Probing Dark Matter from the Galaxy to the Cosmic Web

by

Tianyi Yang

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Science
in
Physics

Waterloo, Ontario, Canada, 2020

Author's declaration

This thesis consists of material all of which I authored or co-authored: see Statement of Contributions included in the thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Statement of contributions

Tianyi Yang was the sole author for Chapter 1 which were written under the supervision of Dr. Michael Hudson and Dr. Niayesh Afshordi.

This thesis consists in part of two manuscripts written for publication. Exceptions to sole authorship of material are as follows:

Research presented in Chapters 2 and Appendix A to D

This research was conducted at the University of Waterloo by Tianyi Yang under the supervision of Dr. Niayesh Afshordi. Dr. Niayesh Afshordi, Tianyi Yang and Supranta Sarma Boruah contributed to study design. Tianyi Yang was the primary coders, contributing to coding and data analysis. Supranta Sarma Boruah participated in coding and analysis. Tianyi Yang drafted the manuscript, which all co-authors contributed intellectual input on. This work has been released on arXiv and has been submitted to Monthly Notices of the Royal Astronomical Society (MNRAS), and is ready to be published.

Citations:

Tianyi Y, Boruah S S, Afshordi N, "Gravitational Potential from small-scale clustering in action space: Application to Gaia DR2", arXiv:1908.02336 [astro-ph.GA].

Research presented in Chapters 3

This research was conducted at the University of Waterloo by Tianyi Yang under the supervision of Dr. Michael Hudson and Dr. Niayesh Afshordi. Dr. Michael Hudson and Tianyi Yang designed the study with assistance from Dr. Niayesh Afshordi. Tianyi Yang contributed to coding and completed the data analysis. Tianyi Yang drafted the manuscript and each author provided intellectual input on manuscript drafts.

Tianyi Y, Hudson J M, Afshordi N, "How dark are filaments in the cosmic web?", paper in preparation.

As lead author of these chapters and appendices, I was responsible for contributing to conceptualizing study design, carrying out data collection and analysis, and drafting and submitting manuscripts. My coauthors provided guidance during each step of the research and provided feedback on draft manuscripts.

Abstract

Dark matter, although invisible, accounts for the majority of matter of the universe. How this invisible component affects cosmic structure formation is one of the primary lines of inquiry in physical cosmology, and the key to understanding its nature. In this thesis, we present two methods to probe dark matter from the Milky Way to the filaments of the cosmic web.

In the first part of this thesis, we demonstrate how to use stellar clustering in action space to probe the underlying gravitational potential of the Milky Way's dark matter halo. Provided that the correct potential is used for the system, integrals of motion such as action variables of small structures (for instance the tidal streams surrounding the galaxy) are conserved during galaxy formation and evolution. If the incorrect potential is applied, action variables will not be conserved, weakening the small-scale clustering in the action space. Conversely, the correct potential is expected to maximize small-scale clustering in action space. After justifying the viability of this idea using simulations, we apply this method to the 2nd data release from Gaia mission, and use it to measure the fraction that the halo contributes to the total centrifugal force at solar position, f_h, and logarithmic slope, α, of a power-law dark matter halo profile. We use stars within 9-11 kpc and 11.5-15 kpc from Galactic centre, and find the power-law potential, which is parametrized by f_h and α, is $(f_h, \alpha) = (0.391 \pm 0.009, 1.835 \pm 0.092)$ and $(0.351 \pm 0.012, 1.687 \pm 0.079)$ respectively. We then use the best-fit potential to compute the total circular velocity of the Milky Way within R = 9-15 kpc. The resulting circular velocity curve is consistent with past measurements (although it is $\sim$ 5-10% lower than previous methods based on masers or globular clusters). To our knowledge, this is the first study that applies this methodology to real data. Furthermore, by constraining the Milky Way potential, our result indirectly shows the existence of dark matter halo around Milky Way.

On cosmological scales, massive dark matter halos are expected to be connected by bridges, known as filaments. Like other large scale structures in our universe, filaments are expected to be dominated by dark matter, making them hard to detect. But are these filaments any darker than other structures, such as voids, clusters and galaxies? In the second part of this thesis, we investigate how "bright" these dark filaments are, which can be characterized by their mass-to-light, or M/L ratios. We first estimate the mass of these dark filaments via weak gravitational lensing: stacking and analyzing the weak lensing signals between Luminous Red Galaxy (LRG) pairs selected from SDSS III BOSS survey. Using the CFHTLens shape measurements, we measure the mass of filaments at a significance level of 4.5 σ. To isolate the filament signal, a catalogue of non-physical projected pairs is constructed. Then, we investigate the average luminosity level of filaments by sub-

tracting the stacked non-physical pairs from the stacked LRG pairs selected from the BOSS survey. We fit a Schechter function over the observed excess galaxy number in filaments, and so, compute the total luminosity in SDSS ${}^{0.1}r$ and ${}^{0.1}g$ band. Then, we calculate the mass-to-light ratio, M/L, and the colour ${}^{0.1}(g-r)$. To investigate the redshift dependence of these parameters, the above analyses are conducted in two independent redshift samples (LOWZ and CMASS as divided by BOSS survey). We find $M/L = 309\pm94$, ${}^{0.1}(g-r) = 0.59\pm0.24$ for LOWZ sample, and $M/L = 435\pm189$, ${}^{0.1}(g-r) = 0.38\pm0.45$ for CMASS sample. If we combine both samples, we find $M/L = 351\pm87$, ${}^{0.1}(g-r) = 0.51\pm0.22$. Due to the uncertainties, we find no significant redshift dependence of these parameters. Our study provides the first measurement of the mass-to-light ratios of filaments of the cosmic web, showing that they are comparable to the cosmic mean value.

Acknowledgements

I would like to thank my adviser Mike Hudson and Niayesh Afshordi for their guidance, patience, and contributions.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

In our universe, matter accounts for approximately 30% of the mass-energy budget of the whole universe, where dark matter makes up more than 80% of the total matter content. This component of the universe, although invisible, plays an important role in cosmic structure formation. Understanding the nature of it has become one of the most significant challenges in modern cosmology in the recent decades. Observational evidence support that our universe can be described by a ΛCDM cosmological model (Planck Collaboration et al., 2016).

Although the cold dark matter (CDM) paradigm has been the most popular model of dark matter, it is difficult to probe the nature of dark matter because it only interacts with ordinary matter via gravity, and furthermore, on small scales ($\lesssim$ few kpc) its effects are difficult to disentangle from the effects of the baryonic component. On these scales, extracting the full 6-dimensional phase space information (that is necessary to infer dark matter mass unambiguously), has been a difficult task. The Milky Way, for example, is believed to be surrounded by a dark matter halo, but measuring the structure of this component has been proven to be challenging. Fortunately, the proliferation of new data on the kinematics of the Milky Way's stars has opened a new avenue to probe the structure of the Milky Way's potential and of the nature of dark matter. Using these data, we are therefore able to constrain the potential of the Milky Way by studying the stellar distribution function in phase space, which indirectly shows the presence of dark matter halo around Milky Way.

On cosmological scales, one of the predictions of the Λ cold dark matter model is that the distribution of dark matter halos, where galaxies, groups and clusters are formed follows a web-like pattern on large scales, which is the so-called "Cosmic Web". The

Cosmic Web is composed of halos, filaments, sheets and also voids, which are underdense regions surrounded by filaments and sheets. In particular, dark matter halos are regarded as nodes of the Cosmic web, and these nodes are expected to be connected by dark-matter dominated filaments. Although the presence of filaments has been traced by the galaxy redshift distribution (see Section 1.3.2.1 for a brief review), their dark matter properties remain poorly understood observationally. Quantities such as the mass-to-light ratio, M/L, which may be considered as the connection between the light and the underlying dark matter distribution, can be evaluated in order to study the properties of filaments. By measuring the M/L of the filament, we can investigate how "bright" the dark-matter dominated filaments are when compared with other cosmic structures.

In this thesis, we present two methods which aim to probe the properties of dark matter from the Galaxy (the Milky Way) to the Cosmic Web (the filaments). In this chapter, we first derive the Collisionless Boltzmann Equation (CBE), then we demonstrate its application in the derivation of linear perturbation theory and extend this into the formation of non-linear structures. Weak gravitational lensing and the three-point correlation function are presented as tools for filament studies. We further show the application of the CBE in describing the galactic dynamics and kinematics. We illustrate the convenience of using action variables to describe a system and its significance in constraining the potential of the Milky Way. At the end of this chapter we summarize recent studies related to constraining the Milky Way potential using action variables, and dark-matter filament detection and measurements of the mass-to-light ratios in different cosmic environments. We finish up this chapter by presenting an outline of this thesis.

1.1 Probing Cosmological Perturbations

To model the kinematics and dynamics of particles, one method is to apply Newton's law of gravity to individual particles of the system. However, this may be computationally expensive considering there are hundreds of millions of particles. Therefore, it is preferable to compute the average motion of particles. This is the motivation for deriving the Collisionless Boltzmann Equation: instead of focusing on individual particles with some positions and velocities, it describes how the probability of a particle occupied in a small volume of d^3xd^3v evolves. The Collisionless Boltzmann Equation describes the dynamics of collisionless systems and it is not only applicable in studying the kinematics and dynamics of stars on galactic scale (which will be discussed in Section 1.2.1), but also useful in describing the evolution of perturbations on cosmological scales. In this section, we first derive the Collisionless Boltzmann Equation and demonstrate its application in the

derivation of linear perturbation theory and extend this to the formation of non-linear structures.

1.1.1 Collisionless Boltzmann Equation

For a galaxy system that is composed of stars and dark matter particles, one way to characterize the equations of motion for this system is to trace the orbits of individual particles. Considering a system with N point masses, the equations of motion for this system can be written as the following equations:

$$\begin{aligned} \frac{d\boldsymbol{r_i}}{dt} &= \boldsymbol{v_i}, \\ \frac{d\boldsymbol{v_i}}{dt} &= \sum_{j\neq i}^{N} Gm_j \frac{\boldsymbol{r_j} - \boldsymbol{r_i}}{|\boldsymbol{r_j} - \boldsymbol{r_i}|^3}, \end{aligned} \tag{1.1}$$

where the summation is over all N point masses, $\boldsymbol{r_i}$ and $\boldsymbol{v_i}$ is the position and velocity of ith point mass, and G is the gravitational constant. However, for a system with a large population of particles, it is more convenient to describe the system statistically: instead of following the motion of each particle, in general, a distribution function, f, can be introduced to represent the density distribution of particles in a system. As we shall see later, with the distribution function, one can determine the density and motion of particles in a particular region at any chosen moment.

To derive the dynamical equations for the distribution function, we follow the procedures outlined in Sparke & Gallagher (2007). Here, we are working in phase space, which is a six-dimensional space and the coordinate of each point is denoted as ($\boldsymbol{x}$, $\boldsymbol{v}$). At time t, for a given probability function $f(\boldsymbol{x}, \boldsymbol{v}, t)$ in phase space, the average number density of particles in a cube centered at position $\boldsymbol{x}$, after marginalizing all possible v_is is thus given by:

$$n(\boldsymbol{x}, t) = \int_{-\infty}^{-\infty}\int_{-\infty}^{-\infty}\int_{-\infty}^{-\infty} f(\boldsymbol{x}, \boldsymbol{v}, t) dv_x dv_y dv_z. \tag{1.2}$$

The mean velocity, $\langle \boldsymbol{v}(\boldsymbol{x}, t)\rangle$ is defined as:

$$\langle \boldsymbol{v}(\boldsymbol{x}, t)\rangle = \frac{\int_{-\infty}^{-\infty}\int_{-\infty}^{-\infty}\int_{-\infty}^{-\infty} \boldsymbol{v} f(\boldsymbol{x}, \boldsymbol{v}, t) dv_x dv_y dv_z}{\int_{-\infty}^{-\infty}\int_{-\infty}^{-\infty}\int_{-\infty}^{-\infty} f(\boldsymbol{x}, \boldsymbol{v}, t) dv_x dv_y dv_z}. \tag{1.3}$$

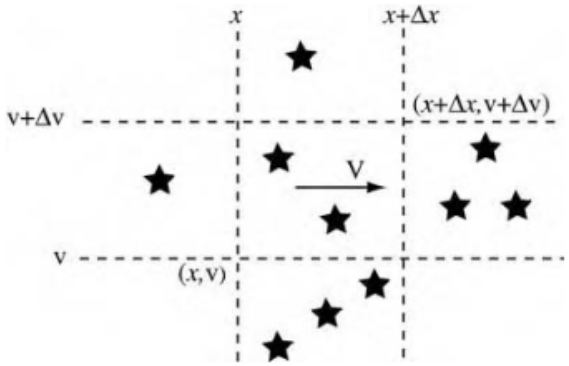

Figure 1.1: Stellar motion in phase space. Here, we focus on the number changes of stars in the central box to derive the CBE. Image is taken from Sparke & Gallagher (2007)

Now we would like to derive how the change in the distribution function relates to the change in the velocity and density. Consider a system where the two-body encounter and relaxations are negligible, with particles are moving under a smooth potential, $\Phi(\boldsymbol{x}, t)$. For simplicity, let us assume that stars move along one direction, x, with velocity, v. We would like to see how the number of stars changes in a particular region in phase space after a particular time Δt.

Figure 1.1 shows how the stars flow in and out in the central box in a two-dimensional phase space. In the central box, stars that lie within position range $[x,\ x+\Delta x]$ are moving at speeds $[v,\ v+\Delta v]$. Along the x direction, after a time interval Δt, stars that were originally located between $[x - v\Delta t,\ x]$ will enter the central box, while stars at the other end within distance $v\Delta t$ will leave the box. Therefore, the number difference within Δt in the box is $\Delta v \Delta t[vf(x, v, t) - vf(x + \Delta x, v, t)]$. However, as stars are affected by a smooth potential, the velocity of stars can change with time. Assume stars are accelerating along the x direction so that $dv/dt > 0$. Due to the change in velocity, after a time interval Δt, stars that were originally moving at speeds between $[v - \Delta t(dv/dt),\ v]$ will enter the box, while stars that were moving with speed $v + \Delta v$ will leave the box due to the acceleration. Therefore, in total, the number of stars in the center box have changed by:

$$\Delta v \Delta x[f(x, v, t + \Delta t) - f(x, v, t)] =$$

$$\Delta v \Delta t\Big[vf(x, v, t) - vf(x + \Delta x, v, t)\Big] + \Delta x \Delta t\Big[\frac{dv}{dt} f(x, v, t) - \frac{dv}{dt} f(x, v + \Delta v, t)\Big]$$

$$\approx -v\Delta x \Delta v \Delta t \frac{\partial f}{\partial x} + \Delta x \Delta t\Big[\frac{dv}{dt} f(x, v, t) - \frac{dv}{dt} f(x, v + \Delta v, t)\Big] \tag{1.4}$$

While in the limit of infinitesimal change of position, velocity and time, equation 1.4 can be rewritten as :

$$\frac{\partial f}{\partial t} + v\frac{\partial f}{\partial x} + \frac{dv}{dt}\frac{\partial f}{\partial v} = 0. \tag{1.5}$$

As the velocity change of particle is due to the presence of smooth potential, where the potential is a function that only depends on the position and time. Therefore,

$$\frac{\partial f}{\partial t} + v\frac{\partial f}{\partial x} - \frac{\partial \Phi(x,t)}{\partial x}\frac{\partial f}{\partial v} = 0, \tag{1.6}$$

where we have replaced dv/dt with $-\partial\Phi(x,t)/\partial t$.

Equation 1.6 is the one-dimensional Boltzmann equation. In general, for a system with full six-dimensional information, the general form of the CBE can be taken as

$$\frac{\partial f}{\partial t} + \boldsymbol{v}\cdot\nabla f - \nabla\Phi\cdot\frac{\partial f}{\partial \boldsymbol{v}} = 0. \tag{1.7}$$

However, solving the full CBE is generally hard as the distribution function f is a function of seven variables ($\boldsymbol{x}, \boldsymbol{v}$ and t). It can be solved in some specific cases (e.g. Gressman & Strain, 2010), but those analytical solutions might not be applicable to real problems. Besides, it is usually not the distribution function f being observed. In real observations, what can be determined are other distributions such as the number density profile or velocity dispersion profile. Therefore, we would like to derive the first moment (velocity-moment) of equation 1.7. Integrating Equation 1.6 over velocity gives:

$$\frac{\partial n(x,t)}{\partial t} + \frac{\partial[n(x,t)\langle v(x,t)\rangle]}{\partial x} - [f]_{-\infty}^{+\infty}\frac{\partial\Phi}{\partial x} = 0, \tag{1.8}$$

where we use the definitions Equation 1.2 and 1.3. Also, we assume that when velocity goes to infinity, the probability distribution goes to zero. This is the equation of continuity that describes mass conservation in phase space.

Multiplying Equation 1.6 by v and integrating over velocity gives:

$$\frac{\partial[n(x,t)\langle v(x,t)\rangle]}{\partial t} + \frac{\partial[n(x,t)\langle v^2(x,t)\rangle]}{\partial x} = -n(x,t)\frac{\partial\Phi}{\partial x}, \tag{1.9}$$

where we use the definition Equation 1.3.

Now define the velocity dispersion as $\langle v^2(x,t)\rangle = \langle v(x,t)\rangle^2 + \sigma^2$. Using the basic rule of integration by parts and the equation of continuity, Equation 1.9 can be reduced to:

$$\frac{\partial\langle v(x,t)\rangle}{\partial t} + \langle v(x,t)\rangle\frac{\partial\langle v(x,t)\rangle}{\partial x} = -\frac{\partial\Phi}{\partial x} - \frac{1}{n(x,t)}\frac{\partial}{\partial x}[n(x,t)\sigma^2(x,t)]. \tag{1.10}$$

Generally, in six-dimensional phase space, the above equation can be written as:

$$\frac{\partial\langle v_j\rangle}{\partial t} + \sum_i \langle v_i\rangle \frac{\partial\langle v_j\rangle}{\partial x_i} = -\frac{\partial\Phi}{\partial x_j} - \frac{1}{n}\sum_i \frac{\partial}{\partial x_i}[n\sigma_{ij}^2] \ (j = 1, 2, 3), \tag{1.11}$$

where the stress tensor σ_{ij}^2 is defined as $\langle v_i v_j\rangle - \langle v_i\rangle\langle v_j\rangle$ and the summation is over all velocity components. This is the Jeans equation and it describes the stellar motion in phase space under a gravitational field. This equation tells us that, for a given population of tracers (for instance, stars in the galaxies), we can determine any of these three quantities: the number density (n), velocity dispersion (σ) and the potential (Φ) as long as the other two are known.

1.1.2 Linear Perturbation Theory

Having derived the CBE and Jeans Equation in previous section, in this section, we first see how these two equations, together with other basic equations, are applied to cosmology and are used to derive the linear perturbation theory.

In an expanding universe, the position of particles are described in the comoving coordinates as these quantities of an object are remained as constants during the expansion of the universe. Here, we define the comoving distance as $\boldsymbol{r} = a(t)\boldsymbol{x}$, where $a(t)$ is the cosmic scale factor evaluated at an arbitrary time t. The scale factor characterizes the expansion of the universe and by our definition, $a(t_0) = 1$.

Then, to characterize the number density and velocity in the cosmological field under a smooth potential, we define the density and velocity as $\boldsymbol{v} \equiv a\dot{\boldsymbol{x}} = \boldsymbol{u} - (\dot{a}/a)\boldsymbol{r}$, and $\rho(\boldsymbol{x}, t) = \overline{\rho_m}(1 + \delta(\boldsymbol{x}, t))$, where $\delta(\boldsymbol{x}, t)$ is the density contrast, $\dot{a}/a \equiv H$ is the Hubble parameter and v is the peculiar velocity which originates from the density fluctuations. The peculiar velocity represents the motion of an object that deviates from the expansion of the universe.

Now we would like to describe the equations of motion in terms of comoving distance. Recall the equation of continuity and Jeans Equation derived in previous section are both evaluated in the proper frame, therefore, we need to transform the coordinate into comoving frame. For an arbitrary function $f(\boldsymbol{x}, t)$, based on the chain rule, the time derivative of that function f at position $\boldsymbol{x}$ is:

$$(\frac{\partial f}{\partial t})_{\boldsymbol{x}} = (\frac{\partial f}{\partial t})_{\boldsymbol{r}} + (\frac{\partial r}{\partial t})_{\boldsymbol{x}} \cdot \nabla_{\boldsymbol{r}} f = (\frac{\partial f}{\partial t})_{\boldsymbol{r}} + \frac{\dot{a}r}{a^2} \cdot (\nabla f). \tag{1.12}$$

Based on this, we can transform the derivative with respect to proper coordinates to the derivative in the comoving frame:

$$\nabla = a\nabla_{\boldsymbol{r}},\ (\frac{\partial}{\partial t})_{\boldsymbol{r}} = (\frac{\partial}{\partial t})_{\boldsymbol{x}} - \frac{\dot{a}}{a}\boldsymbol{x}\cdot\nabla. \tag{1.13}$$

Using the definition of density and Equation 1.13, we can rewrite the equation of Continuity (Equation 1.8), Jeans Equation (Equation 1.11) and Poisson Equation as follows:

$$\frac{\partial\delta}{\partial t} + \frac{1}{a}\nabla\cdot[(1+\delta)\boldsymbol{v}] = 0, \tag{1.14}$$

$$\frac{\partial\boldsymbol{v}}{\partial t} + \frac{\dot{a}}{a}\boldsymbol{v} + \frac{1}{a}(\boldsymbol{v}\cdot\nabla)\boldsymbol{v} = -\frac{1}{a}\nabla\Phi, \tag{1.15}$$

$$\nabla^2\Phi = 4\pi G\overline{\rho}_m a^2\delta \tag{1.16}$$

where Φ is the cosmological gravitational potential.

Under the linear regime, the density and velocity perturbations are small. Therefore, we ignore all high order terms and only leave the first-order linear terms in Equation 1.14 and 1.15:

$$\frac{\partial\delta}{\partial t} + \frac{1}{a}\nabla\cdot\boldsymbol{v} = 0, \tag{1.17}$$

$$\frac{\partial\boldsymbol{v}}{\partial t} + \frac{\dot{a}}{a}\boldsymbol{v} = -\frac{1}{a}\nabla\Phi. \tag{1.18}$$

By plugging Equation 1.18 into the time derivative of Equation 1.17, a second-order differential equation for $\delta(\boldsymbol{x},t)$ can be obtained, which looks like:

$$\frac{\partial^2\delta}{\partial t^2} + 2\frac{\dot{a}}{a}\frac{\partial\delta}{\partial t} = 4\pi G\overline{\rho}_m\delta. \tag{1.19}$$

As this is a second-order differential equation, it has two solutions, $\delta_1(t)$ and $\delta_2(t)$, which satisfy with $\delta_2\dot{\delta}_1 - \delta_1\dot{\delta}_2 \propto a^{-2}$. Based on the relationship between $\delta_1(t)$ and $\delta_2(t)$, as long as one solution is known, the other can be found by solving Equation 1.19. To solve this equation, consider a universe that are composed by matter and dark energy. For a perfect and pressureless fluid, we have the form of Friedmann Equation that can be written as:

$$\frac{\ddot{a}}{a} = \dot{H} + \frac{\dot{a}^2}{a^2} = -\frac{4\pi G}{3}\overline{\rho}_m + \frac{\Lambda c^2}{3}, \tag{1.20}$$

where $\overline{\rho}_m$ is mean matter density of the fluid which goes as $\overline{\rho}_m \propto a^{-3}$, and Λ is the cosmological constant.

