*, 40(9):094001, April 2023. ISSN 1361-6382. doi:10.1088/1361-6382/acbefc. URL http://dx.doi.org/10.1088/1361-6382/acbefc.

Abbé G. Lemaître. A homogeneous universe of constant mass and increasing radius accounting for the radial velocity of extra-galactic nebulæ. *Monthly Notices of the Royal Astronomical Society*, 91(5):483–490, 03 1931. ISSN 0035-8711. doi:10.1093/mnras/91.5.483. URL https://doi.org/10.1093/mnras/91.5.483.

J Lesgourgues and S Pastor. Massive neutrinos and cosmology. *Physics Reports*, 429 (6):307–379, July 2006. ISSN 0370-1573. doi:10.1016/j.physrep.2006.04.001. URL http://dx.doi.org/10.1016/j.physrep.2006.04.001.

Julien Lesgourgues and Sergio Pastor. Neutrino mass from cosmology. *Advances in High Energy Physics*, 2012:1–34, 2012. ISSN 1687-7365. doi:10.1155/2012/608515. URL http://dx.doi.org/10.1155/2012/608515.

Julien Lesgourgues and Sergio Pastor. Neutrino cosmology and planck. *New Journal of Physics*, 16(6):065002, June 2014. ISSN 1367-2630. doi:10.1088/1367-2630/16/6/065002. URL http://dx.doi.org/10.1088/1367-2630/16/6/065002.

Julien Lesgourgues, Gianpiero Mangano, Gennaro Miele, and Sergio Pastor. *Neutrino Cosmology*. Cambridge University Press, 2013.

Noam I. Libeskind, Rien van de Weygaert, Marius Cautun, Bridget Falck, Elmo Tempel, Tom Abel, Mehmet Alpaslan, Miguel A. Aragón-Calvo, Jaime E. Forero-Romero, Roberto Gonzalez, Stefan Gottlöber, Oliver Hahn, Wojciech A. Hellwing, Yehuda Hoffman, Bernard J. T. Jones, Francisco Kitaura, Alexander Knebe, Serena Manti, Mark Neyrinck, Sebastián E. Nuza, Nelson Padilla, Erwin Platen, Nesar Ramachandra, Aaron Robotham, Enn Saar, Sergei Shandarin, Matthias Steinmetz, Radu S. Stoica, Thierry Sousbie, and Gustavo Yepes. Tracing the cosmic web. *Monthly Notices of the Royal Astronomical Society*, 473(1):1195–1217, August 2017. ISSN 1365-2966. doi:10.1093/mnras/stx1976. URL http://dx.doi.org/10.1093/mnras/stx1976.

H. Lietzen, E. Tempel, L. J. Liivamägi, A. Montero-Dorta, M. Einasto, A. Streblyanska, C. Maraston, J. A. Rubiño-Martín, and E. Saar. Discovery of a massive supercluster system atz 0.47. *Astronomy amp; Astrophysics*, 588:L4, March 2016. ISSN 1432-0746. doi:10.1051/0004-6361/201628261. URL http://dx.doi.org/10.1051/0004-6361/201628261.

D. Lynden-Bell, S. M. Faber, David Burstein, Roger L. Davies, Alan Dressler, R. J. Terlevich, and Gary Wegner. Photometry and Spectroscopy of Elliptical Galaxies.

V. Galaxy Streaming toward the New Supergalactic Center. ApJ, 326:19, March 1988. doi:10.1086/166066.

J. C. Mather, E. S. Cheng, D. A. Cottingham, R. E. Eplee, Jr., D. J. Fixsen, T. Hewagama, R. B. Isaacman, K. A. Jensen, S. S. Meyer, P. D. Noerdlinger, S. M. Read, L. P. Rosen, R. A. Shafer, E. L. Wright, C. L. Bennett, N. W. Boggess, M. G. Hauser, T. Kelsall, S. H. Moseley, Jr., R. F. Silverberg, G. F. Smoot, R. Weiss, and D. T. Wilkinson. Measurement of the Cosmic Microwave Background Spectrum by the COBE FIRAS Instrument. ApJ, 420:439, January 1994. doi:10.1086/173574.