Taking time derivative to the above Friedmann equation gives us the following equation:

$$\frac{d^2H}{dt^2} + 2\frac{\dot{a}}{a}\frac{dH}{dt} = 4\pi G\overline{\rho}_m H, \tag{1.21}$$

which is the same as Equation 1.19. Therefore, we can conclude one solution to that second-order differential equation is:

$$\delta_- \propto H(t) \propto t^{-1} \tag{1.22}$$

which is the decaying mode of $\delta(t)$ as this mode decreases with time. With one solution of $\delta(t)$, we can find the other solution by substituting δ_- into the $\delta_1\delta_2$ mentioned previously and write the growing mode of $\delta(t)$ as:

$$\delta_+ \propto H(t)\int_0^t \frac{dt'}{a^2(t')H^2(t')}. \tag{1.23}$$

1.1.3 Higher-Order Perturbation Theory and the Zel'dovich Approximation

The previous section only focus on the linearized version of equations of motion. However, in order to have a more complete picture on the formation and evolution of the universe, we have to consider the structure formation in quasi-linear (when $\delta \lesssim 1$) and non-linear regime. To account for those, it is necessary to include the higher order calculations in perturbation theory to estimate the nonlinear effect, which means that instead of using Equation 1.18, one needs to start with Equation 1.14 and Equation 1.15 to derive how the density and velocity field evolve. Solving these equations can be difficult and tedious, but they can be analytically worked out in Fourier space using perturbation theory. For readers who are interested in the whole derivation, the detailed procedures are discussed in e.g. Peebles (1980), Jain & Bertschinger (1994), Bernardeau (1994) and Bernardeau et al. (2002). Here, we would like to focus on one specific analytic treatment of the formation of non-linear structures, which is the Zel'dovich Approximation (Zel'dovich, 1970).

So far we have been using the density and velocity field to represent the equations of motion of a system. However, as mentioned in the beginning of this chapter, another alternative way to solve the system is tracing the trajectory of single particles. In this framework, the key is to relate the position, $\boldsymbol{x}$, of a mass element to its initial position $\boldsymbol{x}_{\rm ini}$, and this mapping between $\boldsymbol{x}_{\rm ini}$ and $\boldsymbol{x}$ is linked by a displacement field, $\boldsymbol{\Psi}(\boldsymbol{x}, t)$. This displacement field can be further used for defining the local density field by:

$$1 + \delta = \frac{1}{(1-\lambda_1 D)(1-\lambda_2 D)(1-\lambda_3 D)}, \tag{1.24}$$

where D is the growth factor that can be related to the density perturbation, $\delta(\boldsymbol{x}, a) = D(a)\delta_{\rm ini}(\boldsymbol{x})$, and $D(a_{\rm ini}) = 1$. $\lambda_1, \lambda_2, \lambda_3$ are the three eigenvalues of the deformation tensor defined as $\partial\boldsymbol{\Psi}_i/\partial x_j \partial x_k$. Note that the Zel'dovich Approximation predicts that the first non-linear structure to form is the two-dimensional-sheet. The formation of a web-like structure on large scales is actually predicted by Bond et al. (1996). However, in analogy with Zel'dovich theory, from Equation 1.24, we can see that the structure formation depends on the relative amplitude of three different eigenvalues (Hahn et al., 2007): a sheet-like structure (*wall*) can be formed when one of the eigenvalues is positive (only contracting along one direction). Then, prolate structures (*filaments*) can be formed when two eigenvalues are positive. That is when the overdensity matter field can be contracted along two directions. When the eigenvalues are all positive, a collapse can happen in all three directions, which would cause the formation of spherical haloes. Haloes are also the extreme points at two ends of the filaments and they are the places where galaxy clusters are formed at later time. Note that the formation of an underdense region (*void*) is also possible when three eigenvalues are all negative. However, one of the shortcoming of this approximation is that it breaks down after the formation of sheets: when, for instance, two mass elements end up at two different final positions (*shell crossing*).

To study the non-linear gravitational growth of the universe, the most powerful tools are N-body simulations. With direct numerical simulations, we can trace how the build-ups of different structures evolve with time (redshift). Dark matter particles, which are the main mass component in the numerical simulations, are assumed to be weakly interacting particle and interact with each other only through the effect of gravity. The dynamics of these particles will be followed by solving coupled partial differential equations that describe gravity and other astrophysics processes (Springel et al., 2005; Vogelsberger et al., 2014). Bond et al. (1996) predicted that the formation of a web-like pattern on large scales, dubbed the Cosmic Web. They showed that two high density nodes are expected to be connected with filaments. To clearly demonstrate this, Figure 1.2 demonstrates the dark matter distribution of a small part of the universe at z=0 taken from the *Millennium Simulation* (Springel et al., 2005). It can be seen that the web-like pattern is clearly visible: massive halos, where galaxy groups and clusters reside in, are connected by bridge-like filamentary structures, which are presumably dominated by dark matter, and they are surrounded by under-dense void regions. Filaments are one of the most prominent and intriguing structures in the comic web. We can detect them and investigate their properties by utilizing a technique called weak gravitational lensing, which will be introduced in next section.

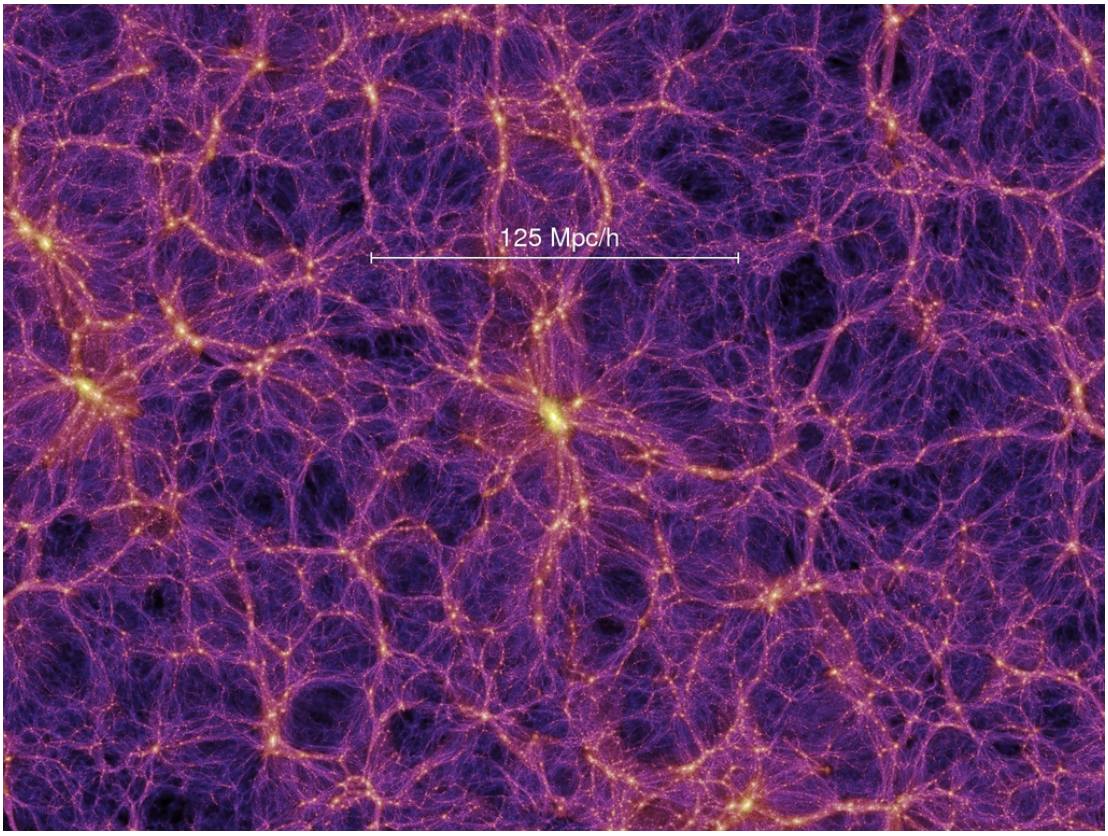

Figure 1.2: Dark matter distribution at $z = 0$ from the Millennium Simulation (Springel et al., 2005). The formation of "Cosmic Web" on large scales is clearly visible.

1.1.4 Weak Gravitational Lensing

As discussed in previous section, the presence of filaments on large scales has been successfully predicted by numerical simulations. Observationally, the presence of filaments on large scales have been revealed by tracing the galaxy distribution using various observational data (see Section 1.3.2.1 for a brief review). However, observations of their dark-matter properties are limited. One way to detect the matter distribution of filaments is through weak gravitational lensing. Gravitational lensing is useful as it is insensitive to the mass type of the object. Based on the prediction from General Relativity, light travels along geodesics in space-time. Massive objects bend the space-time and therefore deflect the path of photons. Gravitational lensing can be divided into two regimes: strong and weak gravitational lensing. Strong lensing occurs when the foreground projected lens mass is so dense that the light emitted from the background sources can be split into multiple light rays, generating multiple images of the background objects. The other case, which is more frequent, is the weak lensing.

To study lensing, we need to first define a system which is composed by an observer, a lensing mass located at redshift z_l (with angular diameter distance to the observer D_l) and a background source galaxy located at redshift z_s (with angular diameter distance D_s). The distance between the lens and the source is denoted as D_{ls}. The geometry of a lensing system is shown in Figure 1.3, where we can see the source galaxy, originally located at β, is deflected by an angle $\hat{\alpha}$ due to the presence of the foreground lens. The source galaxy is observed at position θ. From the diagram, we can obtain the following relationship using the angles and angular diameters defined above:

$$\beta D_s = \theta D_s - \hat{\alpha} D_{ls}. \tag{1.25}$$

Define the reduced deflection angle $\vec{\alpha} \equiv \frac{D_{ls}}{D_s}\hat{\alpha}$, the above equation can be reduced to:

$$\beta = \theta - \vec{\alpha}, \tag{1.26}$$

which is the lens equation. To be specific, the scaled deflection angle is related to the surface mass density via the following equation:

$$\vec{\alpha} = \frac{1}{\pi}\int d^2\theta' \kappa(\vec{\theta'}) \frac{\theta - \theta'}{|\vec{\theta} - \vec{\theta'}|^2}, \tag{1.27}$$

where κ a dimensionless parameter called the convergence. It can be related to the surface mass density through:

$$\kappa = \frac{\Sigma(D_l\vec{\theta})}{\Sigma_{\rm crit}}, \tag{1.28}$$

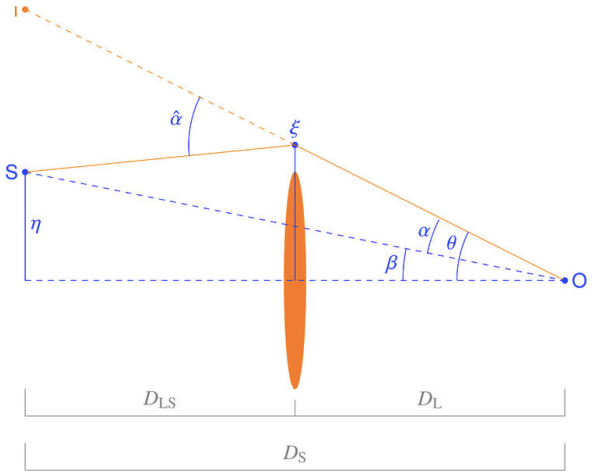

Figure 1.3: A diagram shows the geometry of a lensing system. Image is taken from Narayan & Bartelmann (1999)

where the critical density, $\Sigma_{\rm crit}$ is defined as:

$$\Sigma_{\rm crit} = \frac{c^2 D_s}{4\pi G D_{ls} D_l}. \tag{1.29}$$

As the projected surface density near the lensing matter is related to the convergence, Equation 1.28 indicates that lensing can be divided into strong lensing effect when $\Sigma(D_l\vec{\theta})$ is much larger than $\Sigma_{\rm crit}$ (κ is greater than unity), or weak lensing effect when κ is less than unity.

As mentioned above, in the weak lensing regime, background galaxies are distorted by foreground massive objects. In the study of lensing, this distortion is characterized by shear (γ) . For a circularly symmetric source, this distortion can be described as:

$$g(\vec{\theta}) = \frac{\gamma(\vec{\theta})}{1 - \kappa(\vec{\theta})}, \tag{1.30}$$

where $g(\vec{\theta})$ is the reduced shear, and it can be expressed as a complex quantity including two components, for example $g = g_1 + ig_2$. The effect of shear is manifested by the stretching of the background sources, and the effect of an nonzero convergence on the sources is magnification (or de-magnification if $\kappa < 0$). In the regime of weak lensing, as $\kappa(\vec{\theta})$ is smaller than unity, so $g(\vec{\theta}) \approx \gamma(\vec{\theta})$.

The shear is inferred by measuring the shapes of the background galaxies, which are quantified by the ellipticities. To the lowest-order approximation, the ellipticity of the source galaxy is affected by the shear and can be described as:

$$\epsilon^{\rm obs} \approx \epsilon^{\rm int} + g, \tag{1.31}$$

where $\epsilon^{\rm obs}$ is the actual observed ellipticity, g is the distortion caused by lensing. $\epsilon^{\rm int}$ is the dominated term of noise to the measured lensing signal called the intrinsic ellipticity. Intrinsic ellipticities is caused by the fact that the unlensed source galaxies do not have perfectly circular symmetry, which means that the measured ellipticity ($\epsilon^{\rm obs}$) cannot be directly converted to the shear measurements. However, as the intrinsic ellipticities of galaxies are random, therefore, by taking average over large galaxy samples, $< \epsilon^{\rm int} > \approx 0$. Although the average of $\epsilon^{\rm int}$ is zero, the distribution has finite width. Therefore, this scatter (noise) can contaminate the lensing measurement and introduce uncertainty. This is called the shape noise, and as shall be seen in later chapter (Section 3.2.3), we treat this noise as the dominated source of uncertainty to the lensing measurements of surface mass density.

Although the measured ellipticity caused by lensing is related to the shear measurements, generally, the relationship between the measured shear and surface mass density is hard to interpret. For galaxy-galaxy lensing (when the lens and source are all galaxies), the tangential shear is directly related to the projected excess surface density of lens. However, in some cases, the lens might not be galaxies or in circular symmetry, such as the filaments in our study. Therefore, a method needs to be introduced to connect the measured shear field with the construction of the convergence field (surface mass density field) (Kaiser & Squires, 1993) by the following equation:

$$\kappa(\boldsymbol{\theta}) = -\frac{1}{\pi}\int d^2\theta' \frac{D_i(\boldsymbol{\theta'}-\boldsymbol{\theta})\gamma_i(\boldsymbol{\theta'})}{(\boldsymbol{\theta'}-\boldsymbol{\theta})^2}. \tag{1.32}$$

The above integration may be evaluated using Fast Fourier transformations. $D(\boldsymbol{\theta})$ is the lensing kernel. The Fourier transform of the lensing kernel, $\mathscr{D}_i(\boldsymbol{k})$, is defined as:

$$\mathscr{D}_i(\boldsymbol{k}) \equiv \begin{bmatrix} (k_1^2 - k_2^2)/k^2 \\ 2k_1k_2/k^2 \end{bmatrix}. \tag{1.33}$$

Equation 1.32 indicates that the convergence field can be reconstructed as long as the shear can be measured as a function of the angular position θ on the lens plane.

Weak gravitational lensing only causes slight distortion to the image of background sources (typically at the percent level), and this effect is also vulnerably affected by other contamination. The detection of objects such as filaments, which have relatively low density

contrast, are difficult to detect through weak gravitational lensing. For a given survey, one powerful way to mitigate the problem is to combine the signal coming from many foreground objects over the survey area: one can stack filaments together and obtain the convergence field by averaging them, which can effectively increase the significance of detection and reduce the noise.

1.1.5 Three-point Correlation Function

Comparing the measured signal of filaments to the theoretical model is not straightforward. One way to model the stacked filaments is the galaxy-galaxy-convergence three-point correlation function (3PCF). This function characterizes how a third point is correlated with the other points. In our case, those two points are fixed as the cluster positions (LRG positions), so the calculation of three-point correlation function gives the convergence at another point around the clusters. A detailed derivation of the convergence map using three-point correlation function is presented in Clampitt et al. (2014). Here we only briefly summarize the key points in the derivation.

The projected 3PCF among two haloes at fixed $\vec{\theta_1}$, $\vec{\theta_2}$ and excess mass κ at $\vec{\theta_3}$ is expressed as follow:

$$\zeta_{gg\kappa} \equiv \langle \delta_g(\vec{\theta_1})\delta_g(\vec{\theta_2})\kappa(\vec{\theta_3})\rangle, \tag{1.34}$$

where δ_g is the projected 2D galaxy overdensity and κ is the convergence field. $\vec{\theta_i}$ is the angular diameter that relates the 2D projected distance and line-of-sight comoving distance ($\vec{x_i} \equiv \chi_i\vec{\theta_i}$). δ_g can be expressed in terms of the matter overdensity δ_m:

$$\delta_g = \int d\chi p(\chi)\delta_h^{(3D)}(\chi) = \int d\chi p(\chi) b\delta_m(\chi), \tag{1.35}$$

where the integral is conducted over the line-of-sight separation and b is the linear galaxy bias assumed to be a constant during the calculation. The convergence field is the integral of matter overdensity field along the line-of sight but weighted by the inverse of critical density:

$$\kappa = \int d\chi \Sigma_{\rm crit}^{-1}(\chi, \chi_{\rm source})\overline{\rho_{m,0}}\delta_m(\chi), \tag{1.36}$$

where the critical density is evaluated at a fixed source plane at $\chi_{\rm source}$.

To evaluate the galaxy overdensity, we still need to specify $p(\chi)$, which characterizes the probability distribution of haloes along the line-of-sight. As shall be mentioned in Chapter 3, to accurately locate the positions of LRG halos in the redshift space, we use

the spectroscopic redshift taken from SDSS-BOSS. However, the actual location of clusters (or the true separation between two LRGs in configuration space) can be affected by the peculiar velocities of halos in the large-scale structure. Therefore, to model these uncertainties, one LRG halo is fixed at its known redshift, while the distribution for the other LRG has some finite width around the first LRG due to the presence of uncertainties:

$$p_1(\chi_1) = \delta_{\rm D}(\chi_1 - \chi) \tag{1.37}$$

$$p_2(\chi_2) = \frac{1}{\sqrt{2\pi}\sigma_{\rm scatter}} e^{-(\chi_1-\chi_2)^2/(2\sigma^2_{\rm scatter})}, \tag{1.38}$$

where $\sigma_{\rm scatter}$ is chosen to be a constant ($\sim 8\ h^{-1}$Mpc) estimated by Clampitt et al. (2014) from the peculiar velocity effect. Plugging Equation 1.35 and 1.28 into Equation 1.34, the 3PF can be written as:

$$\zeta_{gg\kappa} = \int d\chi_1 d\chi_2 d\chi_3 p_1(\chi_1) p_2(\chi_2) \Sigma^{-1}_{\rm crit}(\chi_3, \chi_{\rm source}) \overline{\rho_{m,0}} b^2 \langle \delta_m(\chi_1\vec{\theta_1}) \delta_m(\chi_2\vec{\theta_2}) \delta_m(\chi_3\vec{\theta_3}) \rangle, \tag{1.39}$$

and the computation of $\langle \delta_m(\chi_1\vec{\theta_1}) \delta_m(\chi_2\vec{\theta_2}) \delta_m(\chi_3\vec{\theta_3}) \rangle$ is conducted in Fourier space using its Fourier analogue, the bispectrum, denoted as $B_{\rm mmm}$. The bispectrum is computed using the perturbation theory (Bernardeau et al., 2002) as we are studying the weak gravitational lensing caused by the structures in the quasi-linear regime. Eventually, the projected 3PF can be numerically evaluated by integrating over the Fourier k-space, which can be computed using the following formula:

$$\zeta_{gg\kappa} = \frac{\Sigma^{-1}_{\rm crit}(\chi_{\rm lens}, \chi_{\rm source})}{\sqrt{2\pi}\sigma_{\rm scatter}} \overline{\rho_{m,0}} b^2 \frac{1}{(2\pi)^3} \int_0^\infty dk_A \int_0^\infty dk_B \int_0^{2\pi} d\phi \times k_A k_B B^{PT}_{mmm}(k_A, k_B, -k_{AB}) J_0(\sqrt{\alpha^2 + \beta^2}), \tag{1.40}$$

where k_A and k_B are the Fourier conjugate to the real space variable χ_1 and χ_2 (or the so-called wavevectors), and $k_{AB} = \sqrt{k_A^2 + k_B^2 + 2k_A k_B \cos\phi}$ where ϕ is the angle between k_A and k_B. $J_0(x)$ is the zero-th order Bessel function, where $\alpha \equiv k_A R_1 + k_B R_2 \cos(\phi - \theta)$ and $\beta \equiv k_B R_2 \sin(\phi - \theta)$. θ is the angle between the projected R_1 and R_2, which can be computed for given R_1, R_2 and the 2D projected $R_{\rm sep}$ between two LRG haloes.

Using Equation 1.40, for a given value of $R_{\rm sep}$, lensing redshift ($\chi_{\rm lens}$) and source redshift ($\chi_{\rm source}$), we can then predict the convergence field at any point around two clusters. Predictions of the convergence map around haloes for different chosen $R_{\rm sep}$ are presented in Clampitt et al. (2014). To compare with the measured shear data, the modelled convergence map can be further transformed into the predicted shear map using the Kaiser-Squires algorithm (Kaiser & Squires, 1993) introduced in previous section.

1.2 Integrals of Motion and Action Variables

In this section, we introduce the concept of integrals of motion and emphasize its significance in solving the equations of motion of a system. Generally, (isolating) integrals of motion are one-valued function of the phase space coordinate along an orbit. Finding these functions is of great importance as one integral of motion can help to reduce the dimensionality of trajectory by one, which largely reduces the number of unsolved equations of motion of a system. As seen in this section later, the best choice of integrals of motion is the action variable. Since action variables are conserved, the action variables of various structures (such as stream stars) would be clustered on small scales (in action space) due to their common origin. The calculation of action variables depend on the potential of a system. As long as the correct potential is applied, the action variables will remain conserved as long as the host potential evolves adiabatically. This principle can be used as a probe of underlying gravitational potential of a system, which will be discussed in Chapter 2 in more details.

1.2.1 Integrals of Motion

Having seen the application of Equation 1.7 and Equation 1.11 on large cosmological scale, in this section, we would like to see how these two equations help to describe the evolution of small-scale structure (for instance, a galaxy) with some certain assumptions. Still, we do not expect to solve the full 6-D CBE explicitly. However, there is one particular situation (or solution) that might be intriguing. When stars are moving in a time-independent potential field, $\Phi(\boldsymbol{x})$, we introduce a quantity that remains invariant along the orbit of a star. This quantity is called the integral of the motion, $I(\boldsymbol{x}, \boldsymbol{v})$. For an axisymmetric system, the energy, E, and z component of angular momentum, $\boldsymbol{L}_z$, are integrals of the motion. By definition, the full derivative of these quantities should be zero, which gives:

$$\frac{dI}{dt} \equiv \frac{\partial \boldsymbol{x}}{\partial t} \cdot \nabla I + \frac{\partial \boldsymbol{v}}{\partial t} \cdot \frac{\partial I}{\partial v} = 0. \tag{1.41}$$

Comparing the above equation with Equation 1.7, we find that as long as the probability distribution is a time-independent function, then based on the definition of integral of the motion, f is automatically a solution to the CBE. Therefore, we have the following theorem: If the probability function, f, is a integral of motion (or any time-independent functions in phase space that depends on $(\boldsymbol{x}, \boldsymbol{v})$ through integrals of motion only), then f is a solution to the CBE. This is the Jeans theorem. This theorem emphasizes the importance

of finding integrals of motion of a system. Finding the integral of motions helps to specify the distribution function, then in principle, one can use Equation 1.11 (Jean's Equation) to determine the potential of a system. However, finding integrals of motion may not be a easy task in phase space. Here, we choose a specific set of integrals of motion, the action variables, and use them to infer the potential of a system (the Milky Way), which will be discussed in following sections.

1.2.2 Action variables

The orbit of a particle is the most basic element of a collisionless system. In our case, while considering only regular orbits and a system with three degrees of freedom, such a orbit should admit at least three integrals of motion. However, sometimes it is not easy to find the integrals of motion in terms of phase space coordinates. Therefore, a specific type of coordinate transformation (canonical transformation) between two coordinate systems is extremely useful. This transformation can not only transform the original phase space coordinates $(\boldsymbol{x}, \boldsymbol{v})$ (or more generally, $(\boldsymbol{q}, \boldsymbol{p})$ in classical mechanics) to a new set of coordinates $(\boldsymbol{J}, \boldsymbol{\Theta})$, but keeps the Hamiltonian of the system and the area in phase space invariant. In such a coordinate system, the Hamiltonian is just a function of $\boldsymbol{J}$, $H(\boldsymbol{J})$ (originally was $H'(\boldsymbol{q}, \boldsymbol{p})$). Therefore, we have the form of Hamilton's equation of motions as:

$$\dot{J}_i = -\frac{\partial H}{\partial \Theta_i} = 0, \ \dot{\Theta}_i = -\frac{\partial H}{\partial J_i} = \text{constant}, \tag{1.42}$$

where H is the Hamiltonian of the system and in our case, it corresponds to the total energy of the system. $\boldsymbol{\Theta}$ is the angle variables that satisfies with $\Theta_i(t) = \omega_i t + \Theta_i(0)$. This means that an increase of 2π in the angle coordinates brings the point back to the location where $\Theta_i = 0$. $\boldsymbol{J}$ are the action variables and they are conserved quantities, which means they are integrals of motion. For a bound system, the action variables can be calculated as:

$$J_i = \frac{1}{2\pi} \oint_{\gamma_i} \boldsymbol{p} \cdot d\boldsymbol{q}, \tag{1.43}$$

where γ_i is the orbit section where the i-th angle, θ_i increases from 0 to 2π.

1.2.2.1 Calculating the Action Variables—Stäckel Approximation

Finding the action variables can be helpful to solve the states of a system completely, and obviously the calculations of action variables depend on the expression of system potential.

However, not all potential forms allow for an analytic form for actions. Therefore, some approximations and assumptions need to be made during the evaluations of actions. Here, we use "Stäckel approximation" (Binney, 2012), which is an algorithm implemented by `galpy` based on the assumption that an axisymmetric potential can be fitted by a "Stäckel potential" (de Zeeuw, 1985). For such a potential, action variables can be evaluated through numerical integration. Stäckel potential is defined in a confocal ellipsoidal coordinate system, where the original cylindrical coordinates can be transformed to the ellipsoidal coordinate though:

$$R = \Delta \sinh u \sin v; \qquad z = \Delta \cosh u \cos v, \tag{1.44}$$

where R, z are the original coordinates and u, v are the new spheroidal coordinates. In these coordinates, the Stäckel potential can be written as:

$$\Phi(u, v) = \frac{U(u) - V(v)}{\sinh^2 u + \sin^2 v}. \tag{1.45}$$

In such a potential, the action variables can be evaluated using Equation 1.43 after changing of variables:

$$J_r = \frac{1}{\pi} \int_{u_{\min}}^{u_{\max}} p_u(u) du, \tag{1.46}$$

$$J_z = \frac{2}{\pi} \int_{v_{\min}}^{\pi/2} p_v(v) dv, \tag{1.47}$$

where $u_{\min}$ and $u_{\max}$ are the values making $p_u(u) = 0$ and $v_{\min}$ is the root of $p_v(v) = 0$. $p_u(u)$, $p_v(v)$ are defined as:

$$\frac{p_u^2}{2\Delta^2} = E \sinh^2 u - I_3 - U(u) - \frac{L_z^2}{2\Delta^2 \sinh^2 u} \tag{1.48}$$

$$\frac{p_v^2}{2\Delta^2} = E \sin^2 v + I_3 + V(v) - \frac{L_z^2}{2\Delta^2 \sin^2 v}. \tag{1.49}$$

where I_3 is another integral of motion (apart from E and $L_z(J_\phi)$) that can be analytically found in a Stäckel potential. Therefore, for a specific point $(\boldsymbol{x}, \boldsymbol{v})$ in phase space, we can evaluate $p_u(u)$ and $p_v(v)$ using three integrals of motion (E, L_z, I_3), then using Equation 1.46 to calculate J_r and J_z. One problem left is that when transforming the cylindrical coordinates to the spheroidal coordinates (Equation 1.44), we have to evaluate the focal

length, Δ. The focal length can be easily obtained using the second derivative of the potential (Sanders, 2012), which can be calculated as:

$$\Delta^2 = z^2 - R^2 + \left(3R\frac{\partial\Phi}{\partial z} - 3z\frac{\partial\Phi}{\partial R} + Rz\left(\frac{\partial^2\Phi}{\partial R^2} - \frac{\partial^2\Phi}{\partial z^2}\right)\right)/\frac{\partial^2\Phi}{\partial R\partial z}. \tag{1.50}$$

All of the algorithms mentioned above are implemented through `galpy`.

1.2.2.2 Potential Model of the Milky Way

As mentioned in previous section, the calculations of action variables involve the potential of a system. Here, we assume the Milky Way is a simple combination of a central bulge, the disk and the dark matter halo, where the potential of halo is the form to be constrained. The potential model we are using is `MWPotential2014` built-in potential in `galpy` but with a slight modification.

The central bulge is modelled by a power-law density profile but with a exponential cut-off, and the form is given by:

$$\rho_b(r) \propto r^{-1.8} \exp\left[-\left(\frac{r}{1.9\text{ kpc}}\right)^2\right]. \tag{1.51}$$

As shall be seen in the next chapter, data used for analysis has a galactocentric radial range from 9 to 15 kpc. The contribution of bulge is actually negligible at such radii. However, we still include this component for completeness.

We use the Miyamoto-Nagai Potential profile to model the disk component with fixed parameters (Bovy, 2015), and it has a form:

$$\Phi_d(r,z) \propto -\frac{1}{\sqrt{R^2 + \left[3\text{ kpc} + \sqrt{z^2 + (0.28\text{ kpc})^2}\right]^2}}, \tag{1.52}$$

where R is the radial galactocentric cylindrical coordinate and z is the vertical distance to the galactic plane in the cylindrical coordinate.

Then, we use a simple spherical power-law profile to model the dark-matter halo:

$$\rho_{dm}(r) \propto r^{-\alpha}. \tag{1.53}$$

In `MWPotential2014`, the dark-matter halo is modelled as a NFW profile (Navarro et al., 1996). In this study, we use a power law potential instead as the data used for analysis

does not span over a large range, so the NFW profile can be locally well-approximated by a power-law profile in the considered range.

Finally, we set the normalizations of bulge, disk, and dark matter components, $\rho_b, \Phi_d, \rho_{dm}$, so that the fraction of radial force due to dark matter is f_h at the position of the Sun, while the ratio of force due to stellar bulge to disk is fixed to be 1:12. This gives the normalization of halo component as $0.65 f_h/(1\text{-}f_h)$, and the normalization of the bulge and disk component are fixed to be 0.05 and 0.60 respectively. Therefore, in the end we are left with two free parameters, f_h, and the power-law index of the density profile, α, which we aim to constrain using our method.

1.3 Summary of Observations

In this section, we briefly summarize several studies related to our two works presented in this thesis. In Section 1.3.1, we review some recent studies on constraining the Milky Way Potential by maximizing the small-scale clustering in action space. In Section 1.3.2, we first review some works on studying the filaments from either numerical simulations or from observations, then we briefly summarize some studies related to measure the mass-to-light ratio in different cosmic environments.

1.3.1 Constraining the Milky Way Potential in action space

Working in the space of actions can be convenient as the stellar distribution of tidal debris in action space can be much simpler than the distribution in the position-velocity space. As the action variables are adiabatic invariants (the Hamiltonian of the system slowly evolves with time) for an integrable system, stellar components in the tidal debris remain clustered on small scale because they are originated from the same progenitors, which can be helpful to recover the correct potential of the system.

Recently, the idea of constraining the potential by maximizing the small-scale clustering in action space has been used by many studies. Using G-dwarfs selected from SDSS/SEGUE, Bovy & Rix (2013) measured the vertical force ($F_z(R)$) and surface mass density ($\Sigma(R)$ within $|z| < 1.1$ kpc) as a function of galactocentric radius between 4.5 kpc to 9 kpc in the Milky Way. To achieve this, they grouped the selected G-dwarfs based on their metallicity and fitted them separately to a single distribution function in the three-dimensional action space, which is a function of the radial profile of the tracer density, the radial and vertical velocity dispersion with some free parameters. To compute the

actions, they assumed a four-component model of the Milky Way, consisting of a central bulge, a stellar and gas disk and a spherical power law dark-matter halo. Then, a likelihood function was defined in order to find the best-fit parametrized DF model that is consistent with the observed tracer distribution in action space [1]. With the determination of the best-fit radial profile of the vertical force, this profile was further used for fitting to the vertical force profile derived from the potential (with all free parameters varied). However, as the selected samples are located near the galactic disk, there were only weak constraints on the information about the dark-matter halo, where they found $\alpha < 1.53$ at 95% confidence. Although the constraints on the dark-matter halo are weak, this study still shows the feasibility of modelling complex stellar sets with DF in action space.

Adopting the same approach, Bovy et al. (2016) added observations of stream stars from GD-1 and PAL-5 and constrained the triaxial NFW halo profile of the Milky Way. By comparing radial and vertical force derived from real observations and simulated stream stars, they found a more spherical shape of the dark-matter halo than the predictions from numerical simulations. Still adopting the same approach, Bovy (2015) fitted the Milky Way potential as a combination of central bulge with power-law density profile and exponential cut-off, a Miyamoto-Nagai Potential disk, and a dark-matter halo with NFW profile after combining with additional data that can describe the Milky Way on wider scale. Although this profile, denoted as `MWPotential2014` in `galpy` package, is not necessarily the perfect model of Milky Way potential, it should be simple and realistic enough for our study here. As mentioned in Section 1.2.2.2, the potential model considered by this study is a slight modification of `MWPotential2014`.

Buckley et al. (2019) tried to reconstruct the potential by maximizing the phase-space density of tracer stars given an assumed set of parametrizations. They successfully constrained the mass and the King's radius of the M4 globular cluster from both simulations and real Gaia DR2 data. Furthermore, they also included the effect of measurement errors on the final constraints, finding that these errors act as "sources" to increase the phase-space volume (or equivalently, decrease the phase-space density), which would cause a higher estimate of the system's mass. They also demonstrated that for a gravitationally-disrupted systems (for instance, streams or tidal debris), instead of measuring the phase-space density in the position-velocity space, one needs to use action variables to infer the phase-space information and therefore constrain the potential. They tested this idea by using error-free simulated stars drawn from M4 globular cluster and found reasonable constraints on the potential. Due to the lack of complete six-dimensional phase space in-

[1] While evaluating the best-fit DF, parameters of the potential were also varied. However, as the primary goal of this paper was to find the radial profile of the $\Sigma(R)$ and $F_z(R)$, in this step, not all free parameters in the potential were varied

formation and their corresponding errors, in their study, this method was not applied to real data.

Vasiliev (2019) reconstructed the potential of Milky Way based on the model in McMillan (2017), where the Milky way potential model is divided into six axisymmetric components: a central bulge, a dark-matter halo, thin and thick discs plus molecular gas and HI discs. In order to constrain the potential, they used approximately 150 globular clusters and tried to maximize the stellar distribution function (under the assumption that the system is in steady-state equilibrium as the DF is a function of integrals of motion), $f(\boldsymbol{J})$, of these clusters in action space. The calculation of actions are potential-dependent and they also used Stäckel approximation to evaluate action variables for each tracers in a given potential. With the definition of the likelihood, the best-fit parameters can be found by a MCMC method. In this study, a circular velocity curve was also produced using the best-fit potential. They obtained a rotation velocity of 220-250 km/s in the radial range 8-50 kpc, which is slightly higher than the values produced by the default potential from McMillan (2017) but is inconsistent with the results produced by the `MWPotential2014` from Bovy (2015). Their curve also becomes shallower than the NFW profile prediction in the inner part and steeper in the outer part. This study demonstrates the viability of constraining the Milky Way potential using the globular cluster distribution in action space.

Instead of maximizing the stellar "clumpiness" (distribution function) in action space, Sanders & Binney (2013) chose to recover the potential by minimizing the misalignment between the stream track and its original progenitor orbit. To achieve this goal, they simulated a single globular-cluster-like stellar stream using a logarithm potential and constrained the potential by demanding that the angle and frequency differences of stream stars should lie along a straight line. They demonstrated that the correct potential could be recovered almost perfectly with error-free simulations. However, they also showed that the likelihood contour can be complex, presenting multiple peaks, and the addition of observational errors could create significant biases to the constraints. To mitigate these issues, they suggested binning the stream data in the observable space on the sky, and they also argued that their method can be better implemented with multiple and longer streams.

As shall be seen later, the methods closest to ours are those of Peñarrubia et al. (2012) and Sanderson et al. (2015, 2017). Peñarrubia et al. (2012) was among one of the first studies that tested the Minimum Entropy method by minimizing the measured energy distribution of simulated streams. Via simulations, they found that the entropy is indeed minimized at the correct potential, and the changes of system entropy are fairly sensitive to different parametrizations and forms of the adopted potential, suggesting that the Min-

imum Entropy Method could be a promising and powerful tool to accurately constrain the host galaxy potential.

Sanderson et al. (2015, 2017) tried to constrain the proposed Milky Way potential by maximizing the action-space clustering using multiple simulated streams. To characterize the clustering in action space, they utilized the definition Kullback-Liebler divergence (KLD) or "relative entropy" (Kullback & Leibler, 1951), which measures the relative difference between two probability distributions. In their studies, it was the difference between the actual stellar distribution in action space for a given potential parameters and the "shuffled" version of the distribution, where "shuffling" means randomizing different components for each action variable in order to break the correlation between actions. In our case, it is the actual probability distribution of stellar pairs in action space per interval $\mathrm{dln}\, D$ versus the random distribution. The key advantage of for this method is that one does not need to identify stream stars beforehand and fit multiple streams in a given potential simultaneously. In Sanderson et al. (2015), they successfully recovered the potential using both simulated stream stars with no measurements errors and with *Gaia*-like errors. The measurement errors, although they blur the substructures, do not erase the action-space clusterings. They also showed that a system with more streams and smaller observational errors could provide more accurate and precise constraints on the potential, which is in accordance with the expectations. In Sanderson et al. (2017), they additionally found that by finding the alignment between the stream stars and their orbit integrated in a trial potential could indeed roughly constrain the parameters of the potential, but it is neither accurate nor precise. In all, both of the two studies applied this idea using simulated tidal streams only, and the toy models used by the studies were spherical symmetric potential, where only one action variable is sensitive to the choices of potential (J_R).

To recover the potential, the studies mentioned above are not free of some *ad hoc* assumptions, such as virial/Jean's equilibrium, using tracers near the Galactic plane (circular motion), or some stream-finding algorithms. Even for the works that are close to our study, the same methodology has never been applied to real data. In our work, we present a method that uses two-point correlation of stellar pair distribution as a measure of small-scale clustering in action space. This approach is free from identifying any stream membership beforehand as well as the *ad hoc* assumptions mentioned before. Also, we further apply our method to real data, which is a more exciting step and a highlight of our work presented in Chapter 2.

1.3.2 Studies of Filaments and Mass-to-Light Ratio Under Different Environments

In this section, we briefly review some important works on the study of filaments and M/L ratio measurements. In Section 1.3.2.1, we first review some studies on investigating the properties of filaments using numerical simulations, then we summarize some works on tracing the filaments by mapping the galaxy redshift distribution using large surveys and we also review some works on studying the filaments with weak gravitational lensing. In section 1.3.2.2, we review some studies on measuring the mass-to-light ratio in different cosmic environments.

1.3.2.1 Current Studies of Filaments

Bond et al. (1996) illustrated the dynamical relationship between the formation of filaments and the high density nodes (peak-patches), where the massive clusters would from. They stressed that the formation and evolution of the structure is closely tied to the tidal force field. To illustrate this, they developed a peak-patch theory and showed that when the orientation of the shear tensor for a patch was specified, the surrounding patches would have the directions of shear tensor preferentially aligned with it, leading the formation of bridge-like structures between clusters. Using this theory, they showed the filament-dominance in the structure formation via simulations and demonstrated that high density nodes with appropriate separations could have strong filament between them.

Using N-body simulation, Colberg et al. (2005) attempted to find filaments between clusters. Based on their simulations, they found that close pairs with cluster-cluster connecting length smaller than 5 h^{-1} Mpc are always connected by filamentary structures, and clusters can posses one or more filaments: the number of filaments is on average increasing with the mass of clusters. They also investigated the radial density profile of straight filaments, which closely followed an r^{-2} profile (where r is the distance from the cluster-cluster axis) after a well-defined radius r_s between 1 h^{-1} Mpc and 2 h^{-1} Mpc. This density profile is generally treated as the theoretical model by other studies while determining the density and enclosed mass of the filaments.

Higuchi et al. (2014) used N-body simulations to find filaments between massive haloes with masses greater than $10^{14}h^{-1}\mathrm{M}_{\odot}$ and found $\sim 20\%$ of filament candidates are straight filaments, which is consistent with the results obtained by Colberg et al. (2005). They also confirmed the viability of detecting filaments using stacking method. For a HSC-like survey, significant lensing detection of filaments (signal-to-noise ratio > 5) can be achieved by stacking around 300 filaments.

Due to the complex spatial pattern presented in simulations, classifying these structures accurately and efficiently can be challenging in practice. Several sophisticated filament finder algorithms, such as percolation analysis, characterizing the structure based on the tidal force field, or tracing the mass distribution evolution in phase space, have been developed over the past few years. A detailed comparison of different structure identification algorithms are presented in Libeskind et al. (2018) (and relevant references therein). A diversity is found while investigating the properties of Cosmic Web using different algorithms, which is expected due to different means and definitions of structure features employed by various methods. However, similarities are also found across the methods, such as voids take up the most volume of the Cosmic Web and massive haloes indeed reside in the highest density region of the Cosmic web, etc.

Observationally, the presence of filaments have been proved by many surveys through tracing the redshift distribution of galaxies on the sky. Gregory et al. (1981) used 116 galaxies from the Perseus supercluster showing that a filamentary structure is clearly seen in the supercluster. Zeldovich et al. (1982) found a string of galaxies between Virgo and Coma supercluster from the a slice of supergalactic Y and Z plane. Using galaxies with $m_B \leq 15.5, 8^h \leq \alpha \leq 17^h$ and $26.5^\circ \leq \delta \leq 32.5^\circ$, de Lapparent et al. (1986) plotted the observed velocity and right ascension of those galaxies and showed that galaxies are distributed in elongated structures surrounded by empty regions. Oort (1983) summarized several supercluster observations from previous works and found most superclusters are elongated and contain complex filamentary structures. He noticed that clusters in the superclusters are tend to be aligned with their neighbouring clusters and are possibly also aligned with the elongation of superclusters. They also found evidence that between two superclusters, they might be connected with filaments as well. Since then, the presence of filaments has been observationally confirmed with many large galaxy redshift surveys. From the Las Campanas Redshift Survey, Shectman et al. (1996) plotted the galaxy distribution as a function of RA and redshift. Structures such as filaments and voids are more visible and clearly demonstrated in their plot. Gott et al. (2005) projected SDSS galaxies, detected within declination $-2^\circ < \delta < 2^\circ$, on a flat plane using their comoving distance (assuming a flat-ΛCDM model) and right ascension. They found a prominent web-like structure (the SDSS Great Wall) extending from 8.7h to 14h at a median comoving distance of ~ 310 Mpc from us. Similarly, Lietzen et al. (2016) also detected a large-scale overdensity pattern (the BOSS Great Wall) at a mean redshift of 0.47 by using the CMASS data from BOSS catalogue. From their smoothed luminosity-density map (shown in sky coordinates), the galaxy distribution clearly indicates the presence of elongated large-scale structures. The same foam-like pattern on large scales is also manifested in the map of galaxy distribution from the completed 2dFGRS (Colless et al., 2003) survey, the 2MASS redshift survey

(Huchra et al., 2012) at relatively low redshift range, also from other surveys such as VIPERS field at larger redshift (Guzzo et al., 2014, using $\sim$ 55,000 galaxies between $0.5 \lesssim z \lesssim 1.2$).

However, due to their low density contrast, observational evidence of the dark matter content is limited. Even using tools such as weak gravitational lensing, stacked lensing is necessary. As predicted by Bond et al. (1996), the detection of filaments between clusters should be much easier. Therefore, the target for finding filaments was mainly focusing on massive clusters. Dietrich et al. (2012) reported the detection of a dark-matter filament between two massive clusters, Abell 222 and Abell 223, via weak gravitational lensing at 4.1σ significance level. In order to demonstrate that the mass bridge is indeed caused by a filamentary structure instead of the overlap of cluster haloes. They fit the observations with NFW profiles (for clusters) plus an additional filament component, and the final results suggested that the observations are highly preferred the model with filament than the model with halo triplet only. From the reconstructed κ map, the measured filament mass is $M_{\rm fil} = (6.5 \pm 0.1) \times 10^{13} {\rm M}_{\odot}$, which is consistent with the result obtained from the model fit within one standard deviation ($M_{\rm fil} = (9.8 \pm 4.4) \times 10^{13} {\rm M}_{\odot}$).