David Merritt, Alister W. Graham, Ben Moore, Jürg Diemand, and Balša Terzić. Empirical models for dark matter halos. i. nonparametric construction of density profiles and comparison with parametric models. *The Astronomical Journal*, 132 (6):2685–2700, January 2006. ISSN 1538-3881. doi:10.1086/508988. URL http://dx.doi.org/10.1086/508988.

Andrea Morandi, Ming Sun, William Forman, and Christine Jones. The galaxy cluster outskirts probed by Chandra. In *IAU General Assembly*, volume 29, page 2244785, August 2015.

Marston Morse. *The Calculus of Variations in the Large*, volume 18 of *Colloquium Publications*. American Mathematical Society, New York, 1934.

S. Nadathur, S. Hotchkiss, J. M. Diego, I. T. Iliev, S. Gottlöber, W. A. Watson, and G. Yepes. Self-similarity and universality of void density profiles in simulation and sdss data. *Monthly Notices of the Royal Astronomical Society*, 449(4):3997–4009, April 2015. ISSN 1365-2966. doi:10.1093/mnras/stv513. URL http://dx.doi.org/10.1093/mnras/stv513.

Julio F. Navarro, Carlos S. Frenk, and Simon D. M. White. A universal density profile from hierarchical clustering. *The Astrophysical Journal*, 490(2):493–508, December 1997. ISSN 1538-4357. doi:10.1086/304888. URL http://dx.doi.org/10.1086/304888.

Dmitri Novikov, Stéphane Colombi, and Olivier Doré. Skeleton as a probe of the cosmic web: the two-dimensional case. MNRAS, 366(4):1201–1216, March 2006. doi:10.1111/j.1365-2966.2005.09925.x.

Danny C. Pan, Michael S. Vogeley, Fiona Hoyle, Yun-Young Choi, and Changbom Park. Cosmic voids in sloan digital sky survey data release 7: Cosmic voids in sloan digital sky survey data release 7. *Monthly Notices of the Royal Astronomical Society*, 421(2):926–934, February 2012. ISSN 0035-8711. doi:10.1111/j.1365-2966.2011.20197.x. URL http://dx.doi.org/10.1111/j.1365-2966.2011.20197.x.

Chan-Gyung Park, Hwasu Hyun, Hyerim Noh, and Jai-chan Hwang. The cosmological principle is not in the sky. *Monthly Notices of the Royal Astronomical Society*,

469(2):1924–1931, April 2017. ISSN 1365-2966. doi:10.1093/mnras/stx988. URL http://dx.doi.org/10.1093/mnras/stx988.

Jennifer L. Pauls and Adrian L. Melott. Hierarchical pancaking: why the Zel'dovich approximation describes coherent large-scale structure in N-body simulations of gravitational clustering. MNRAS, 274(1):99–109, May 1995. doi:10.1093/mnras/274.1.99.

A. A. Penzias and R. W. Wilson. A Measurement of Excess Antenna Temperature at 4080 Mc/s. ApJ, 142:419–421, July 1965. doi:10.1086/148307.

Luis A. Pereyra, Mario A. Sgró, Manuel E. Merchán, Federico A. Stasyszyn, and Dante J. Paz. Detection and analysis of cluster-cluster filaments. MNRAS, 499(4): 4876–4886, December 2020. doi:10.1093/mnras/staa3112.

Planck Collaboration. Planck 2013 results. I. Overview of products and scientific results. A&A, 571:A1, November 2014. doi:10.1051/0004-6361/201321529.

Planck Collaboration. Planck 2018 results. VI. Cosmological parameters. A&A, 641: A6, September 2020. doi:10.1051/0004-6361/201833910.

Jesus Prada, Jaime E. Forero-Romero, Robert J. J. Grand, Rüdiger Pakmor, and Volker Springel. Dark matter halo shapes in the Auriga simulations. MNRAS, 490 (4):4877–4888, December 2019. doi:10.1093/mnras/stz2873.