Jauzac et al. (2012) detected a large-scale filament within the massive galaxy cluster MACS J0717.5+3745 (at $z = 0.55$). Instead of using the traditional "parametric" fitting method, by applying a multi-scale reconstruction method, which is designed for irregular mass distributions, the resulting measured surface mass density is found as $(2.92 \pm 0.66) \times 10^{8} h_{74} {\rm M}_{\odot} {\rm Mpc}^{-2}$ with a 2D projected separation of $4.5 h_{74}^{-1}$Mpc. With the total mass estimate, by assuming a constant stellar-mass-to-light ratio for the galaxy population of massive cluster, they further measured a stellar mass fraction along the filament to $(0.9 \pm 0.2)\%$.

Over the past few years, lensing surveys with high precision and accuracy has been designed, and meanwhile, the survey area and source density of surveys are sufficiently wide and high for the stacking process for filament studies. Clampitt et al. (2014) studied the weak lensing signal between two halos by stacking approximately 135,000 LRG pairs selected from SDSS catalogue. To eliminate the shear contamination by halo pairs, they assumed spherical symmetry of LRG haloes and used four chosen points on the grid to null the spherically symmetric halo signals (A detailed procedure is outlined in their paper). They detected a residual shear signal of filaments at $\sim 4\sigma$ significance. They also compared the observations to two filament models: thick model based on the prediction of three-point correlation function, and a thin model using a 1D line of NFW haloes aligned with the filaments (halo-pair connecting line). They demonstrated that the thick model is a better fit to the data, suggesting that three-point correlations should be more reliable in modelling the filaments.

Kondo et al. (2019) detected the weak lensing measurements of filaments between CMASS LRG pairs. Using the galaxy shape catalogue in the Subaru Hyper Suprime-Cam (HSC) first-year data, they reported an excess lensing surface mass density detection at $\sim 3.9\sigma$ significance. Using the same lensing estimator and halo subtraction method as outlined in Clampitt et al. (2014), they measured the $\Delta\Sigma$ values as a function of distance perpendicular to the pair connecting line, which were fairly consistent with the theoretical prediction using the mock CMASS LRG pairs and source galaxies. From the mock data, they also demonstrated that the high statistical power of HSC survey should be able to provide more precise weak lensing measurements in the future.

Most recently, Xia et al. (2019) presented the weak lensing measurements of filaments between LRG pairs selected from BOSS LOWZ sample. They adopted a smaller $R_{\rm sep}$ selection criterion for LRG physical pair selection (3-5 h^{-1} Mpc instead of 6-10 h^{-1} Mpc as adopted by this study). For source galaxies selection, they combined three lensing surveys: KV450, RCSLenS, and CFHTLenS, where the galaxy shape measurement were reduced by $lens fit$. Using similar nulling techniques developed in Clampitt et al. (2014), they measured a residual shear field of filaments at 3.3σ significance. By fitting the observations to the parametric filament density profile motivated by Colberg et al. (2005), they obtained the total mass by integrating the best-fit profile from connecting line to infinity, and they found $M_{\rm fil} = (4.7 \pm 2.2) \times 10^{13} {\rm h}^{-1} {\rm M}_{\odot}$.

Apart from using weak gravitational lensing, Tanimura et al. (2019) reported a detection of filaments between two LRG haloes based on the thermal Sunyaev-Zel'dovich (tSZ) effect arisen from the inverse Compton scattering between CMB photons and inter-cluster hot gas. To enhance the significance of detection, the averaged Planck tSZ map was generated by stacking against $\sim$ 260,000 LOWZ LRG pairs, and the halo contributions were subtracted by using the outside regions of the LRGs. They obtained a 5.3σ detection of filaments between two LRG haloes from the stacked tSZ map. Using a similar method, de Graaff et al. (2019) reported a detection of filaments using Planck tSZ map stacked against CMASS LRG pairs, and they found a $\sim 3\sigma$ detection. From the tSZ map, they also evaluated the baryon content in filaments. In particular, de Graaff et al. (2019) used the stacked CMB lensing map from Planck to estimate the baryon density of filaments, suggesting that approximately 10% of the total baryons in the universe from the filament environment. However, this estimate is biased and incomplete due to the specific selection cuts while finding the filament.

Our work is a follow-up study of Epps & Hudson (2017). In their study, they reported a weak lensing measurements of stacked filaments between LRG pairs at $\sim 5\sigma$ level. LRG galaxies were selected from BOSS survey and source galaxies were selected from CFHTLens. To identify filaments, LRG physical pairs were constructed with redshift separation < 0.002

and 2D projected separation between $6h^{-1}$ and $10h^{-1}$ Mpc (which are also the selection criteria adopted by other studies). The excess signal was measured from the subtraction between physical and non-physical map, and they found $M_{\rm fil} = (1.6\pm0.3)\times10^{13}{\rm M}_{\odot}$ between LRG pairs. They further modelled the filament with three-point correlation function and found fair consistency with observations.

Apart from the mass estimate of the filament, in our study presented here, we also evaluate the luminosity density of filaments by stacking LRG pairs selected from entire SDSS-BOSS coverage. This large catalogue of sample provides us a clean luminosity map of filaments between LRGs. We conduct the luminosity analysis in two different SDSS bands in order to have a rough estimate to the stellar type of filament galaxies. Finally we compute the mass-to-light ratio of filaments and compare our results with the M/L values obtained from different cosmic environments. In next section we briefly summarize previous measurements of mass-to-light ratios at different environments.

1.3.2.2 Current Studies of Mass-to-Light Ratios

One straightforward way to investigate how dark the filaments might be is to measure the mass-to-light ratio (M/L, or stellar mass fraction, M_*/M) of the filaments. In reality, we do expect that M/L values should not be the same in different cosmic environments. To estimate how the M/L ratios vary with B-band luminosity and halo mass, Marinoni & Hudson (2002) assumed that the mass-to-light ratio is related to the luminosity of the virialized systems via a power law ($\Upsilon \propto L^{\gamma}$). By comparing their measured luminosity function to the Press-Schechter mass prediction (halo abundance matching), they found that for ΛCDM model, M/L ratios decrease as $L^{-0.5\pm0.06}$ till a minimum around L_* ($m \sim 10^{13}M_{\odot}h_{75}^{-1}$), then increase as $L^{0.5\pm0.26}$, which is consistent with the expectation. A similar trend has also been found by Yang et al. (2007), where they developed a group-finder algorithm and applied it to SDSS DR4 catalogue. By running a group-finder algorithm, the stellar mass and r-band luminosity of groups were obtained. To estimate the group halo mass, they assumed that the halo masses and the estimated luminosities/stellar masses are linearly correlated (still using halo abundance matching). They found that for a halo with mass $10^{13}M_{\odot}$, the average M/L ratio is approximately 240 $M/L_{\odot,r}$, and the M/L ratios rise with increasing halo mass, then flatten out for the most massive halos.

Sheldon et al. (2009) measured the mass-to-light ratio around MaxBCG galaxy clusters selected from SDSS. Mass density profile around the central BCG as a function of radial bin (2D projected distance to the cluster center) was determined by measuring the tangential shear of the background sources. The 2D projected luminosity density profile was measured by including all galaxies around the central cluster, then subtracting the

luminosity measurements around randomized central cluster to exclude the background contamination. The luminosity was measured in i band at $z = 0.25$ and only galaxies with luminosity values greater than $10^{9.5}h^{-2}L_{\odot}$ were included in the analysis. The 2D profile was further de-projected into 3D via Abel inversion to get the integrated mass-to-light ratio in the range from $25h^{-1}$ kpc to $30h^{-1}$ Mpc. Finally, they got that within r_{200}, the M/L of the cluster within the cluster versus M goes roughly as a power low with index $= 0.33 \pm 0.02$, and the M/L at large separations, which should approximate to a universal value, is $362 \pm 54hM_{\odot}/L_{\odot}$. Using the same method, by analysing $> 10^5$ SDSS MaxBCG groups and clusters, Bahcall & Kulier (2014) found the M/L value at large separation is $409 \pm 23hM_{\odot}/L_{\odot}$, which is consistent with previous studies. They further investigated the cumulative stellar mass fraction M_*/M as a function of radial bin and found the asymptotic value at large separation is around $1.0 \pm 0.4\%$.

Parker et al. (2005) estimated the mass-to-light ratio of galaxy groups (with mean $N_{\rm gal}$ per group ~ 4) at B band. They stacked a sample of ~ 120 galaxy groups identified from CNOC2 redshift survey to measure the tangential shear as a function of angular bins. Mass were then obtained by fitting an isothermal profile to the observed tangential shear profile. Light profile were traced by the luminosity of each galaxy that belongs to the group with proper weights. The averaged mass-to-light ratio integrated to $1h^{-1}$ Mpc is $185 \pm 28hM_{\odot}/L_{B,\odot}$, and this value does not significantly vary with the distance from the group center.

1.4 Outline of this Thesis

In this work, we present two studies focusing on probing the dark matter from the galaxy to the Cosmic Web. The thesis is divided into two part. In the first part of the thesis, we shall constrain the Milky Way potential by exploiting the small-scale clustering in action space (Chapter 2). In that chapter, we present how the correct potential maximizes the clumpiness of stellar distribution in action space. We test this method in both simulations (Section 2.3.1) and real data (Section 2.3.2). Using the best-fit potential model, we also recover the circular velocity curve within 9-15 kpc from Galactic centre (Section 2.4). Finally we compare our results with other previous studies and present some discussions regarding our methodology and results. This work has been released on arXiv (Yang et al., 2019).

In the second part of the thesis, we focus on investigating the mass and light of the filaments in Cosmic Web (Chapter 3). In that chapter, we start off by reconstructing the surface mass density map via weak gravitational lensing. We outline the stacking procedure

and other major procedures for implementing this purpose in Section 3.2 Then, following similar logic, we present the approach of evaluating the surface luminosity density of the filaments. The resulting Σ_L and Σ_{M_*} map of the filaments are shown in Section 3.3.3, and we also compute the total luminosity of the filaments by fitting a Schechter function over the observed number of filament galaxies, which shall be discussed in Section 3.3.4. At the end of this chapter, we discuss our results and further compare our measured quantities (M/L and M_*/M) with other studies and cosmic environments.

Throughout this work, we adopt a ΛCDM model with $\Omega_{m,0} = 0.3$, $\Omega_{\Lambda,0}= 0.7$ and $h \equiv H_0/(100 \text{ km s}^{-1} \text{ Mpc}^{-1}) = 0.7$

Chapter 2

Constraining Potential from Small-Scale Clustering in Action Space

In this chapter, we develop a novel idea that uses clustering in action space as a probe of the underlying gravitational potential: the correct potential should maximize the small-scale clustering in action space. Current theories and observational evidence suggest that the growth of structure in our universe is hierarchical, where smaller structures merge to form larger ones. During the formation of galaxies, however, smaller structures are tidally disrupted and, due to various relaxation mechanisms at play, the memory of their common origin in configuration space is erased. This makes identifying stars with a common origin nearly impossible. Nevertheless, the information regarding their common origin may still be present in the phase space of action variables because the action variables remain conserved as long as the host potential evolves adiabatically. When smaller structures are tidally disrupted in the Milky Way potential, their spread in action space remains much less than the rest of the stars in the Milky Way. Therefore, we expect small scale structure of the action space to contain the hierarchical tidal disruption/assembly history of the Milky Way (Afshordi et al., 2009). Since action variables are conserved, due to their common origin, the action variables of the various structures would be clustered on small scales (in action space). This principle can be used to infer the potential of the Milky way. If the action variables are estimated using an incorrect potential, the resulting quantity will not be conserved with the dynamical evolution. Therefore, the clustering of the stars, in action space, on small scales will be destroyed if we use the wrong potential. Conversely, using the correct potential will maximize the small scale clustering in this space. The application

of this principle in constraining the Milky Way potential is discussed in this chapter.

This chapter is structured as follows: in Section 1.2 of Chapter 1, we introduced the required theoretical background for our methodology, including parametrized models of the Milky Way potential and the computation of action variables. In Section 3.1, we briefly discuss the data sets we used and the selection cuts imposed on the raw data. Details related to the two-point correlation function and the likelihood test used in the action space are presented in Section 2.2. To check the viability of our method, we first apply it to simulations with only stream stars, where the streams are simulated using the Python package `galpy` (Version 1.3.0. See Bovy, 2015, for more details). Then we proceed to apply our method to the real observations taken from *Gaia* DR2, with the selection cuts listed in Section 3.1.3 from two different galactocentric radial bins. The results of the analysis are presented in Section 2.3.2. In Section 2.4, we discuss the shortcomings of, and future prospects for, our method. This work has been released on arXiv (Yang et al., 2019)

2.1 Observational Data

In this section, we would like to include a brief introduction and summary to the surveys used in this study, including the *Gaia* mission for the investigation of stellar kinematics in the Milky Way and the selection criteria applied to the real data analysis.

2.1.1 *Gaia*

To study the stellar clustering in the six-dimensional phase space, one requires the precise measurements of 6D information of a star. *Gaia* mission aims at understanding the formation and evolution on our own galaxy by mapping the three-dimensional spatial and three-dimensional velocity distribution of stars in the Milky Way. Programmed by the European Space Agency (ESA), the satellite as launched on 19 December, 2013, and the first data release, referred as *Gaia* Data Release 1, was published on September, 2016. It includes the position and *Gaia* "G" band magnitude measurements (with a magnitude limit at least brighter than 20) for around 1.1 billion stars, and also information related to the parallax and proper motions for more than 2 million stars. After another 22 months observation, the second data release, *Gaia* Data Release 2, was published on April, 2018. Compared with Data release 1, the second release includes the position and magnitude measurements (with a magnitude limit at least brighter than 21) for around 1.7 billion stars, and the parallax and proper motions measurements for nearly 1.3 million stars. A

full description of *Gaia* mission can be found in Gaia Collaboration et al. (2016) and a full description about Gaia data release 2 can be found in Gaia Collaboration et al. (2018).

Another improvement of data release 2 compared to release 1 is that the second release also includes the radial velocity measurements for more than 7 million stars, which allows us obtain the full 6D information of stars without cross correlating with other surveys. The Radial velocity Spectrometer (RVS) on *Gaia* mission is designed to have unprecedented precision at measuring the radial velocity and the chemical abundance of stars. The bandpass for RVS is within wavelength range 847-874 nm with an average spectral power resolution around 10,500 - 12,500. A full description about the design and the performance of this instrument is presented in Cropper et al. (2018).

2.1.2 Data selection

As mentioned in Section 2.1.1, with *Gaia* DR2 data set only, we can determine the full six-dimensional phase space information of stars. However, we still apply some quality cuts to the raw data in order to avoid some outliers with large measurement errors. In this regard, we retain samples with relative errors on the parallax, proper motion and radial velocity measurements smaller than 20 % for further analysis:

$$\frac{\Delta p}{p}, \frac{\Delta \mu_{\rm RA}}{|\mu_{\rm RA}|}, \frac{\Delta \mu_{\rm dec}}{|\mu_{\rm dec}|}, \frac{\Delta V_{\rm radial}}{|V_{\rm radial}|} < 0.2, \tag{2.1}$$

where, $p, \mu_{ra}, \mu_{\rm dec}, V_{\rm radial}$ are the parallax of a star, proper motion along the right ascension direction and declination direction, and the radial velocity measured by RVS (Section 2.1.1). There were approximately 7 million stars in total included in the catalogue (with radial velocity measurements). Around 5.6 million stars remained after applying these cuts. These stars are further transformed into galactocentric coordinates using `galpy` library (Bovy, 2015), where we assume the solar information $R_\odot$ = 8.122 kpc, $z_\odot$ = 0.025 kpc, $V_{x,\odot}$ = -11.1 km/s, $V_{y,\odot}$ = 245.8 km/s and $V_{z,\odot}$ = 7.8 km/s (Gravity Collaboration et al., 2018; Eilers et al., 2019). We shall discuss the choice of this usage of solar information and its impact on the final results in Section 2.4.

Apart from the quality cuts imposed to the raw data, we also applied the galactocentric radial cut ($R_{\rm gal}$ cut) in order to select two independent subsamples. We choose two different radial ranges: 9 kpc < R < 11 kpc and 11.5 kpc < R < 15 kpc. Besides, we also include a vertical distance cut (with respect to the galactic plane, z cut) to these sample: only stars with $|z| > 1$ kpc are kept. We hope the above $R_{\rm gal}$ and z cuts could be helpful to include more halo stars, which should presumably be better traces to the halo potential

than disk stars. After applying all of those cuts, approximately 61,000 and 16,000 stars are remained in each sample. We shall test the effect of these $R_{\rm gal}$ and z cuts, as well as the 20 % quality cuts on the constraints in Section 2.4 later.

2.2 Methodology

In this section, we discuss the approach adopted by this study, including the definition of the two-point correlation function in action space and how does this correlation function relate to the likelihood (or equivalently, the entropy of the system) in the parametrized space (f_h and α in the power-law potential).

2.2.1 Distance in Action Space

We are interested in the small scale clustering of the stars in the action space. There are many different measures of clustering which are useful for different purposes. Here, we use two-point correlation function as our measure of clustering, which should be one of the most straightforward ones. However, to define the correlation function, we need to have a measure of the distance. While this choice is not unique, we shall use the following measure to find the distance of two stars in action space

$$D = \sqrt{(\Delta J_R/\sigma_{J_R})^2 + (\Delta J_\phi/\sigma_{J_\phi})^2 + (\Delta J_z/\sigma_{J_z})^2}, \tag{2.2}$$

where ΔJ_i denotes the difference in the action coordinates of the two stars, while σ_{J_i}'s are standard deviations of J_i's over all stars. The reason why we normalize the difference in action by the standard deviation in the action space is that stream stars, due to their common origin, should be significantly more clustered than the background, i.e. $\Delta J \ll \sigma_J$. Since the background could be anisotropic in the action space, this normalization provides a more appropriate distance measure. We further discuss this choice in Section 2.4 below.

The calculation of σ_{J_i} can be affected by the outliers in the raw data or numerical artifacts in `galpy` action approximation. Therefore, another constraint is added to effectively exclude outliers out of the sample with $\frac{|J_i - \bar{J}_i|}{\sigma_{J_i}} > 3$. This choice of 3σ seems relatively arbitrary. However, we confirm that the constraints are not affected by this choice, as long as those outliers are safely removed. Also, there is only a small fraction of stars being cut off by the "3σ-cut" from the original catalogue. Therefore, we do not believe our results are significantly biased by this choice. We then use the remaining action variables that satisfy the above criterion to re-calculate the standard deviation.

2.2.2 Two-Point Correlation Function and Likelihood Definition

For points distributed randomly with a uniform distribution in a three dimensional action space, the probability of finding pairs at a separation between D and $D + dD$ is given by

$$\begin{aligned} &\mathcal{P}(D)|_{\text{uniform}} dD \propto D^2 dD \\ \implies &\mathcal{P}(\ln D)|_{\text{uniform}} d\ln D \propto D^3 d\ln D. \end{aligned} \tag{2.3}$$

However, the actual probability distribution $\mathcal{P}(\ln D)$ will be different from $\mathcal{P}(\ln D)|_{\text{uniform}} \propto D^3$ due to clustering in the action space. This clustering can be quantified using the 2-point correlation function $\xi(\ln D)$:

$$1 + \xi(\ln D) \equiv \frac{\mathcal{P}(\ln D)}{\mathcal{P}(\ln D)|_{\text{uniform}}} = \frac{D_{\text{max}}^3 \mathcal{P}(\ln D)}{3D^3 \int_{-\infty}^{\ln D_{\text{max}}} \mathcal{P}(\ln D') d\ln D'}, \tag{2.4}$$

where we used the fact that both $\mathcal{P}$ and $\mathcal{P}|_{\text{uniform}}$ should integrate to unity over the range $\ln D \in (-\infty, \ln D_{\text{max}})$.

The blue solid line in Figure 2.1 shows $\ln\left[\frac{P(\ln D)}{D^3}\right]$ as a function of $\ln(D)$ calculated by using the Gaia DR2 real data from galactocentric radius 11.5-15 kpc with [$f_h = 0.34$, $\alpha = 1.66$][1], where D is the normalized distance of pairs of stars in the action space (as shown in Equation 2.2), while $\mathcal{P}(\ln D)$ is its probability density over all pairs of Gaia DR 2 within our sample. According to Equation 2.3, if stars are uniformly distributed, a plateau is expected at small values of $\ln D$, and deviations from this plateau would correspond to clustering. As can be seen from this figure, at small distance, $\mathcal{P}(\ln D)$ roughly obeys uniform distribution as stated in Equation 2.3, and the probability drops significantly due to the lack of stellar pairs at large values of $\ln D$.

As it turns out, with certain assumptions, the statistical likelihood of any action-space distribution can be expressed in terms of $\xi(\ln D)$. The key idea here is to assume the star distribution in the action space is the Poisson sampling of a near-uniform background plus a random gaussian field. The correlation function of this random gaussian field encodes all the clustering information at small scale in the action space. This model is agnostic about the distribution function, $f(\mathbf{J})$ and instead relates the likelihood to the correlation function in the action-space $\xi(\ln D)$, after marginalizing over all possible $f(\mathbf{J})$'s. More explicitly,

[1] As shall be seen in Section 2.3.2, this corresponds to the best-fit halo potential recovered by our method

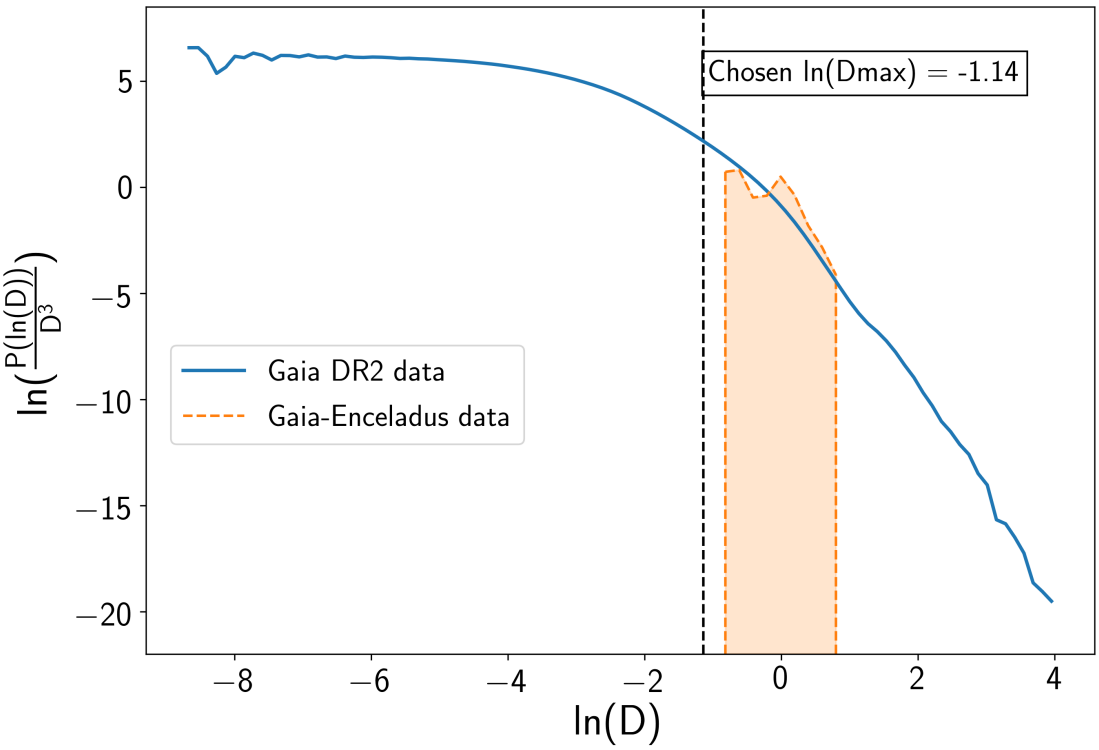

Figure 2.1: $\ln\left[\frac{P(\ln D)}{D^3}\right]$ versus $\ln(D)$ calculated by using the Gaia DR2 real data from galactocentric radius 11.5-15 kpc with $f_h = 0.34$, $\alpha = 1.66$. Here D is the normalized distance of pairs of stars in the action space, while $P(\ln D)$ is its probability density over all pairs of Gaia DR 2 within our sample. The correlation function computed using *Gaia-Enceladus* data (Myeong et al., 2018b) is over-plotted on the same figure (orange dashed line with shaded area). For comparison, the optimum chosen value of $D_{\rm max}$ for likelihood estimate is also shown as black vertical dashed line. Please see Section 2.4 for more details.

we find that the log-likelihood for a potential is given by

$$\ln \mathcal{L}(\text{data}|f_h, \alpha) \simeq \left\langle \sum_{\text{pairs}} \ln\left[1 + \xi(\ln D_{\text{pair}})\right] \right\rangle_{\text{pairings}} = N_{\text{pairs}} \int_{-\infty}^{\ln D_{\max}} \mathcal{P}(\ln D) \ln\left[1 + \xi(\ln D)\right] d\ln D, \tag{2.5}$$

A detailed derivation of this expression is presented in Appendix A. In this equation, N_{pairs} is the total number of pairs, i.e. half of the number of stars in the sample. Also, as shown in Figure 2.1, $\mathcal{P}(\ln D)$ obeys the scaling of uniform distribution only at small values of $\ln D$. Therefore, when evaluating the value of likelihood function, integration is terminated at a chosen value of $D_{\max}$. This $D_{\max}$ characterizes the scale of homogeneity in the action space background, and we shall discuss the choice of $D_{\max}$ in later section.

We further note that relative entropy (Kullback & Leibler, 1951) of the distribution $\mathcal{P}(\ln D)$, with respect to the uniform distribution, is defined as

$$S_{\text{relative}} \equiv -\int \mathcal{P}(\ln D) \ln\left[\frac{\mathcal{P}}{\mathcal{P}|_{\text{uniform}}}\right] d\ln D = -\frac{\ln \mathcal{L}(\text{data}|f_h, \alpha)}{N_{\text{pairs}}}, \tag{2.6}$$

i.e. the maximization of the likelihood function corresponds to minimizing the entropy relative to the uniform pair distribution. In other words, the best-fit values for the dark matter halo density produce the most non-uniform distribution of pairs in the action space. We should note that while this is similar to the criterion proposed by Sanderson et al. (2015), their relative entropy is based on phase space density in the action space $f(\mathbf{J})$, while our derivation in Appendix A shows that likelihood depends on the relative entropy of the pair distance probability distribution $\mathcal{P}(\ln D)$.

2.3 Results

Following the method introduced in previous section, in this section, we would like to focus on the discussion of our results. In Section 2.3.1, we first justify the viability of our result by implementing our approach to a simulation, which is composed by three streams simulated by using the `galpy` package. Subsequently, in Section 2.3.2, we continue to apply our method to real data selected from *Gaia* DR2 based on the criteria outlined in Section 3.1.3, managing to constrain the halo potential in two independent galactocentric radii.

Table 2.1: Initial conditions of progenitor for the simulation of stream stars (where galactocentric radius and velocity are normalized by solar radius value.)

Stream Number	Initialization							
	R	ϕ	z	V_R	V_T	V_Z	σ_v(km/s)	$t_{disrupt}$ (Gyr)
Stream 1	1.56	0.12	0.89	0.35	−1.15	−0.48	0.3	2
Stream 2	1.00	−0.05	0.001	−0.60	0.51	0.0086	0.3	2
Stream 3	1.20	−0.05	−1	−0.30	0.51	0.16	0.3	2

2.3.1 Simulation

To validate our method, we simulate the orbits of a few stars in a known parametrized potential of the same form and then applied the above mentioned analysis to check if we can recover the true parameters of the potential. The simulation includes three groups of tidal stream stars with different initial conditions of progenitors. This is achieved by using the built-in modelling method in `galpy` package (Bovy, 2014). One can specify the gravitational potential that stars evolve in, the method for action variables calculation, the initial conditions of progenitors, the velocity distribution of progenitors and the time when the disruption began. Initial conditions of the progenitors' orbit for three streams are tabulated in Table 2.1. The header of the table is organized in the order of R, ϕ, z, v_R, v_T, v_z, velocity dispersion (σ_v) and the disruption time ($t_{\rm disrupt}$). Each of the three streams consist of 3000 stream stars. With the simulated stellar trajectories, the action variables of stars can be calculated based on the Stäckel approximation as explained in Section 1.2.2.1. We now wish to test whether our proposed likelihood function (Equation 2.5) leads to constraints that are consistent with parameters that are used in our simulated host potential.

There are two free parameters in the expression of the dark matter halo density profile (Equation 1.53), f_h which fixes the normalization, and the logarithmic slope α. We choose the mass fraction of the halo $f_h = 0.35$ based on Table 1 in Bovy (2015), and we produce two sets of stream simulations with different choices of $\alpha = 1.70$ and 2.00 respectively. The progenitor stars are evolved in these two host gravitational potentials respectively. We then compute the action variables on a grid in the (f_h, α) space, and compute the corresponding likelihood function using Equation 2.5. The likelihood functions evaluated with 9000 simulated stream stars for both potentials are shown in Figure 2.2 (assuming $\ln D_{\rm max} \sim -1$). For each case, we find clear constraints on both parameters as expected,

in reasonable agreement with input parameters of the simulations, subject to caveats that we discuss next.

To determine the location and the uncertainties of the measurements at each $D_{\max}$, we fit the log-likelihood distribution with a quadratic function around its maximum. The assumption made by this procedure is that the likelihood only has a single peak that can be approximated by a gaussian distribution. In order to check the validity of this assumption, we plot the posterior of parameter at different values of $D_{\max}$, which is shown in Appendix B. As can be seen from Figure B.1, both of the posterior distributions for f_h and α have a single peak that can be reasonably approximated by gaussian. Also, the probability distributions do not drastically vary with the choices of $D_{\max}$. Therefore, we conclude that, at least for our simulated streams, our likelihood distribution is well approximated by gaussian statistics:

$$\chi^2 = \chi^2_{min} + (X_i - \bar{X}_i)F_{ij}(X_i - \bar{X}_i)^T, \tag{2.7}$$

where χ^2_{min} is given by the likelihood peak value by assuming $\mathcal{L} \propto e^{-\frac{\chi^2}{2}}$, and F_{ij} represents the Fisher information matrix for $X_1 = f_h$ and $X_2 = \alpha$. The covariance of the parameters is then given by the inverse of the Fisher matrix, F^{-1}_{ij}. There are six parameters to be determined in Equation 2.7. In practice, we fit for these parameters using a 3×3 grid around the peak of the likelihood.

As can be seen from Figures 2.2, the final measurements of mass fraction and index do not significantly vary with the choices of $D_{\max}$, with small error bars that further shrink by increasing the number of stars or $D_{\max}$. Based on these plots, we conclude that the measurement error for the simulations are dominated by systematic error, which is at the level of 1% for mass fraction f_h and 4% for the power law index α. This systematic error, while small, might arise due to the use of the Stäckel approximation to compute our action variables. We need to keep this in mind when we apply our method to real data, highlighting where further improvements may be needed when other sources of error are small.

As a further sanity check, we can verify that the parameters of potential found by the maximum likelihood test do correspond to the most clustering in the action space. In other words, the two-point correlation function defined in Equation 2.3 should be maximized for the correct potential. To verify this, we show how a 2D projection of stellar distribution in the action space varies with different choices of potential for both simulations. Figures are shown in Appendix B. As expected, the most compact distributions occur when parameters approach the correct values for the simulation, which is also reflected in the behavior of the two-point correlation function.

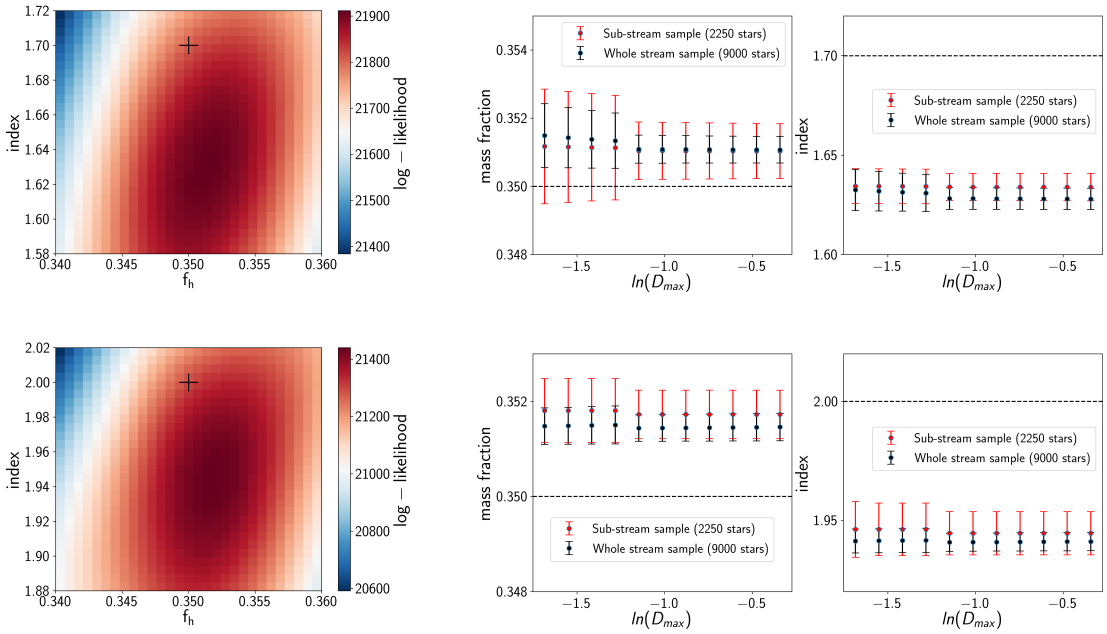

Figure 2.2: Top panel: likelihood test and error bar plot using $[f_h = 0.35, \alpha = 1.70]$. Bottom panel: likelihood test and error bar plot using $[f_h = 0.35, \alpha = 2.0]$. The initial inputs of the potential for both cases are indicated as "+" sign on the likelihood plot. Constraints given by the maximum likelihood are $(f_h = 0.35, \alpha = 1.63)$ and $(f_h = 0.35, \alpha = 1.95)$ respectively. On the right panel, error bars are determined based on the gaussian fitting discussed in Section 2.3.1, where the error bars determined with the whole 9000 stream stars and with the sub-sample are both shown. Obviously, the parameter recovered from likelihood function is inconsistent with the initial inputs, with approximate 4% systematic error for α estimate, and around 1% systematic error for f_h evaluation. Here, $\ln D_{\max}$ are chosen as $\sim$ -1.0.

To summarize, we have confirmed that our method to maximize likelihood (Equation 2.5) based on clustering in action space can yield reasonable constraints on simulated potentials, subject to small systematic errors of 1% (4%) on normalization and logarithmic slope. There also does not seem to be any significant dependence on the maximum separation of included pairs in action space $D_{\max}$. A more exciting step is to apply our method to real Gaia DR2 data to see how well it can constrain the Milky Way potential, which we shall do next.

2.3.2 Real Data

After confirming the reliability of the method, we proceed with our analysis using real data from Gaia DR2. The criteria for data selection were already discussed in Section 3.1. Let us now introduce some additional selection cuts. Recall that in the derivation of likelihood function (Appendix A), we assume the stellar distribution in the action space is a uniform background plus fluctuations. This assumption is more appropriate for halo stars in our galaxy, as disk stars have $J_z \simeq 0$. Additionally, as we are trying to constrain dark matter profile, which mostly occupies the Milky Way halo, halo stars should be better candidates compared to disk stars. Due to these considerations, we only select stars that have vertical distance to the galactic plane > 1 kpc[2].

Figure 2.3 shows the galactocentric distance and the tangential velocity distribution (in cylindrical coordinate) for all the data with relative measurement error smaller than 20% and $|z| > 1$ kpc. There are around 337,022 stars in total. As expected, the peak of radial distribution is around solar radius and the peak of V_T distribution is around the value of circular velocity at solar radius. Here, we assume $R_\odot = 8.122$ kpc, the vertical distance to the galactic plane $z_\odot = 0.025$ kpc, and the galactocentric velocity of the Sun $V_{x,\odot} =$ -11.1 km/s, $V_{y,\odot} = 245.8$ km/s, $V_{z,\odot} = 7.8$ km/s as taken in Eilers et al. (2019), but we shall discuss this choice further in Sec. 2.4. Due to the limitation of the computational time, another galactocentric radius cut is also applied to the data: We choose two different radial ranges 9 kpc $<$ R $<$ 11 kpc (hereafter real-data-9-11) and 11.5 kpc $<$ R $<$ 15 kpc (hereafter real-data-115-15). After applying all of these cuts, there are approximately 61,000 and 16,000 stars in each sample, respectively.

Now, taking the NFW profile as reference, the expected value of the α in the power law density profile should be within 1 to 3: For $r \gg r_s$, the density is proportional to r^{-3}, while for $r \ll r_s$, it goes to r^{-1}. Furthermore, Bovy & Rix (2013) used the assumption of Jean's

[2]We will discuss the effect of this cut, as well as the measurements error cuts on the final results later in Section 2.4

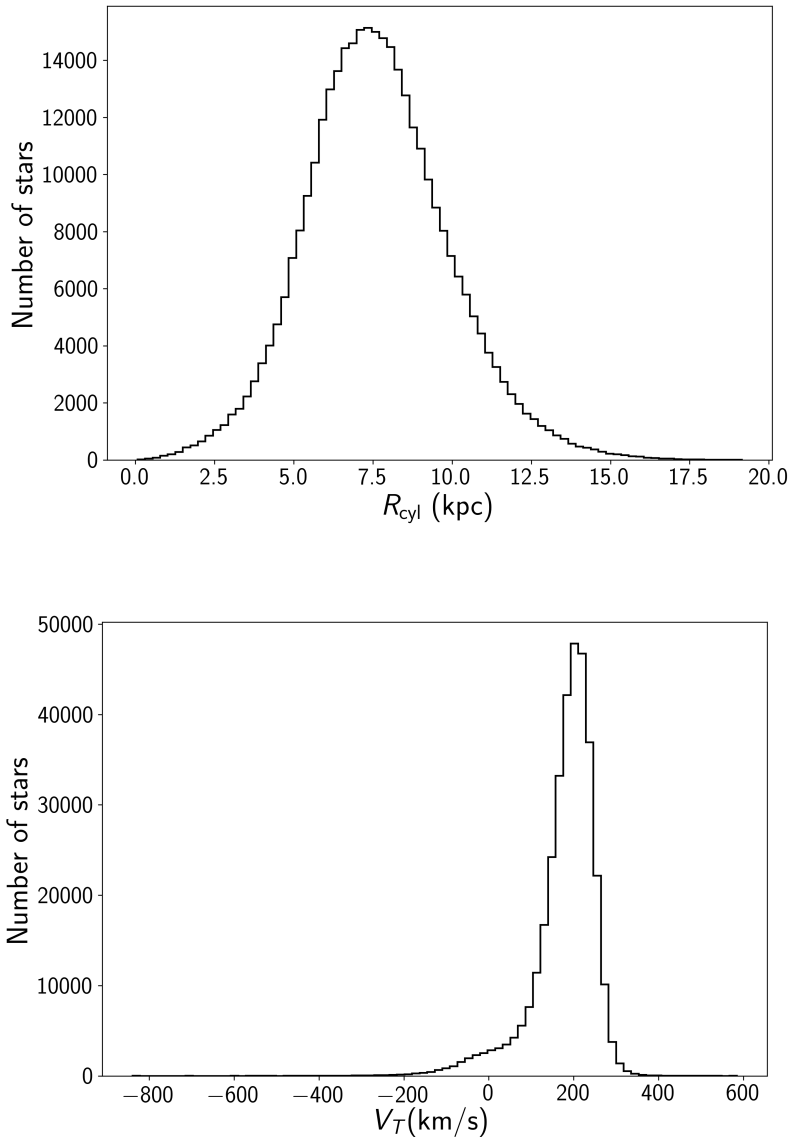

Figure 2.3: The galactocentric radius and tangential velocity distribution in cylindrical coordinates for the selected Gaia DR2 catalogue. Calculations are all conducted in cylindrical coordinate.

equilibrium for G-dwarfs from SEGUE survey to constrain $\alpha < 1.53$ (at 95% confidence) between R=4 kpc and 9 kpc. Therefore, to allow for a conservative prior, we consider the range:

$$\begin{aligned} 0.5 < \quad \alpha \quad &< 2.5, \\ 0.25 < \quad f_h \quad &< 0.55, \end{aligned} \tag{2.8}$$

for our dataset within 9 kpc to 15 kpc.

The likelihood plots showing the constraints on the mass fraction and the index for both radial samples are shown in the left panels of Figure 2.4. Although calculated within the same $\ln D_{\rm max} \sim -1$, the log-likelihood values for "real-data-9-11" is larger than those of "real-data-115-15" as there are more stellar pairs included in the more nearby sample. Furthermore, the error bar plots in Figure 2.4 show how the likelihood peaks (black solid points) and median constraints on parameters (blue hollow circles with error bars) vary with different choices of $D_{\rm max}$. We see that for $\ln D_{\rm max} \lesssim -1$, the f_h and α constraints are stable and robust to the choice of free parameter $D_{\rm max}$.

For the choice of $\ln D_{\rm max}$, we notice that for the more distant sample "real-data-115-15", there is a jump in the error bar plots for both parameters when $\ln D_{\rm max}$ is smaller than -1.5. Taking this into account, we treat -1.14 as our final choice of $\ln D_{\rm max}$. Results estimated at this point have the smallest uncertainties and the constraints are consistent (within error bars) for all $\ln D_{\rm max} \lesssim$ -1.14. We shall refer to this value as $\ln D_{\rm max,\ optimum}$. To be consistent, we use the same value of $D_{\rm max}$ for both Gaia samples. More discussions about choosing an appropriate $\ln D_{\rm max}$ are presented in Section 2.4 and Appendix C.

Having seen statistical constraints on both parameters from the likelihood plots, we would like to evaluate the true uncertainties of the measurements. However, the determination of uncertainties is more subtle compared with the simulations. Unlike simulations, where we found the posterior distribution of parameters had a sharp gaussian peak, we notice that the likelihood 2D plots have multiple peaks for real Gaia data. This can be seen more clearly in the 1D posterior distributions in Figure 2.5, where (depending on the choice of $D_{\rm max}$) there can be multiple peaks. As a result, it is no longer appropriate to simply assume the likelihood distribution is approximated by a gaussian. In particular, the jump in f_h around $\ln D_{\rm max} \simeq$ -1.5 in "real-data-115-15" sample is due to the change in relative heights of the two main peaks in posteriors shown in the top panel of Figure 2.5. Therefore, we calculated the median of the parameters using the full posterior distribution within our prior range (Equation 2.8), as it is a more robust statistical estimator than average whenever multiple peaks or outliers are presented in the distribution. The 68% confidence interval (68% CI) around the median, which can be also computed from the posterior distribution, is treated as the error on the parameter.

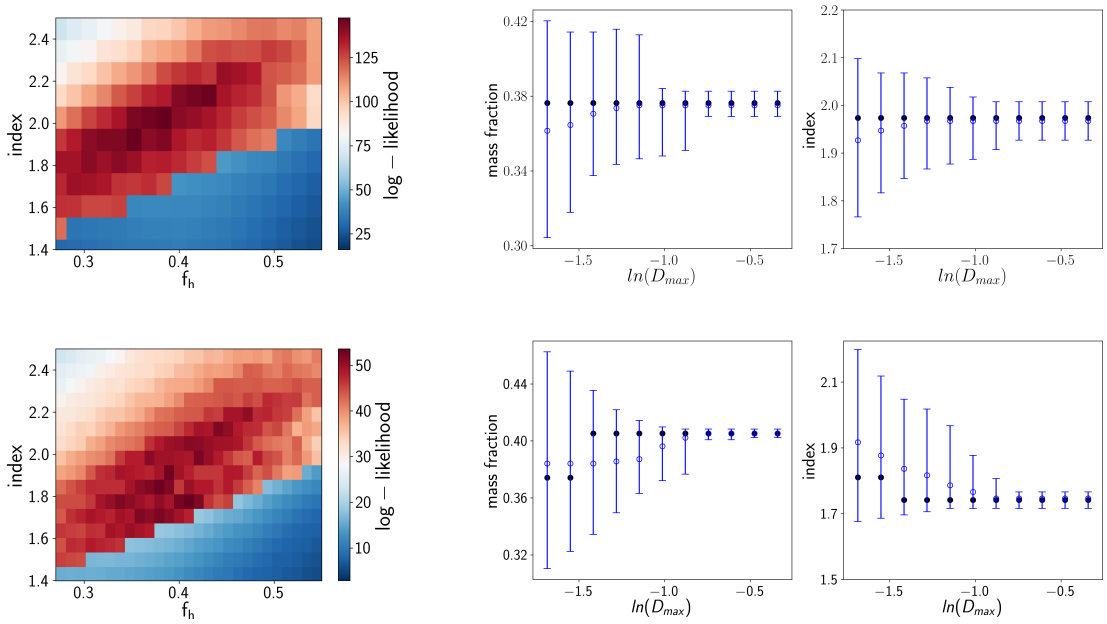

Figure 2.4: Top panel: likelihood test and error bar plot using sample from 9 kpc< R < 11 kpc. Bottom panel: likelihood test and error bar plot using stars from 11.5 kpc< R < 15 kpc. Constraints given by the maximum likelihood (f_h = 0.376, α = 1.974) for nearby sample and (f_h = 0.405, α= 1.741) for farther sample. On the right panel, error bars are determined based on the 68% CI around the best-fit median from the posterior distribution. Here, ln $D_{\rm max}$ are chosen as $\sim$ -1.0.

Error Type ($\|z\| > 1$ kpc)	[9.0 kpc <R<11.0 kpc] f_h = 0.375	α = 1.967	[11.5 kpc <R<15.0 kpc] f_h = 0.387	α = 1.786
Stochastic	$^{+0.029}_{-0.038}$	$^{+0.090}_{-0.070}$	$^{+0.024}_{-0.027}$	$^{+0.070}_{-0.181}$
Systematic	± 0.004	± 0.079	± 0.004	± 0.071
Total	± 0.034	± 0.112	± 0.026	± 0.144

Table 2.2: Constraints on (f_h, α) for both samples with selection cuts. Here, all sources of errors are summarized in this table: the statistical errors from the 68% CI around the median, the systematic errors from stream-only simulation, and the total error combining the statistical errors with systematic errors.