H. P. Robertson. Kinematics and World-Structure. *Astrophys. J.*, 82:284–301, 1935. doi:10.1086/143681.

Nico Schuster, Nico Hamaus, Alice Pisani, Carmelita Carbone, Christina D. Kreisch, Giorgia Pollina, and Jochen Weller. The bias of cosmic voids in the presence of massive neutrinos. *Journal of Cosmology and Astroparticle Physics*, 2019(12): 055–055, December 2019. ISSN 1475-7516. doi:10.1088/1475-7516/2019/12/055. URL http://dx.doi.org/10.1088/1475-7516/2019/12/055.

N. Scoville, H. Aussel, M. Brusa, P. Capak, C. M. Carollo, M. Elvis, M. Giavalisco, L. Guzzo, G. Hasinger, C. Impey, J. P. Kneib, O. LeFevre, S. J. Lilly, B. Mobasher, A. Renzini, R. M. Rich, D. B. Sanders, E. Schinnerer, D. Schminovich, P. Shopbell, Y. Taniguchi, and N. D. Tyson. The Cosmic Evolution Survey (COSMOS): Overview. ApJS, 172(1):1–8, September 2007. doi:10.1086/516585.

Ruchika Seth and Somak Raychaudhury. Evolution of galaxies in groups in the coma supercluster. *Monthly Notices of the Royal Astronomical Society*, 497(1):466–481, June 2020. ISSN 1365-2966. doi:10.1093/mnras/staa1779. URL http://dx.doi.org/10.1093/mnras/staa1779.

Sergei F. Shandarin. Morphological statistics of the cosmic web. In Antonaldo Diaferio, editor, *IAU Colloq. 195: Outskirts of Galaxy Clusters: Intense Life in the Suburbs*, pages 45–50, July 2004. doi:10.1017/S1743921304000110.

Stephen A. Shectman, Stephen D. Landy, Augustus Oemler, Douglas L. Tucker, Huan Lin, Robert P. Kirshner, and Paul L. Schechter. The Las Campanas Redshift Survey. ApJ, 470:172, October 1996. doi:10.1086/177858.

T. Sousbie. The persistent cosmic web and its filamentary structure - I. Theory and implementation. MNRAS, 414(1):350–383, June 2011. doi:10.1111/j.1365-2966.2011.18394.x.

Volker Springel. The cosmological simulation code GADGET-2. MNRAS, 364(4): 1105–1134, December 2005. doi:10.1111/j.1365-2966.2005.09655.x.

Volker Springel, Naoki Yoshida, and Simon D. M. White. GADGET: a code for collisionless and gasdynamical cosmological simulations. New A, 6(2):79–117, April 2001. doi:10.1016/S1384-1076(01)00042-2.

Volker Springel, Carlos S. Frenk, and Simon D. M. White. The large-scale structure of the Universe. Nature, 440(7088):1137–1144, April 2006. doi:10.1038/nature04805.

R. S. Stoica, V. J. Martínez, J. Mateu, and E. Saar. Detection of cosmic filaments using the Candy model. A&A, 434(2):423–432, May 2005. doi:10.1051/0004-6361:20042409.

P. M. Sutter, Guilhem Lavaux, Benjamin D. Wandelt, and David H. Weinberg. A public void catalog from the sdss dr7 galaxy redshift surveys based on the watershed transform. *The Astrophysical Journal*, 761(1):44, November 2012. ISSN 1538-4357. doi:10.1088/0004-637x/761/1/44. URL http://dx.doi.org/10.1088/0004-637X/761/1/44.

E. Sánchez. The dark energy survey: Status and first results. *Nuclear and Particle Physics Proceedings*, 273-275:302–308, 2016. ISSN 2405-6014. doi:https://doi.org/10.1016/j.nuclphysbps.2015.09.042. URL https://www.sciencedirect.com/science/article/pii/S2405601415005313. 37th International Conference on High Energy Physics (ICHEP).

Max Tegmark, Michael R. Blanton, Michael A. Strauss, Fiona Hoyle, David Schlegel, Roman Scoccimarro, Michael S. Vogeley, David H. Weinberg, Idit Zehavi, Andreas Berlind, Tamás Budavari, Andrew Connolly, Daniel J. Eisenstein, Douglas Finkbeiner, Joshua A. Frieman, James E. Gunn, Andrew J. S. Hamilton, Lam Hui, Bhuvnesh Jain, David Johnston, Stephen Kent, Huan Lin, Reiko Nakajima, Robert C. Nichol, Jeremiah P. Ostriker, Adrian Pope, Ryan Scranton, Uroš Seljak, Ravi K. Sheth, Albert Stebbins, Alexander S. Szalay, István Szapudi, Licia Verde, Yongzhong Xu, James Annis, Neta A. Bahcall, J. Brinkmann, Scott Burles, Francisco J. Castander, Istvan Csabai, Jon Loveday, Mamoru Doi, Masataka Fukugita, J. Richard Gott, III, Greg Hennessy, David W. Hogg, Željko Ivezić, Gillian R. Knapp, Don Q. Lamb, Brian C. Lee, Robert H. Lupton, Timothy A. McKay, Peter Kunszt, Jeffrey A. Munn, Liam O'Connell, John Peoples, Jeffrey R. Pier, Michael Richmond, Constance Rockosi, Donald P. Schneider, Christopher Stoughton, Douglas L. Tucker, Daniel E. Vanden Berk, Brian Yanny, Donald G. York, and SDSS

Collaboration. The Three-Dimensional Power Spectrum of Galaxies from the Sloan Digital Sky Survey. ApJ, 606(2):702–740, May 2004. doi:10.1086/382125.

E. Tempel, R. S. Stoica, R. Kipper, and E. Saar. Bisous model-Detecting filamentary patterns in point processes. *Astronomy and Computing*, 16:17–25, July 2016. doi:10.1016/j.ascom.2016.03.004.

Giovanni Verza, Alice Pisani, Carmelita Carbone, Nico Hamaus, and Luigi Guzzo. The void size function in dynamical dark energy cosmologies. J. Cosmology Astropart. Phys., 2019(12):040, December 2019. doi:10.1088/1475-7516/2019/12/040.

Giovanni Verza, Carmelita Carbone, Alice Pisani, and Alessandro Renzi. DEMNUni: disentangling dark energy from massive neutrinos with the void size function. J. Cosmology Astropart. Phys., 2023(12):044, December 2023. doi:10.1088/1475-7516/2023/12/044.

Matteo Viel, Martin G. Haehnelt, and Volker Springel. The effect of neutrinos on the matter distribution as probed by the intergalactic medium. J. Cosmology Astropart. Phys., 2010(6):015, June 2010. doi:10.1088/1475-7516/2010/06/015.

N. Vittorio. *Cosmology*. CRC Press, Taylor & Francis Group, 2017.

Mark Vogelsberger, Federico Marinacci, Paul Torrey, and Ewald Puchwein. Cosmological simulations of galaxy formation. *Nature Reviews Physics*, 2(1):42–66, January 2020. doi:10.1038/s42254-019-0127-2.

A. G. Walker. On milne's theory of world-structure*. *Proceedings of the London Mathematical Society*, s2-42(1):90–127, 01 1937. ISSN 0024-6115. doi:10.1112/plms/s2-42.1.90. URL `https://doi.org/10.1112/plms/s2-42.1.90`.

Steven Weinberg. *Cosmology*. Oxford University Press, Oxford, UK, 2008. ISBN 978-0-19-852682-7.

R. L. Workman et al. Review of Particle Physics. *PTEP*, 2022:083C01, 2022. doi:10.1093/ptep/ptac097.

Ia. B. Zel'dovich and S. F. Shandarin. Large-scale distribution of matter in the universe. In L. V. Keldysh and V. Ia. Fainberg, editors, *Problems in Theoretical Physics and Astrophysics*, pages 338–352. 1989.

Ia. B. Zeldovich, J. Einasto, and S. F. Shandarin. Giant voids in the Universe. Nature, 300(5891):407–413, December 1982. doi:10.1038/300407a0.

Ya. B. Zel'dovich. Gravitational instability: An approximate theory for large density perturbations. A&A, 5:84–89, March 1970.