As can be seen in the error bar plots in Figure 2.4, the maxima of likelihood (black points) are all consistent with the median values within 68% CI (blue hollow points with blue error bars), and the error bars become larger for smaller $D_{\max}$, where fewer stellar pairs are included. When choosing $\ln(D_{\max})$ as -1.14, the constraints we get under both situations are summarized in Table 2.2.

Furthermore, recall that from simulated measurements in Section 2.3.1, we do expect an additional 4% percent systematic discrepancy for index measurements and an 1% off in mass fraction measurement. We do include these estimates in Table 2.2 as systematic errors, which can be combined with our stochastic errors to obtain the total expected uncertainties.

Let us now perform the same consistency checks we did in Section 2.3.1 for simulated data, and see how stars from real data are distributed in the action space. Figures 2.6 and 2.7 shows how (2D projections of) the stellar distribution in the action space, as well as its two-point correlation function change as we vary f_h or α in the Milky Way halo potential. As expected, the correlation function values for both radial ranges are maximized while the potential approaches the parameters that maximize the likelihood. However, unlike in simulations, instead of a compact cluster, stars in the action space present a more "lath-shaped" distribution, where most stars concentrate around $J_R \sim 0$ for the best-fit potential. In this situation, stars are extended along J_ϕ axis with no distinct J_R, J_z contributions, which indicates the property of circular (or disk) motion for most stars. This is not surprising as can be seen from the tangential velocity distribution in Figure 2.3; the disk stars are still the dominant component in our real data samples even though we performed a $|z| > 1$ kpc cut[3]. Just as in the simulations, stars within the potential that maximizes the small-scale clustering statistics (Figure 2.7) present the

[3]Note that the inclusion of disk stars does not contradict with our assumption made in the likelihood.

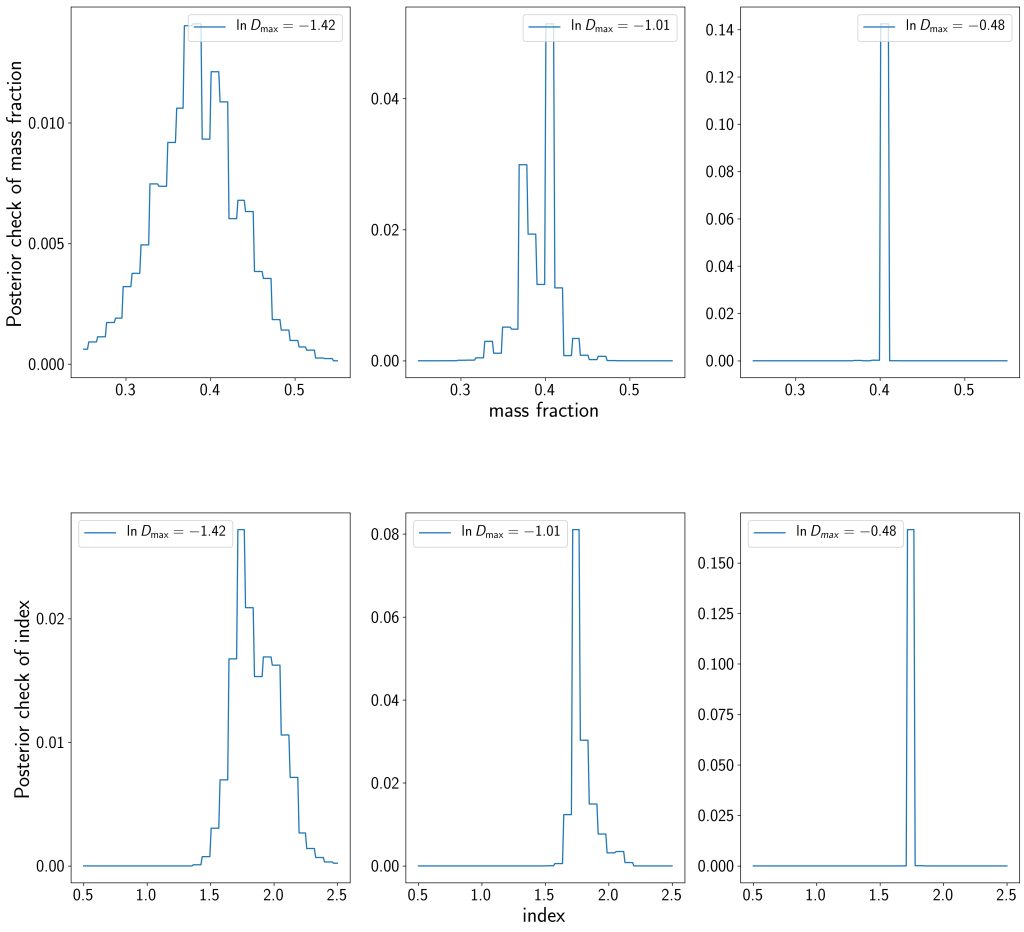

Figure 2.5: The posterior distribution of f_h (upper panel) and α (lower panel) using sample from 11.5 kpc< R < 15 kpc. Distributions are computed at three different values of $D_{\max}$.

most compact distribution in the action space. The fact that stellar distribution reduces to circular motion for the best-fit potential is in practice consistent with the traditional assumption of circular motion for disk stars, in order to estimate the mass of Milky Way galaxy. However, our method does not explicitly make this assumption, and thus can account for deviations from circular motion, effectively combining (thin+thick) disk+halo stars.

2.4 Discussion

In the previous analysis, we used some measurement error cuts and a vertical distance cut to real data. However, selection cuts to the raw data could cause unexpected biases in the measured parameters. To investigate the degree to which our results are sensitive to an arbitrary choice of z-cut, we randomly choose 90,000 stars from 9-11 kpc sample[4] and take all data from 11.5-15 kpc ($\sim$ 90,000 stars in total), without imposing any of the previous error or distance cuts [5]. The error bar plots are shown in Figure 2.8. For comparison, we also overplot the results obtained before using the sample with selection cuts. Generally, at same value of $D_{\rm max}$, the uncertainties on parameters are significantly reduced when using the data samples without selection cuts. To be consistent, for both radial ranges, we still take ln $D_{\rm max} \simeq$ -1.14 and check the corresponding constraints on f_h and α. The results are tabulated in Table 2.3.

Compared the constraints at the same $D_{\rm max}$ obtained previously but with error selection cuts with the results from full data set, we find a consistency in f_h constraint for both radial ranges. However, for α, we notice a 11% systematic discrepancy within 9-11 kpc, and an 7% discrepancy for the 11.5-15 kpc. For both radial ranges, the index estimates for the uncut sample are lower than those of the cut sample.

This systematic shift is primarily due to the selection cut to the vertical distance, z. Although some measurement error cuts are also imposed on the raw data, z distance cut seems to be the most severe: 80% of raw data survives the measurement error cuts, while

Disk is a mixture of correlated structures and uniform background, so the disk component can also contribute to the small-scale clustering signal. Also, our distance metric defined above has already included the effect of disk component (Equation 2.2).

[4]There are 607,257 stars in total, but we only choose a subset of the catalogue due to the limitation of computational time.

[5]However, we did apply some minimal cuts to the raw data in order to get rid of the unreliable observations, including $|z| < 10$ kpc and the absolute values of all three components of velocity in cylindrical coordinates are smaller than 500 km/s

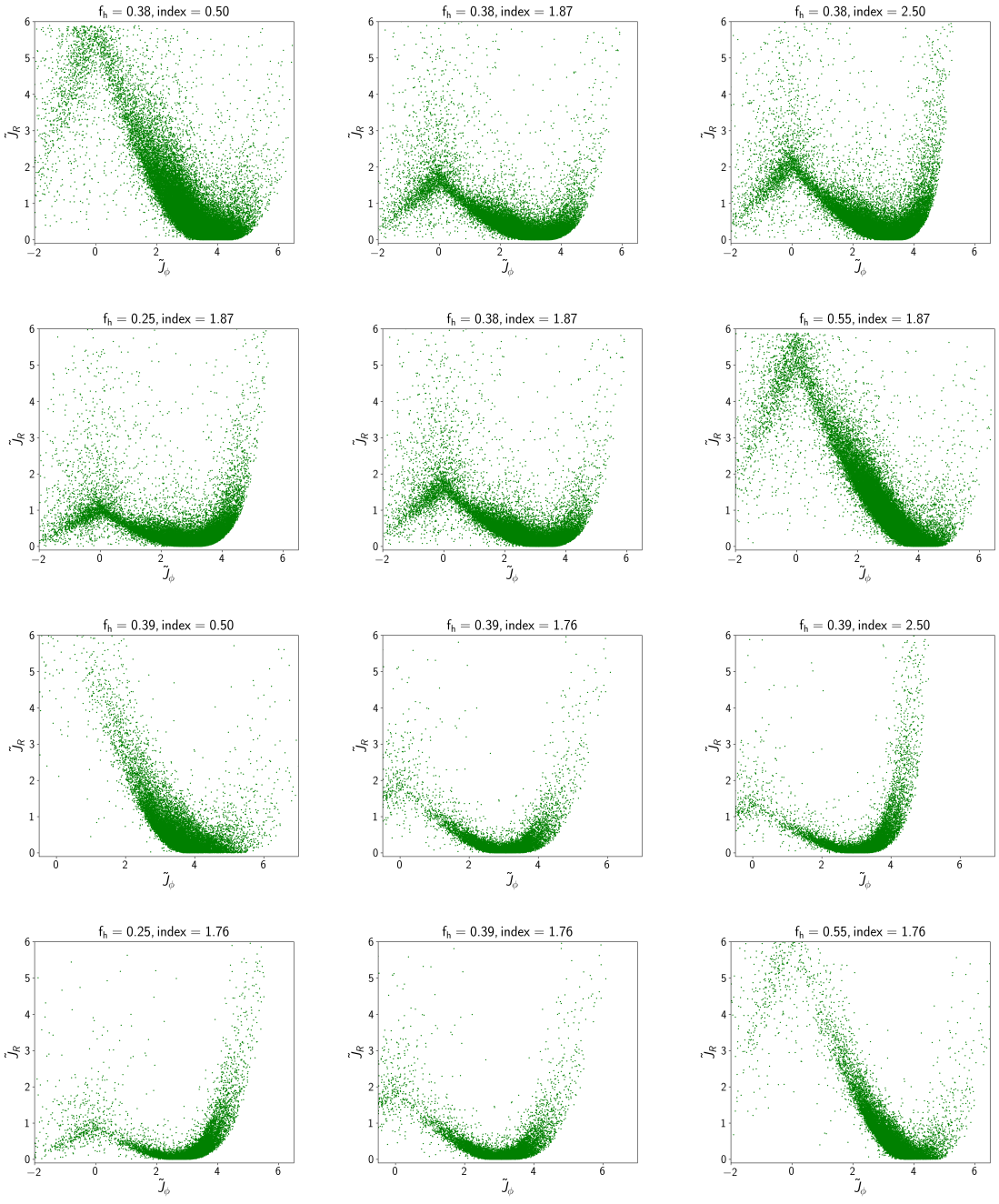

Figure 2.6: Stellar distribution in the $\tilde{J}_R$ and $\tilde{J}_\phi$ 2D projected plane varying with different choices of potential. First two rows show the stellar distribution for the first case of real data with fixed f_h (α) in the first (second) row. Last two rows show the result for the second case of real data. Interestingly, while approaching the potential that maximizes the likelihood (middle panels), stars are tend to be more disk-like and display the properties of circular motion.

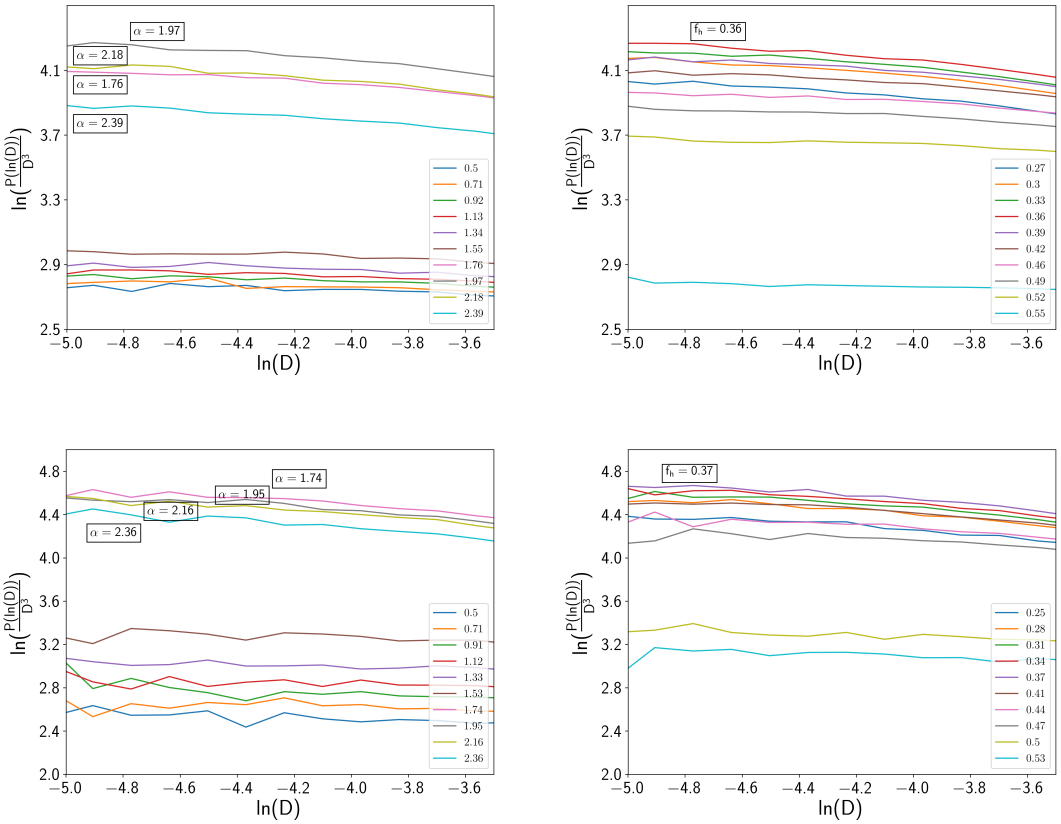

Figure 2.7: Correlation function, $\frac{P(\ln D)}{D^3}$, as a function of $\ln D$. Top panel: $\frac{P(\ln D)}{D^3}$ vs $\ln D$ for nearby sample with fixed f_h (α) on the left (right). Different colors indicate different parametrizations of potential. Bottom panel: $\frac{P(\ln D)}{D^3}$ vs $\ln D$ for farther sample with fixed f_h (α) on the left (right). As expected, while varying the choices of potential, the two-point correlation function is maximized around the correct value of parameter.

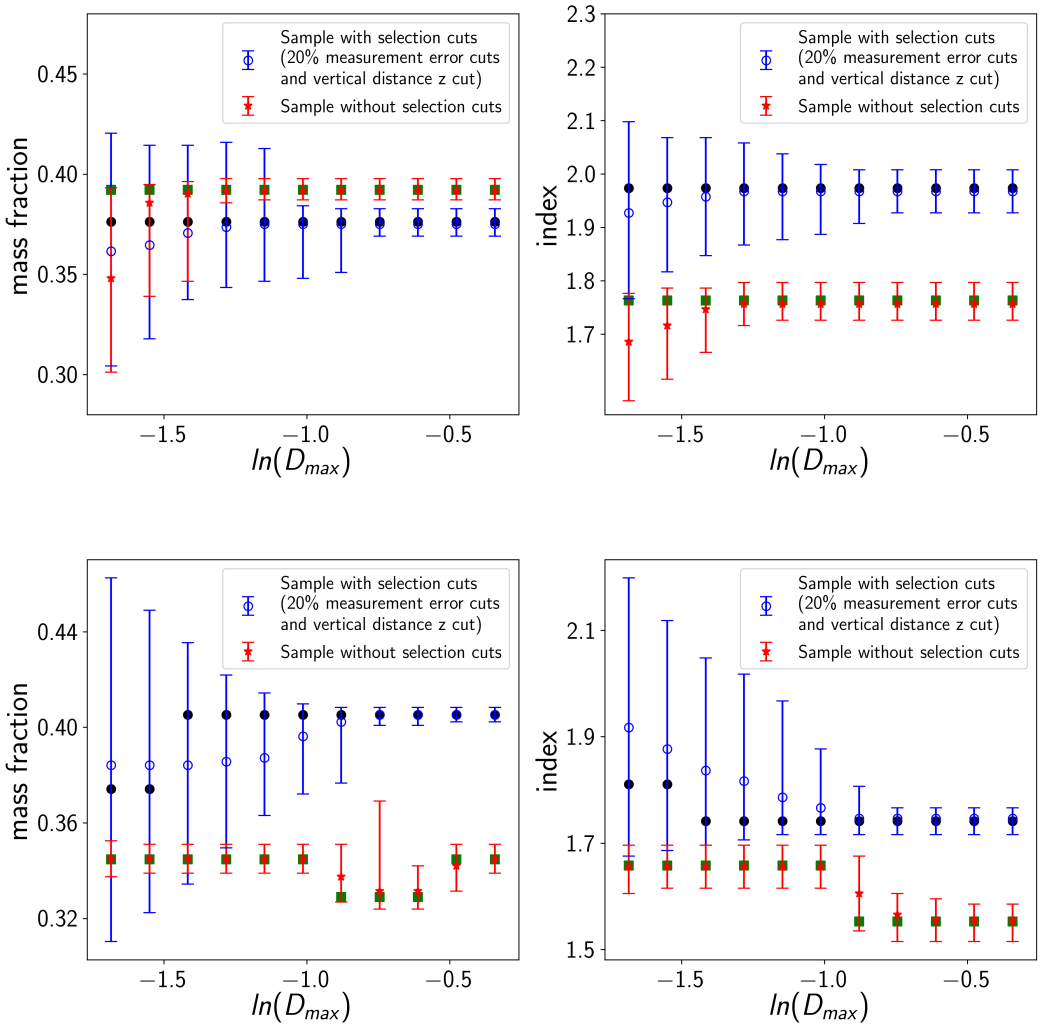

Figure 2.8: Error bar plot showing the comparison between the constraints on (f_h, α) obtained from sample with selection cuts and from sample without cuts. Top panel shows the results using stars from 9 to 11 kpc and bottom panel shows the results using stars from 11.5 to 15 kpc. Green square points shows the corresponding likelihood peak. The red star point and red line represent the best-fit median values. For comparison, the constraints on two parameters determined from the sample with selection cuts are also over-plotted on the same figure (blue hollow points with blue error bars).

Error Type (no $\|z\|$ cut)	[9.0 kpc <R<11.0 kpc] f_h = 0.392	α = 1.756	[11.5 kpc <R<15.0 kpc] f_h = 0.345	α = 1.656
Stochastic	$^{+0.004}_{-0.006}$	$^{+0.030}_{-0.040}$	± 0.006	± 0.040
Systematic	± 0.004	± 0.070	± 0.003	± 0.066
Total	± 0.006	± 0.078	± 0.007	± 0.077

Table 2.3: Constraints on (f_h, α) for both samples without selection cuts. Here, all sources of errors are summarized in this table: the statistical errors from the 68% CI around the median, the systematic errors from stream-only simulation, and the total error combining the statistical errors with systematic errors.

only around 6% remain after the $|z|$ >1 kpc cut is imposed. One possible reason for this systematic difference could be the inaccuracy of the simple analytic model for the disk potential used in Equation 1.52. It remains to be seen whether a more realistic model (e.g. using other datasets), or including the disk parameters in the likelihood marginalization, could lead to more consistent (and realistic) estimates. In order to account for additional potential systematic errors due selection cuts, we use the probability function defined in Appendix D, which yields our final constraints on the mass fraction of dark matter and the index in the localized density profile in Table 2.4.

After obtaining the constraints on both parameters in the dark matter halo density profile, we can translate them to less model-dependent constraints by computing the rotation curve (circular Keplerian velocity) of the Milky Way, as a function of distance from the centre. This result can then be compared to other studies that use different parametrizations and methods. To obtain a more robust estimation to the circular rotation curve, we evaluate the average and standard deviation of $v_{\rm circ.}(R) \equiv \sqrt{\frac{\partial \Phi_{\rm tot}(R,z=0)}{\partial \ln R}}$ given the likelihoods found from our different Gaia samples (Equation 2.5) over our prior range of f_h and α (Equation 2.8).

As the expression of disk potential is fixed (where we also fix $z = 0$ in Equation 1.52), its contribution to the total rotation curve (as well as that of the bulge) can be simply added to the halo part in quadrature. Right panel of Figure 2.9 displays the circular velocity curve obtained from this work (purple and black solid line) and its corresponding uncertainty shaded area. For comparison, the results obtained from Bovy (2015), McMillan (2017), Vasiliev (2019), and Eilers et al. (2019) are also shown in the same figure. McMillan (2017) used kinematic data from maser observations with (expected) near-circular motion to fit a Milky Way model with an NFW spherical halo, a stellar and gas disk plus a central

	[9.0 kpc <R< 11.0 kpc]	[11.5 kpc <R< 15.0 kpc]
f_h	$0.375^{+0.029}_{-0.038}$	$0.387^{+0.024}_{-0.027}$
α ($\|z\|$ >1 kpc, stochastic error only)	$1.967^{+0.090}_{-0.070}$	$1.786^{+0.070}_{-0.181}$
f_h	$0.392^{+0.004}_{-0.006}$	0.345± 0.006
α (no $\|z\|$ cut, stochastic error only)	$1.756^{+0.030}_{-0.040}$	1.656± 0.040
f_h	0.391± 0.009	0.351± 0.012
α (combined, stochastic+systematic error)	1.835± 0.092	1.687± 0.079

Table 2.4: Constraints on (f_h, α) from samples with and without selection cuts. Here, the top four rows tabulate results for different cuts (where only stochastic errors are included), while the final two rows show the combination of these results using the probability function as defined in Appendix D

bulge. Using a nearly identical model, Vasiliev (2019) assumed Jeans equilibrium of Milky Way globular clusters in Gaia data to constrain the gravitational potential. Eilers et al. (2019) also used Jeans equilibrium for Gaia luminous red-giant stars to determine the circular velocity of the Milky way over radial range 5 kpc $< R <$ 25 kpc. Although we approximate the localized halo density profile as a simple power law, our result is consistent with the estimation from Eilers et al. (2019) and relatively close to (but around 4% higher than) the best-fit NFW dark matter potential found in Bovy (2015) (`MWPotential2014`). However, the circular velocity (radial force) is about 5-10% (10-17%) smaller than the other two studies. Compared with these studies, our method might be more robust as it does not rely on assumptions of circular motion or Jeans equilibrium, and can be equally applied to halo or disk stars.