Andrea Zonca, Leo Singer, Daniel Lenz, Martin Reinecke, Cyrille Rosset, Eric Hivon, and Krzysztof Gorski. healpy: Python wrapper for HEALPix. Astrophysics Source Code Library, record ascl:2008.022, August 2020.

# Appendix

## Stacked Filaments: 2sFR Method

The stacked dark matter (DM) halo density contrast in a 200 $\times$ 40 $Mpc$ grid at the positions of the reconstructed filaments for all the redshift slices of sample B are provided in the following figures: Figure 5.1, Figure 5.2, Figure 5.3, and Figure 5.4.

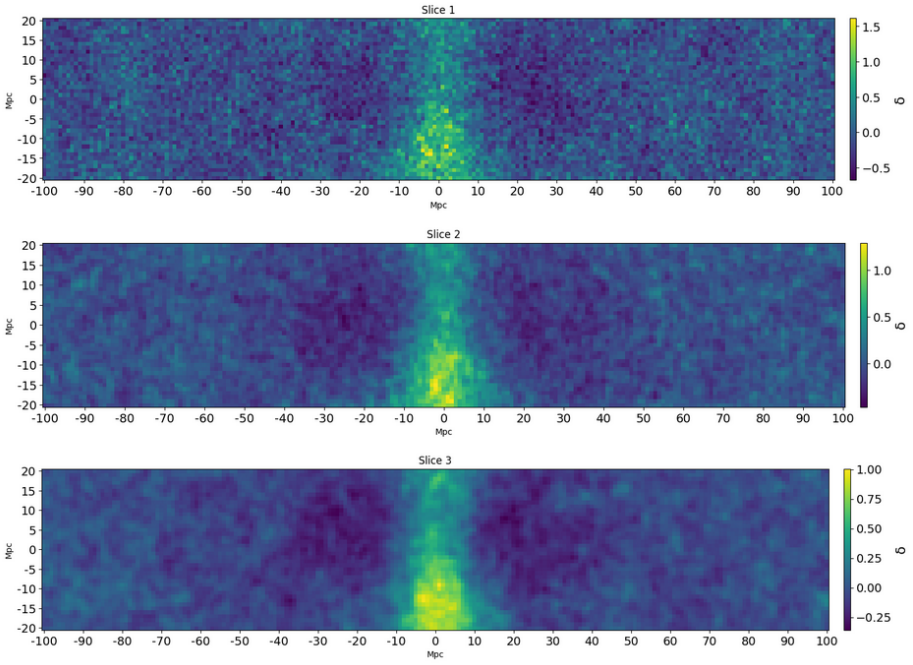



Figure 5.1: Stacked DM halo contrast at the positions of the reconstructed filaments in redshift slices: Slice 1, Slice 2 and Slice 3 of sample B.