An important consideration for comparison to other measurements of circular velocity is our choices of the solar coordinates in the coordinate transformation. If we make the choice of solar coordinates consistent with the analysis of Bovy (2015) ($V_{\phi,\odot} = 220$ km/s and $R_\odot = 8$ kpc), as shown in the left panel of Figure 2.9, the recovered circular velocity curve is comparable to Bovy (2015)'s estimation but $\sim$ 9-17% lower than other three curves. However, the measurements of $V_\odot$ and $R_\odot$ have been progressively improving. For example, if we take the GRAVITY results (Gravity Collaboration et al., 2018) used in Eilers et al. (2019) to convert the coordinates, it does bring our curve close to their measurements. As shown in the right panel of the figure, measurements using GRAVITY solar coordinates are higher than Bovy (2015)'s and comparable to those of Eilers et al. (2019) from Gaia DR2, but are still significantly lower than maser and globular cluster measurements by $\sim$ 5-10%.

Let us now comment on our choice of distance (or metric) in the action space (Equation 2.2). The reason why we normalize action variables by their standard deviation to compute distance is partly due to the assumption we made in the likelihood derivation in Appendix A. Our derivation starts from a uniformly distributed background plus gaussian fluctuations which model clustering in action space. Therefore, the structures we consider should be on smaller scale than the background distribution in the action space. Since the extent of the background could be different in different directions in the action space, the normalization has the effective role of making the distribution homogeneous and isotropic, at least for $D \ll 1$, i.e. close pairs.

Note that here we ignore the covariance between different action variables. As a sanity check, we modified our distance definition to $D^2 = \sum_{i,j=1}^{3} \Delta J_i \Delta J_j \sigma_{ij}^{-2}$ (where σ_{ij}^{-2} is the inverse covariance matrix of J_i's over the entire sample) accounting for the correlation between action variables. We checked this using both our real data and simulations, and found no significant change in our results (e,g., for 11.5<R<15 kpc and same $\ln D_{\max}$, the

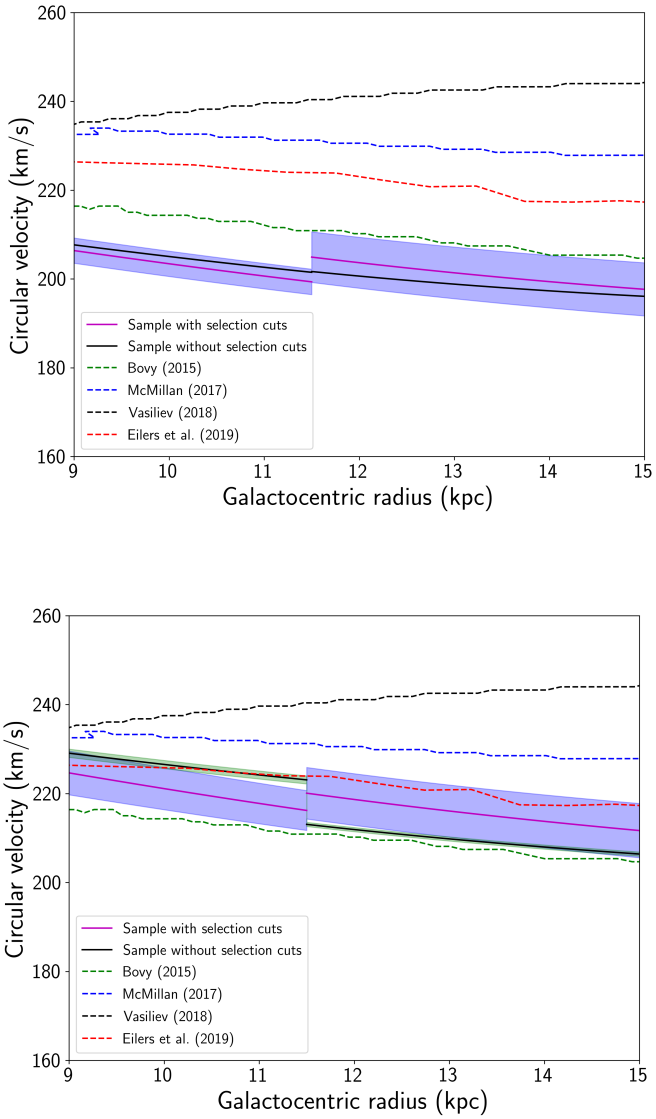

Figure 2.9: The total recovered rotation curves from 9 kpc to 15 kpc. Curve from this study is indicated as purple solid line (for the sample with selection cuts) and black solid line (or the sample without selection cuts). The shaded area indicates the 1-σ uncertainty. For comparison, curves computed from other studies are over-plotted. Top: curve calculated using the solar positions in Bovy (2015). Bottom: curve calculated using the solar information from new GRAVITY measurements.

relative changes in f_h and α are less than 1%).

For the choice of the free parameter $\ln D_{\rm max}$, the main criterion is that we do not expect the constraints on the parameters to significantly vary with $\ln D_{\rm max}$. Therefore, when $\ln D_{\rm max} \lesssim \ln D_{\rm max,optimum}$, the constraints on both parameters should not be a strong function of $\ln D_{\rm max}$ and also be self-consistent within error bars. Meanwhile, the measurements are better to be the least uncertain at the $\ln D_{\rm max,optimum}$. One of the reasons why we plot the median values (with error bars) and two-point correlation function, $\ln\left[\frac{P(\ln D)}{D^3}\right]$, as a function of $\ln D_{\rm max}$ (or generally, $\ln D$) is to see which $\ln D_{\rm max}$ value can give us stable *and* reliable constraints. Therefore, we do not expect the optimum choice of $\ln D_{\rm max}$ to be necessarily the same for different systems. This criterion is further explored in Appendix C, where we include a background in our simulations, leading to a different $\ln D_{\rm max,optimum}$.

Here, for real data analysis, we choose $\ln D_{\rm max,\ optimum}$ = -1.14 as mentioned in previous section, and we also choose the same range for the stream-only simulation for consistency (as under this specific case, neither parameters drastically change with $\ln D_{\rm max}$). To give more intuition about what the chosen $D_{\rm max}$ physically means, we calculate the values of two-point correlation function using the *Gaia-Enceladus* globular clusters data (Myeong et al., 2018b), which is shown as orange dashed line in Figure 2.1. By doing this comparison, we can see that maximum separations of the pairs we considered (within the considered $D_{\rm max}$) are slightly smaller, but comparable to the size of *Gaia-Enceladus* globular clusters distribution (also known as Gaia Sausage) in the action space, which indicate that the largest size of the structure contributing to the likelihood estimate is close to the characteristic size of the *Gaia-Enceladus* structure. We suggest that the choice of $\ln D_{\rm max}$ needs to be inferred from the behavior of two-point correlation function (as a point of transition from clustered streams to background) and $f_{h,\rm fit}(\alpha_{\rm fit})$ as a function of $\ln D$ or $\ln D_{\rm max}$, which does help to determine the point where our method and the estimations of parameters are still reliable.

As the assumption made in our likelihood derivation (Appendix A) is a uniform background plus a random gaussian field, one might be skeptical about the validity of systematic error estimations with a stream-only simulation as presented in Section 2.3.1. To improve this, we conducted another simulation with the inclusion of a background. Background stars are directly taken from Gaia DR2 and their action variables are calculated in a simulated host potential with [$f_h = 0.35$, $\alpha = 1.70$] using the 'Stäckel approximation'. Based on this, we subsequently randomize the stellar distribution in action space by adding a random gaussian scatter to each J_i respectively, then stars with randomized action distribution are transformed back to ($\vec{X}$,$\vec{V}$) in cylindrical coordinate using the TorusMapper code (Binney & McMillan, 2016). We combine three original streams with the simulated

background and evolve them with different choices of (f_h, α). We would like to test whether the parameters recovered by the likelihood function is consistent with the initial input of the simulated host potential (which is [$f_h = 0.35$, $\alpha = 1.70$]). A detailed analysis is presented in Appendix C. As can be seen from the error bar plots (Figure C.3), wherever the constraints do not significantly depend on $\ln D_{\rm max}$, the inclusion of background actually improves the measurements (at $\ln D_{\rm max,optimum}$, the constraint we get is $f_h = 0.352 \pm 0.003$, $\alpha = 1.678 \pm 0.058$ from quadratic fit, while for a stream-only simulation, we have $f_h = 0.352 \pm 0.001$, $\alpha = 1.634 \pm 0.014$). Therefore, we conclude that the systematic error estimated from a stream-only simulation should be conservative and can be propagated to further analysis.

There are other proposals to use the action-angle (or similar) variables to constrain the potential. Sanders & Binney (2013) use the correlations in the angle-frequency[6] space for stars of a single stream to constrain gravitational potential. For a true potential, the angle and frequency differences of stars in a long narrow stream should lie along a straight line. An incorrect potential could cause a misalignment between the stream orbit and the underlying progenitor orbit. By minimizing this misalignment, which is potential-dependent, they manage to recover the expected constraints to a spherical logarithm potential using a simulated tidal stream. While this method uses more information (i.e. angle variables) than ours, and thus can be potentially more precise, it requires identifying only stream stars and relies on the assumption of a cold stream, which does limit its precision and accuracy. Magorrian (2014) relates the clumpiness of the stellar action-space distribution to the potential and define a likelihood, where the stellar action-space distribution is drawn from a Dirichlet process. This study justifies the viability of constraining the potential using stellar action-space distribution. However, instead of assuming a specific functional form of stellar distribution in the action space, our method is independent of f($\mathbf{J}$) as we marginalized over all possible distribution of f($\mathbf{J}$) in the derivation of likelihood function (or more precisely, what we assumed here is the probability functional P{f($\mathbf{J}$)} is gaussian with an arbitrary 2-point function that only depends on D.). Methods introduced in Peñarrubia et al. (2012), Sanderson et al. (2015), Sanderson et al. (2017), and Buckley et al. (2019) are the closest compared with our methodology, which *minimize* relative entropy (or KLD) of a system in the space of action variables (or more generally, integrals of motion). However, these studies do not directly connect their statistical representations to the two-point correlation function in the action space as proposed in this study. Indeed, our derivation in Appendix A suggests that relative entropy of the distance distribution $P(\ln D)$ [rather than $f(\mathbf{J})$] is more directly related to the likelihood. On a more practical

[6]For an integrable system, frequencies can be simply thought as another coordinate system in the action space.

note, given that the density of stars (or pairs of stars) is discrete, the answer does depend on the coarse-graining procedure. However, since there are many more stellar pairs than stars ($N(N-1)/2$ vs N), our likelihood computation is much more robust to coarse-graining. Furthermore, to our knowledge, none of these methods have yet been applied to real data.

Finally, to be fair, we should also highlight some of the caveats in our study. Several assumptions are made in the derivation of our likelihood test in Appendix A, most importantly that of a uniform background with statistically uniform gaussian fluctuations in the action space. How much do gravitationally bound structures or non-uniformity of the background can bias our finding? While the latter effect is partially captured by the dependence on $D_{\rm max}$, a more systematic test using numerical simulations of galaxy formation may be more satisfactory. Another point of concern is the dependence of the best-fit parameter on the selection cuts. While, this could signal the inadequacy of our current potential model (either for stellar disk or dark matter halo, which may need more free parameters), it could also signal deeper problems such as errors in computing the action variables for Milky Way potential, or their non-adiabatic evolution.

Chapter 3

Measuring the Mass and Light of the Filaments in Cosmic Web

In the previous chapter, we studied the structure of the dark matter halo around Milky Way. In this chapter, we turn to probing the properties of dark matter on larger scales. As discussed in Chapter 1, on large scales, the galaxy distribution follows a web-like pattern, which is known as the Cosmic Web. Massive halos, where galaxy groups and clusters are formed, are expected to be connected by dark-matter dominated bridges, which are the so-called filaments. In this chapter, we investigate how "dark" the filament is compared to other cosmic structures. To evaluate this, we compute the mass-to-light ratio, M/L, of an ensemble of filaments. The determination of M/L ratios of filaments is of great importance because the halo mass function also depends on environment, being different in filamentary environments than in knots (Libeskind et al., 2018). Therefore, the mass-to-light ratio in a filament is likely to reflect this dependence and may be different from the global mean. Of course, it is also possible that the environment affects the mass-to-light ratio of haloes at fixed halo mass. The goal of this work is to measure the mass-to-light and stellar-to-total-mass ratios, as a first step towards understanding these effects.

We estimate the mass of filaments using weak gravitational lensing. We stack and analyze the weak lensing signal between Luminous Red Galaxy (LRG pairs). To isolate a clean filament signal, we construct a catalogue of non-physical (projected) pairs and subtract the stacked weak lensing signal from that of the stacked LRG pairs. We then calculate the average luminosity of filaments between LRG pairs using a similar approach. To compute the total luminosity of the stacked filament, we fit a Schechter function (Schechter, 1976) to the observed excess galaxy numbers in the stacked filament. In this way, the M/L ratios of filaments can be measured.

While there have now been both studies of the luminous content of filaments and of their dark matter content, to our knowledge, there have been no attempts to measure both simultaneously in a consistent way and interpret the results therefrom. Our work presents the measurements of M/L ratios of filaments in the first time, showing that these structures, although are dominated by dark matter, are not entirely dark.

This chapter is structured as follows: in Section 3.1, we discuss the data selection of LRG pairs and source galaxies, where LRGs are selected from BOSS and ellipticities of source background galaxies are taken from CFHTLenS. In Section 3.2, using weak lensing data, we reproduce the 2D projected mass maps for stacked filaments following the methodology presented in Epps & Hudson (2017). In Section 3.3, adopting the same LRG pair selection, we then stack the galaxy light and calculate the average luminosity of filaments. We present the resulting 2D projected luminosity density map and measured luminosity of stacked filaments in Section 3.3.3. We compare our M/L measurements for filaments with other literature in Section 3.3.5.

3.1 Observational Data

In this section, we include a brief summary of the surveys used in this study: the Sloan Digital Sky Survey (SDSS), the Baryon Oscillation Spectroscopic Survey (SDSS-BOSS) and the Canada France Hawaii Telescope Lensing Survey (CFHTLenS). We also summarise the selection criteria for the LRG pair construction and the background (and also, foreground) galaxies of the stacked light and stellar-mass maps.

3.1.1 CFHTLens

In the regime of weak gravitational lensing, the ellipticity of a source background galaxy will be distorted by foreground massive objects at a level of few percent. Therefore, it is extremely challenging to remove unnecessary contamination from the point-spread function in the observational data and produce accurate galaxy shape measurements.

The CFHTLenS is designed to have shape and photometric redshift measurements down to $i'_{AB} \sim 24.7$. Derived from the Wide component of the Canada France Hawaii Telescope Legacy Survey (CFHTLS-wide), it covers 154 deg^2 under five different optical bands ($u^*g'r'i'z'$) with sub-arcsecond seeing. The image data are taken with MegaCam ($\sim 1^\circ \times 1^\circ$ field of view) mounted on the CFHT. Observations are conducted in four different patches: W1 with 72 pointings centered around RA = 34.5°, DEC = -7.0°; W2

with 33 pointings centered around RA = 133.5°, DEC = -4.3° ; W3 with 49 pointings centered around RA = 214.5°, DEC = 54.5° and W4 with 25 pointings centered around RA = 333.3°, DEC = 1.3°. Galaxies with available shear and redshift measurements have a number density level of approximately 17 galaxies per square arcminute. A detailed description related to the data reduction is presented in Erben et al. (2013).

The shape measurements of galaxies are estimated using the *lensfit* algorithm developed by Miller et al. (2013). The algorithm performs a Bayesian fit to the image data. A galaxy model with seven free parameters is convolved with the point spread function (PSF), which is then fit to the observation. After marginalizing over the galaxy size, flux, position and the bulge fraction, the ellipticity (e_1, e_2) and a weight is assigned to each galaxy based on the likelihood of the model posterior probability. A detailed description of the method is presented in Miller et al. (2013), along with the discussion of the required calibrations as estimated from the image simulations (Heymans et al., 2012; Miller et al., 2013).

Photometric redshift estimate is carried out using the Bayesian Photometric Redshift (BPZ) code (Benítez, 2000) with a slightly modified prior. The mean redshift for the sample is $\sim \bar{z} = 0.75$ when limiting the sample to the range $0.2 < z_{\rm photo} < 1.3$. Galaxies within this redshift range are considered to have accurate photo-z measurements with an average scatter of $\sigma_z \sim 0.04(1+z)$ and an average outlier rate below 4% (Heymans et al., 2012). A detailed description for the determination of photo-z is outlined in Hildebrandt et al. (2012).

3.1.2 SDSS-BOSS

As the weak lensing signal from an individual filament is relatively weak, we need to stack many LRG pairs to enhance the signal and suppress background noise. Furthermore, although the photometric redshift measurements are reliable enough for lensing studies, they are still not sufficiently accurate to locate LRG pairs in redshift space. In this regard, the BOSS survey is useful as it has unprecedented accuracy for the spectroscopic redshift measurements of luminous red galaxies (LRG) with $\sigma_{\rm spec} \sim 10^{-4}$ (Bolton et al., 2012). BOSS LRG targets are selected based on the SDSS multi-band ($ugriz$) imaging data, and are used for measuring the BAO scale using 1.5 million luminous galaxies to redshift $z < 0.7$. The coverage of the whole survey is approximately 10,000 degree^2 and the apparent magnitude limit of the spectroscopic sample reaches $i = 19.9$. LRGs are selected based on their color-magnitude cuts as outlined in Dawson et al. (2013) and the samples are further divided into LOWZ ($0.15< z <0.43$) and CMASS ($0.43< z <0.70$) catalogue based on their redshift distribution. A full descriptions of the survey can be found in Eisenstein

et al. (2011) and Dawson et al. (2013), and a detailed description of the reduction of spectroscopic redshift measurements are presented in Bolton et al. (2012).

LRG galaxies are early type galaxies with intrinsically red colors. They are believed to have old stellar populations with a prominent 4000 Å break in their spectral energy distribution (SED). This makes them relatively easy to select for surveys such as BOSS by using multi-color measurements, and this also makes them have relatively more accurate redshift measurements. More importantly, LRG galaxies are bright and massive so the majority of then are located in the center of their host halos (Mandelbaum et al., 2006; Parejko et al., 2013), which makes them good proxy for the galaxy rich groups/clusters. Therefore, they are regarded as good candidates for galaxy-galaxy lensing studies.

Fortunately, the CFHTLens survey overlaps with the latest version of the BOSS data (see Miyatake et al., 2015a). Using these LRG pairs and the shape measurements from CFHTLens within the same region, we can therefore conduct the lensing analysis and estimate the mass of filaments between LRG pairs, which will be discussed in detail later. For the light map, we select LRG pairs from almost the entire BOSS coverage to enhance the detected signal. However, as the BOSS spectroscopic sample is too shallow compared with the SDSS photometric sample ($i = 19.99$ vs $r \sim 22$), fainter galaxies can be missing if we were only using the BOSS spectroscopic catalogue. Also, due to fibre collision, which is due to the fact that the minimum allowed angular separation between two spectroscopic fibres on the focal plane is limited to 62 arcseconds (Dawson et al., 2013), the number of galaxies might be underestimated. Given all of those problems, we choose to use the SDSS photometric catalogue for the generation of stacked light map, which is a deeper survey and free from issues such as fiber collision.

3.1.3 Data selection

Filaments are the bridges between galaxy groups and clusters, but only pairs that are relatively close to each other are expected be connected by filaments. As LRGs are considered to be good proxies for rich groups (Mandelbaum et al., 2006), we first focus on the selection of LRGs in Section 3.1.3.1. Pairs of LRGs that are close in redshift are referred to as *physical* LRG pairs, and these are used to identify filaments. Furthermore, to isolate the filament signal, we also need a control sample of pairs that are not physically connected (close in redshift), which are referred to as *non-physical* (projected) pairs. The criteria for the pair selection are outlined in Section 3.1.3.2.

3.1.3.1 LRG selection

To study the filaments between physical LRG pairs, accurate LRG redshifts are necessary. Here, we use the spectroscopic redshifts provided by SDSS-III/BOSS to select LRGs. BOSS is a spectroscopic survey covering over 10,000 square degrees derived from the SDSS imaging, including the spectra and redshift measurements for 1.5 million galaxies with redshift $z < 0.7$. A full description and summary of the survey are provided in Eisenstein et al. (2011) and Dawson et al. (2013).

The weak lensing signal from the filament connecting a single LRG pair is weak and very noisy, so we stack LRG pairs together to enhance the significance of detection. Therefore, we extracted all object redshifts from 0.15 to 0.7 that are flagged with SourceType = "LRG" and Z_NOQSO>0. In order to investigate how the properties of filaments evolve with redshift, the whole LRG sample is further divided into two independent redshift bins: LOWZ and CMASS. In BOSS, LOWZ (low redshift) aims to selecting galaxies at lower redshift with coverage $0.15 < z < 0.43$ and CMASS (constant mass) is designed to select galaxies within $0.43 < z < 0.7$. In SDSS DR14, LRGs which satisfy with the LOWZ or CMASS criteria are flagged with "GAL_LOZ" or "GAL_CMASS_ALL" respectively. From now on, we shall have three samples available for further analysis: the whole BOSS sample including all LRGs from $z = 0.15$ to 0.7, and the LOWZ and CMASS subsamples.

3.1.3.2 LRG physical and non-physical pairs

For consistency with Clampitt et al. (2014) and Epps & Hudson (2017), physical LRG pairs are constructed by selecting pairs that have redshift separation $\Delta z_{\rm sep} < 0.002$ and projected 2D R separation within $6h^{-1}$Mpc $\leqslant R_{\rm sep} \leqslant 10h^{-1}$Mpc. They are physically "real" pairs as they are not only close in projection, but are also close in redshift space. In other words, two LRG haloes with such a small redshift separation are expected to be connected by filament.

To isolate the mass associated with the filament (as opposed to mass that may be clustered with LRGs), we also define a catalogue of non-physical LRG pairs. Non-physical LRG pairs, which are close in projection only, are selected by the same $R_{\rm sep}$ projection criterion as for physical LRG pairs but with larger redshift separation: $0.033 \leqslant z_{\rm sep} \leqslant 0.05$. Galaxy groups/clusters which have this amount of separation in the redshift space are unlikely to have filaments between them, so the signal from non-physical pairs only yields the contribution from two isolated LRG haloes. Therefore, the residual signal from the filament should remain after the subtraction between the stacked maps of physical and non-physical pairs.

3.2 Lensing Mass Map

Following the procedures outlined in Epps & Hudson (2017), in this section, we reproduce the 2D stacked surface mass density map of physical and non-physical LRG pairs using weak gravitational lensing. A major difference between our study and Epps & Hudson (2017) is that we also repeat the same analysis with BOSS LOWZ and CMASS subsamples separately. We begin with the LRG pair and lensing source galaxies selection in Section 3.2.1 and 3.2.2. We present the procedures for the generation of lensing map in Section 3.2.3. Then we summarise and discuss the lensing results in Section 3.2.4.