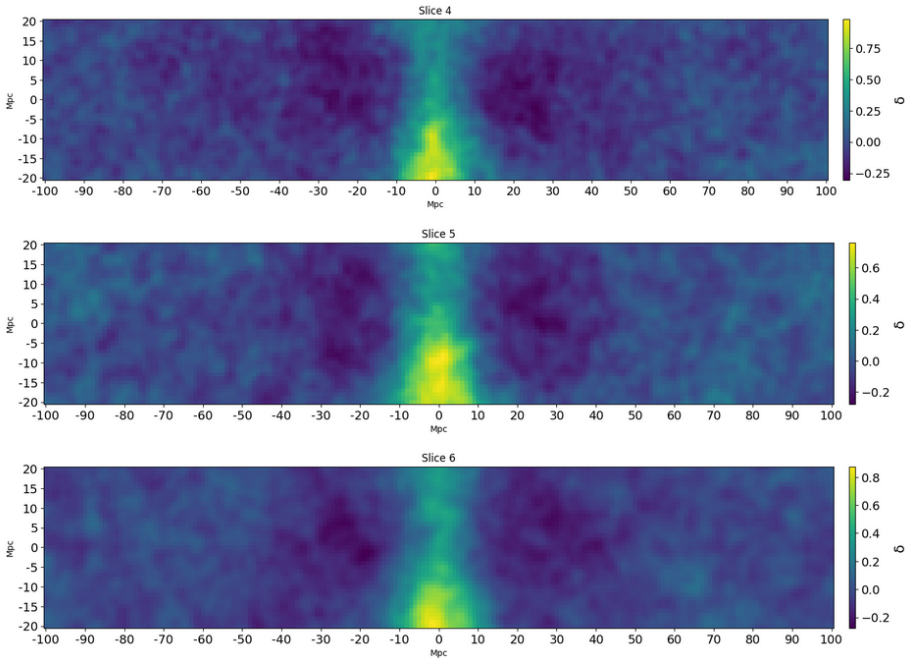



Figure 5.2: Stacked DM halo contrast at the positions of the reconstructed filaments in redshift slices: Slice 4, Slice 5 and Slice 6 of sample B.

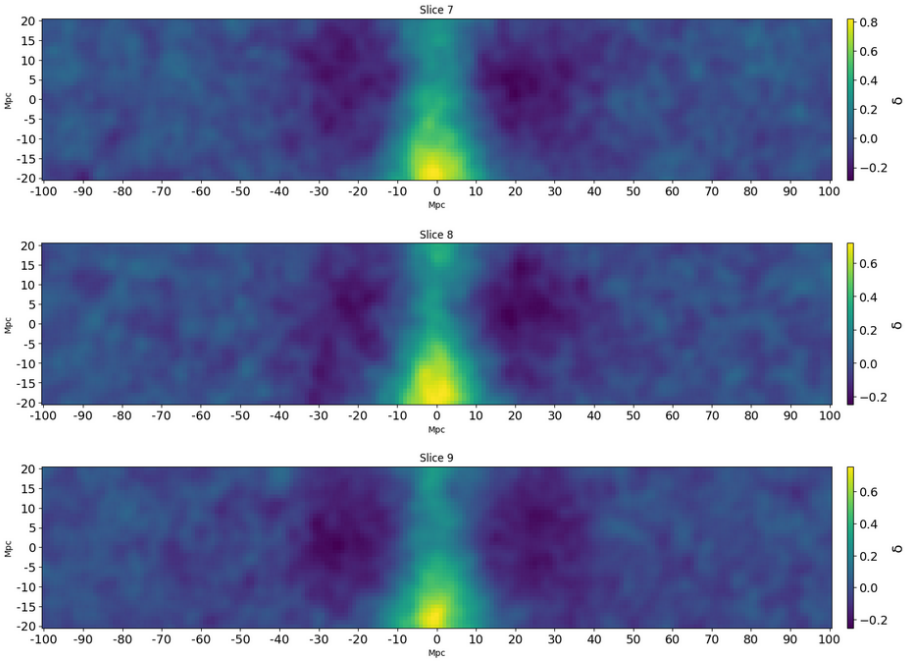



Figure 5.3: Stacked DM halo contrast at the positions of the reconstructed filaments in redshift slices: Slice 7, Slice 8 and Slice 9 of sample B.

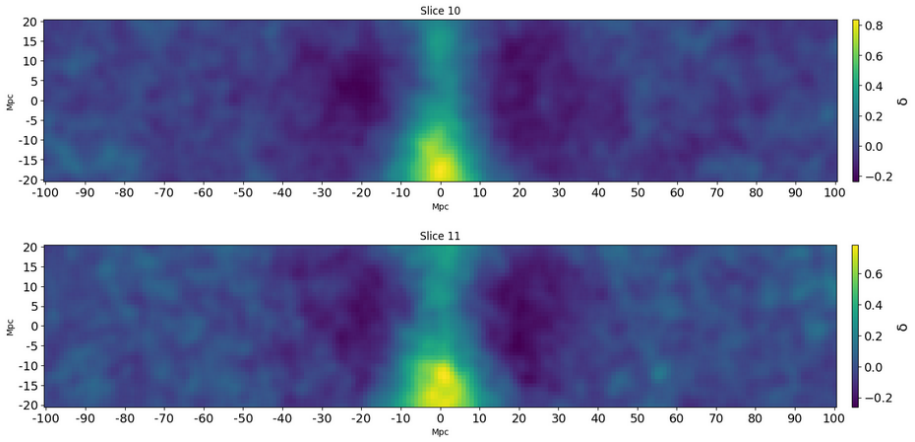



Figure 5.4: Stacked DM halo contrast at the positions of the reconstructed filaments in redshift slices: Slice 10 and Slice 11 of sample B.

# Estimation of Filament Width

The FWHM of the stacked profile of the filaments are used as a proxy to study the filament width. We estimate the FWHM of the profile by following method:

The profile of the filament is obtained by taking the average density contrast of DM halos along the direction of the axis of the stacked filaments in the $200 \times 40$ $Mpc$ grid where each pixel corresponds to 1 $Mpc$ scale. We compute half of the maximum of the profile and we find the pixel indices where the value of the profile is less than half of the maximum. Then we compute the difference between the larger pixel index and the smaller pixel index values, the difference will correspond to the FWHM in $Mpc$ scale. In Figure 5.5, we can see the way the FWHM is estimated in the stacked profile of the reconstructed filaments in the intermediate redshift slice of Sample A smoothed with a Gaussian kernel of 3 $degrees$.

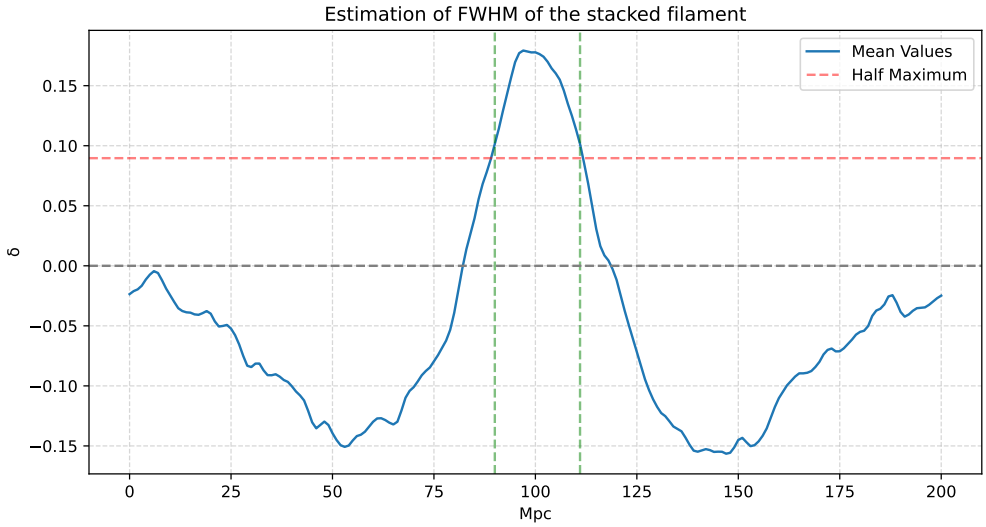



Figure 5.5: Estimation of the FWHM of a stacked profile of reconstructed filaments in the intermediate redshift slice of Sample A smoothed with a Gaussian kernel of 3 $degrees$.

# Acronyms

**ΛCDM** - Λ Cold Dark Matter

**CMB** - Cosmic Microwave Background

**CNN** - Convolutional Neural Network

**DE** - Dark Energy

**DEMNUni** - Dark Energy and Massive Neutrino Universe

**FLRW** - Friedmann-Lemaître-Robertson-Walker metric

**FWHM** - Full Width Half Maximum

**GR** - General Relativity

**LSS** - Large Scale Structure

**LSST** - Large Synoptic Survey Telescope

**ML** - Machine Learning

**MMF** - Multiscale Morphology Filter

**MST** - Minimum Spanning Tree

**SCMS** - Subspace Mean Shift Algorithm

**SKA** - Square Kilometer Array

**2sFR** - Two-step Filament Reconstruction

**WMAP** - Wilkinson Microwave Anisotropy Probe

# List of Figures

# List of Tables