3.2.1 Lensing pair selection

The physical and non-physical pairs are selected based on the criteria outlined in 3.1.3.2. However, some extra position cuts need to be included. The lensing map is derived from ellipticity measurements from CFHTLens survey, so only LRGs in the regions of sky covered by CFHTLens survey are selected. This yields 15,254 LRG physical pairs in total, with a mean redshift $< z_{\rm pair} > \sim 0.47$, a mean 2D projected distance $< R_{\rm sep} > \sim 8.10 h^{-1}$ Mpc and a mean critical density $\overline{\Sigma}_{\rm crit} = 3471 M_{\odot}/{\rm pc}^2$. For the LOWZ sample, there are 2,752 LRG physical pairs with $< z_{\rm pair} > \sim 0.30$, $< R_{\rm sep} > \sim 8.10\ h^{-1}$ Mpc and $\overline{\Sigma}_{\rm crit} = 3320 M_{\odot}/{\rm pc}^2$, and for the CMASS sample, there are 12,497 LRG physical pairs with $< z_{\rm pair} > \sim 0.52$, $< R_{\rm sep} > \sim 8.04\ h^{-1}$ Mpc and $\overline{\Sigma}_{\rm crit} = 3565 M_{\odot}/{\rm pc}^2$.

3.2.2 Source galaxies for lensing map

The weak lensing source galaxies are selected from the CFHTLenS catalogues. CFHTLenS is derived from the Wide component of the Canada-France-Hawaii Telescope Legacy Survey (CFHTLS) and covers 154 square degrees including the photometric information from five optical bands (u_*, g', r', i', z') in four patches: $W_1 - W_4$, where $W1, W3$ and $W4$ have substantial overlap with BOSS over an area of approximately 100 deg^2 (Miyatake et al., 2015b). The galaxy shape measurements are from the *lensfit* algorithm and the photometric redshifts are obtained using the Bayesian Photometric Redshift code with a scatter of $\sigma_z \sim 0.04(1+z)$. General information about the data reduction and the survey design can be found in Heymans et al. (2012) and Erben et al. (2013), and a detailed description of the determination of the ellipticity and redshift measurements are presented in Miller et al. (2013) and Hildebrandt et al. (2012) respectively.

Based on the criteria outlined in Hudson et al. (2015), we select all source galaxies from unmasked regions with reliable photometric redshift measurements ($0.2 < z_{\rm p} < 1.3$) and with FITCLASS=0. Only galaxies with weight > 0 are used. This selection yields approximately 5.6×10^6 background sources. We use source galaxies that have photometric redshifts greater than 0.1 of that of the LRG pair, in order to avoid contamination by intrinsically-aligned "source" galaxies within the filaments themselves.

3.2.3 Methodology

As is the case with galaxy-galaxy lensing, it is necessary to stack LRG pairs to extract a statistically significant signal. Lensing by LRG pairs, however, is more complicated than lensing by single galaxies because the pairs are not circularly symmetric. Furthermore, LRG pairs have random orientation angles on the sky and each galaxy pair has its own projected separation. Therefore we need a coordinate transformation to normalise the position of each LRG pair in order to compare and stack the lensing signals. To achieve this, we follow the procedure outlined in Epps & Hudson (2017). First, all LRG pairs are rotated with respect to their centres such that the two LRGs lie along the x-axis in the new coordinate system. Then, pairs are rescaled by their 2D projected separation and the coordinate of individual LRG is transformed to $(x_{\rm L}, y_{\rm L}) = (-0.5, 0)$ and $(x_{\rm R}, y_{\rm R}) = (0.5, 0)$, where L and R denote left and right. Finally, all LRG pairs are stacked together. The coordinate transformation of the source galaxies follows the same procedure. Following this procedure, the shear map can be generated by stacking the ellipticities of the background source galaxies.

The resulting shear map, however, is difficult to interpret. Therefore, we convert the shear map into a convergence map, which is directly proportional to the surface mass density of the field, using the method of Kaiser & Squires (1993). With the convergence map, we can subsequently evaluate the residual surface mass density of the central region after subtracting the physical and non-physical maps.

3.2.4 Lensing results

The resulting shear map using the whole LRG sample is presented in Figure 3.1, and it is apparent that near the galaxy halos, the shear signals are distributed tangentially. The projected surface mass density maps are shown in Figure 3.2, which demonstrates the map reconstructed by using physical pairs connected by a central filament. Figure 3.3 shows the projected surface mass density around non-physical pairs. In contrast with

the physical pairs, there is no obvious filamentary structures between them. The excess projected surface mass density after subtraction is shown in Figure 3.4 with a rectangular box showing the region used for the mass measurement of the filaments. The proper size of the filament box is $\sim 7.0\ h^{-1}$Mpc $\times$ $2.4\ h^{-1}$Mpc, following Epps & Hudson (2017). The size of the box is the projected $R_{\rm sep}$ after averaging all LRG pairs, and it has excluded a region extending to 3 r_{200} around the LRG haloes.

The uncertainty in the measured filament mass is mainly due to "shape noise". We quantify this by first defining the variance of the shear in a given pixel at location (x, y) in the new coordinate system, $\gamma(x, y)$, is as follows:

$$\sigma^2_\gamma(x,y) = \left[\left(\frac{1}{\Sigma_l \Sigma_{j\in(x,y)} W_{lj}}\right)^2 \Sigma_l \Sigma_{j\in(x,y)} (W_{lj})^2 \sigma^2_{e_{lj}}\right] / (1+K)^2, \tag{3.1}$$

where $(1 + K)$ is an additional multiplicative correction for the shear (Miller et al., 2013). The sum is taken over all source galaxies j that belong to a pixel, then sum over all LRG pairs, l. The scatter for a single lens-source pair lj is $\sigma^2_{e_{lj}} = \frac{\sigma^2_{\rm int}}{(w/w_{\rm max})}$, where w is the weight assigned to each galaxy from *lensfit* algorithm. The intrinsic scatter, $\sigma_{\rm int}$, and the maximum weight, $w_{\rm max}$, are two constants taken as 0.28 and 16 respectively in CFHTLens survey. $\Sigma_l \Sigma_{j\in(x,y)} W_{lj}$ is equivalent to

$$\Sigma_l \Sigma_{j\in(x,y)} W_{lj} = \Sigma_l \Sigma_{j\in(x,y)} w_j \Sigma^{-2}_{{\rm crit};l,j} \tag{3.2}$$

as defined in Hudson et al. (2015) while calculating the shear. Then, this artificial scatter representing the shape noise is drawn from a Gaussian distribution with $N(0, \sigma^2_\gamma(x, y))$ and then added to the calculated shear value per pixel. This noise is propagated to the subsequent convergence calculations and finally to the subtraction between the two maps. After repeating the above procedure 1000 times, the average value and dispersion of the mass distribution measured from the enclosed box are taken as the final result.

The enclosed mass of the filament in the rectangular box is $(4.25 \pm 0.91) \times 10^{13} M_\odot$, which is a detection at the $\sim 4.5\sigma$ significance level. Another empirical way of estimating the variance of the excess mass is to place boxes of the same dimensions on the resulting excess convergence map. Specifically, boxes are chosen with identical size of the filament region (0.86×0.30 in the standardized coordinates), and they are placed in the region between $\pm$ 4 along the x axis and $\pm$ 4 along the y axis. A square centred at (0,0) with an area of 4 are excluded in order to avoid the filament region. There are 210 such boxes, and the standard deviation of the excess mass is 1.30, which is larger than but still comparable to the uncertainty estimated from the shape noise. A histogram of the masses in these

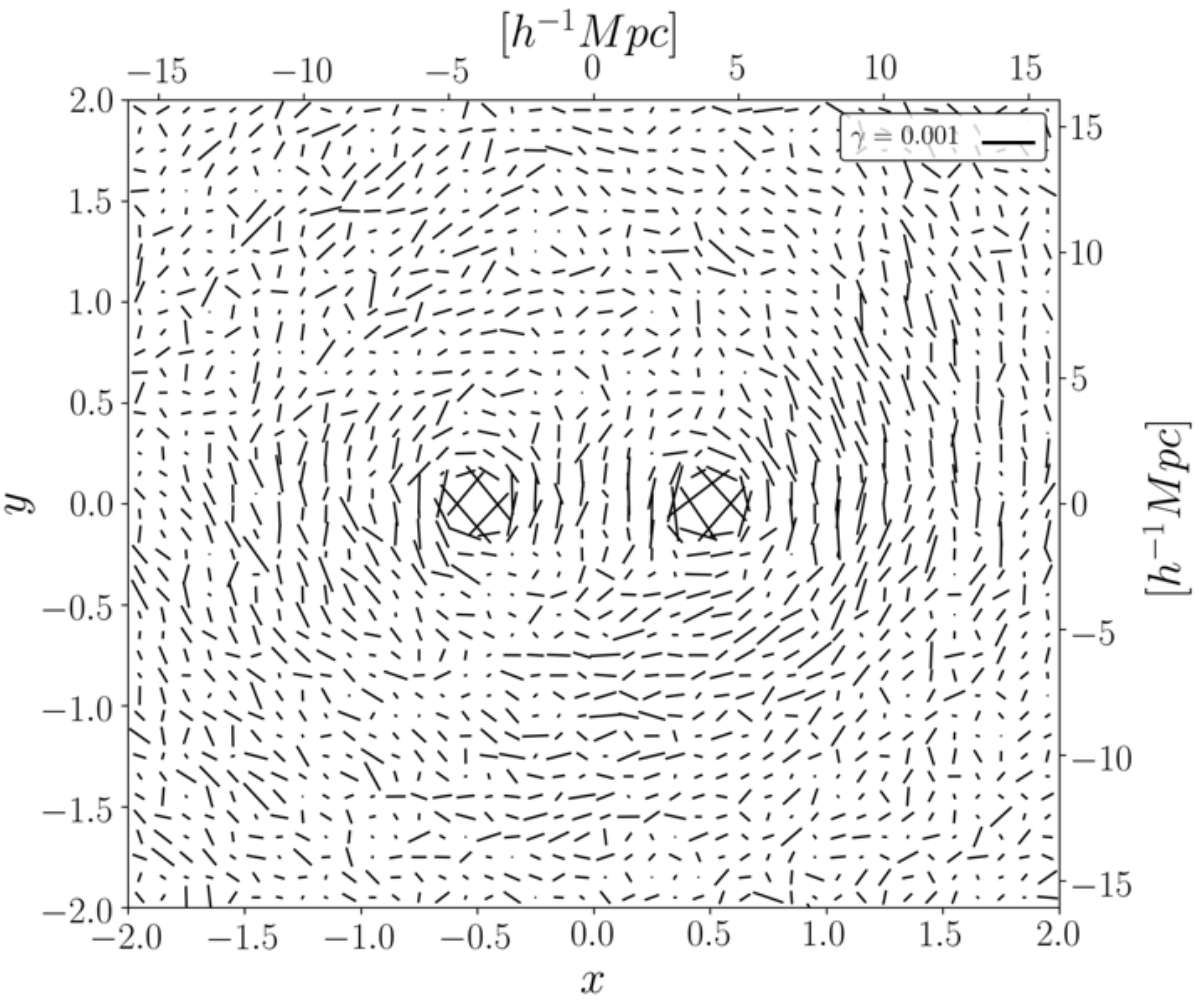

Figure 3.1: The resulting shear map after stacking approximately 17,000 LRG physical pairs from redshift $0.15 < z < 0.7$ with 2D R seperation $6h^{-1}\text{Mpc} \leqslant R_{\text{sep}} \leqslant 10h^{-1}\text{Mpc}$. The tangential shear signal around two LRG halos are obvious.

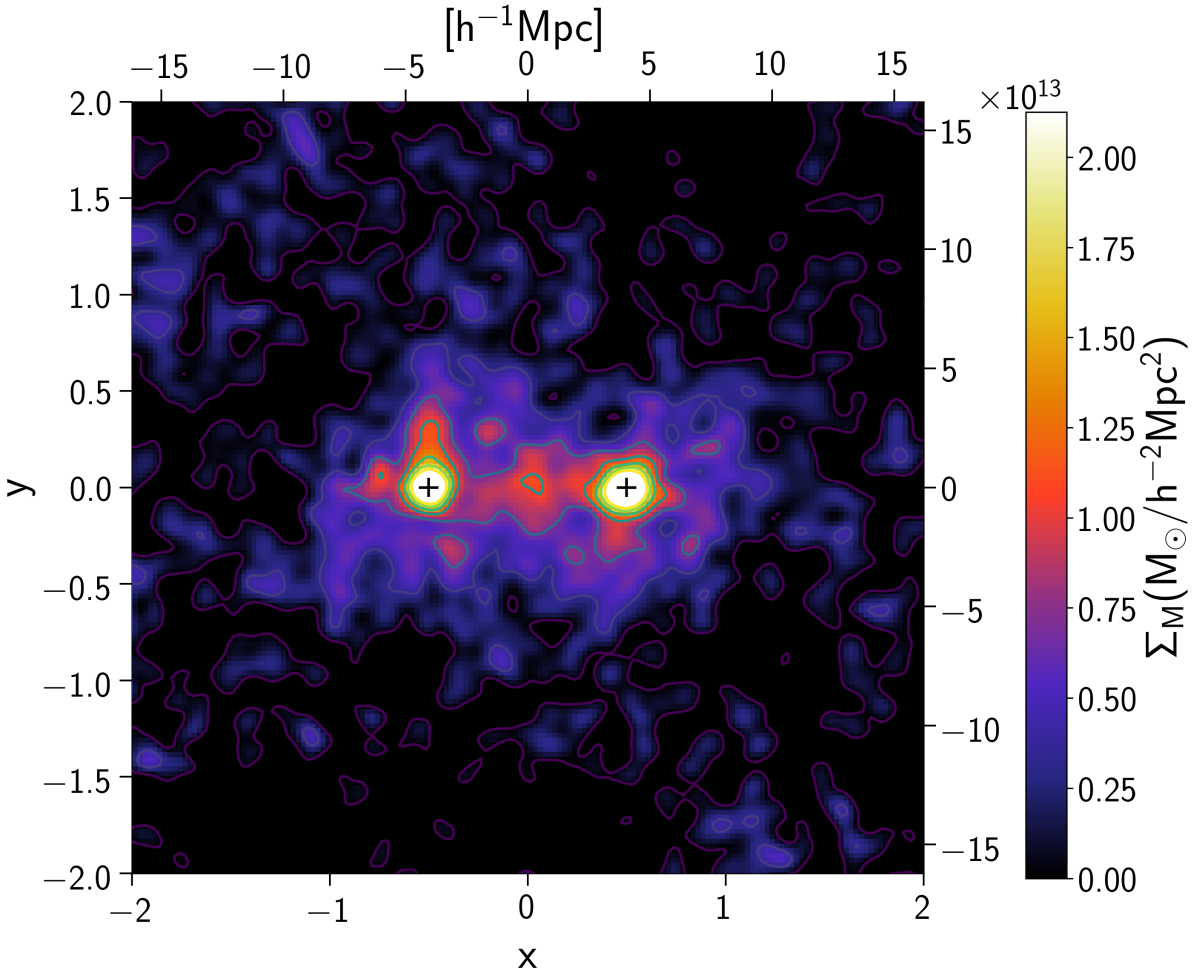

Figure 3.2: The surface mass density map (Σ_M) obtained from the shear map using the Kaiser & Squires (1993) algorithm. This is the map constructed using physical LRG pairs, where the central bridge-like filamentary structure is visible. The locations of LRG haloes in the standardized coordinate system are marked as two "+" signs. A Gaussian filter of smoothing scale 0.40 h^{-1}Mpc (0.05 in units of x, y) has been applied to the image for illustration. Note that the left and bottom coordinates are in rescaled units, while the top and right are in h^{-1} Mpc. The colorbar shows the surface mass density in units of solar mass per h^{-2}Mpc^2.

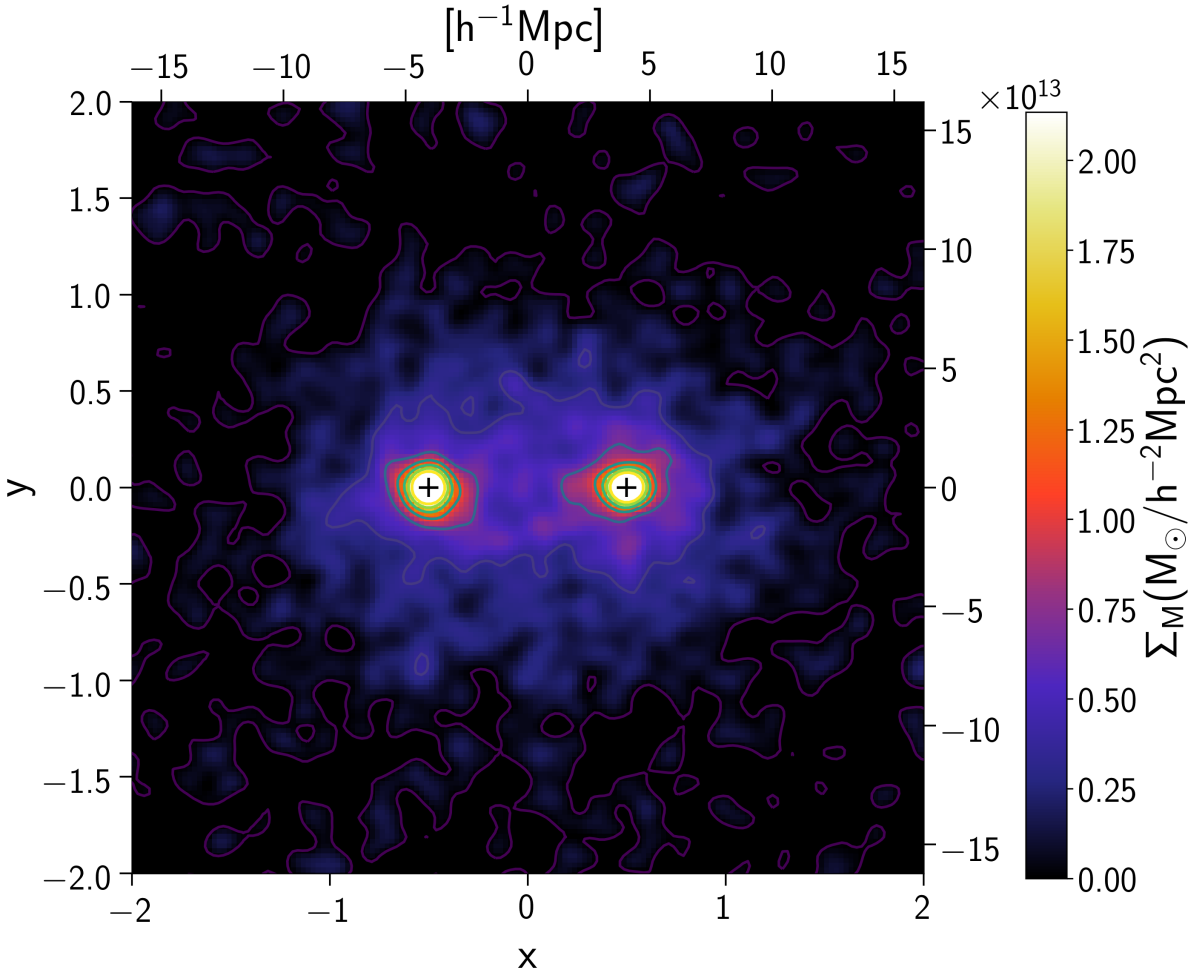

Figure 3.3: As in Figure 3.2, but for Σ_M map from non-physical LRG pairs.

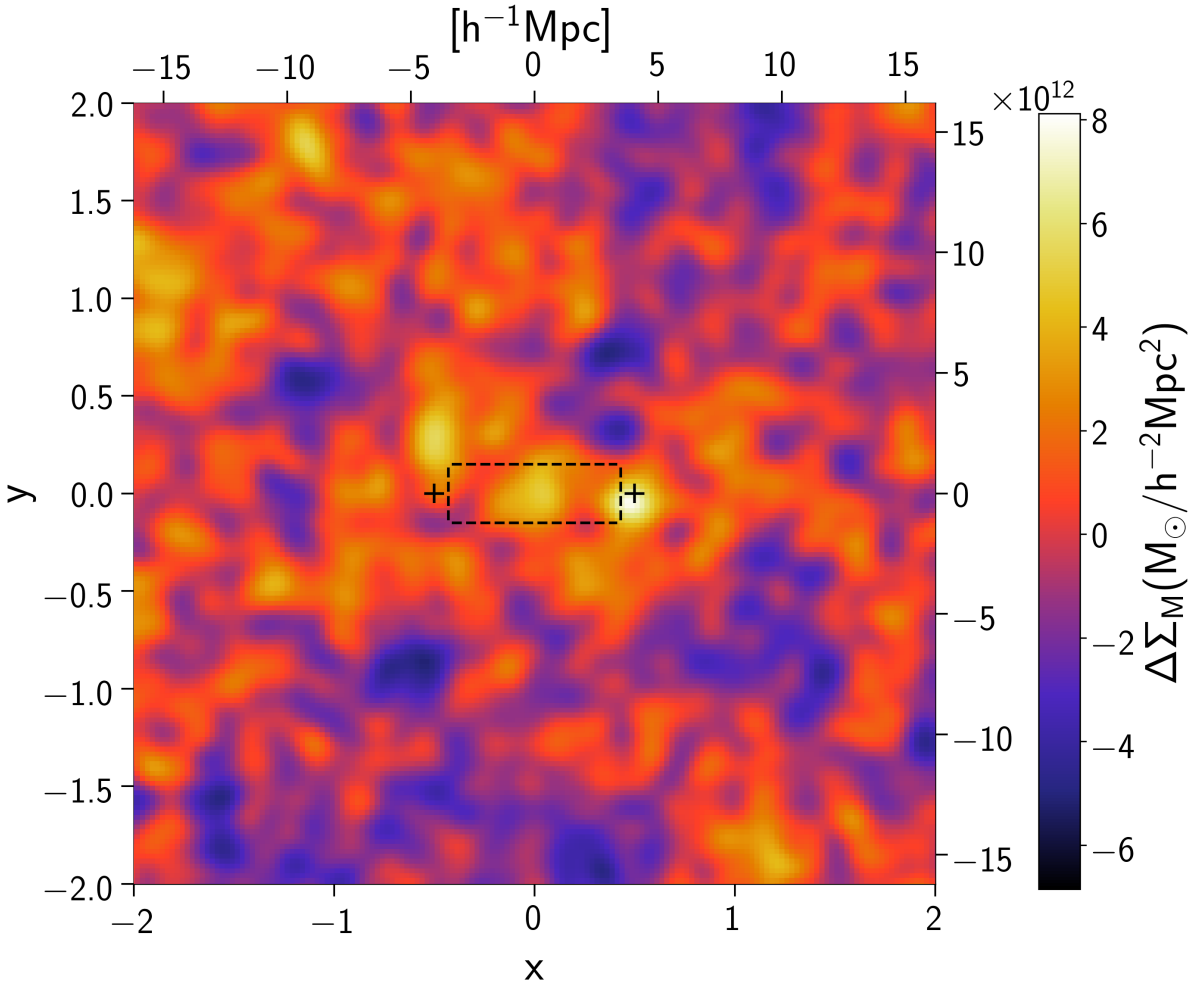

Figure 3.4: The resulting excess surface density map after the subtraction the Σ_M map of the non-physical pairs from that of the physical pairs. The region which is used for the mass measurement of the filament is shown by the black dashed rectangular box. The size of the filament box is 0.86 × 0.30 $R_{\rm sep}$ (Epps & Hudson, 2017) in the new coordinate system, which corresponds to $\sim$ 7.0 h^{-1}Mpc × 2.4 h^{-1}Mpc in proper size.

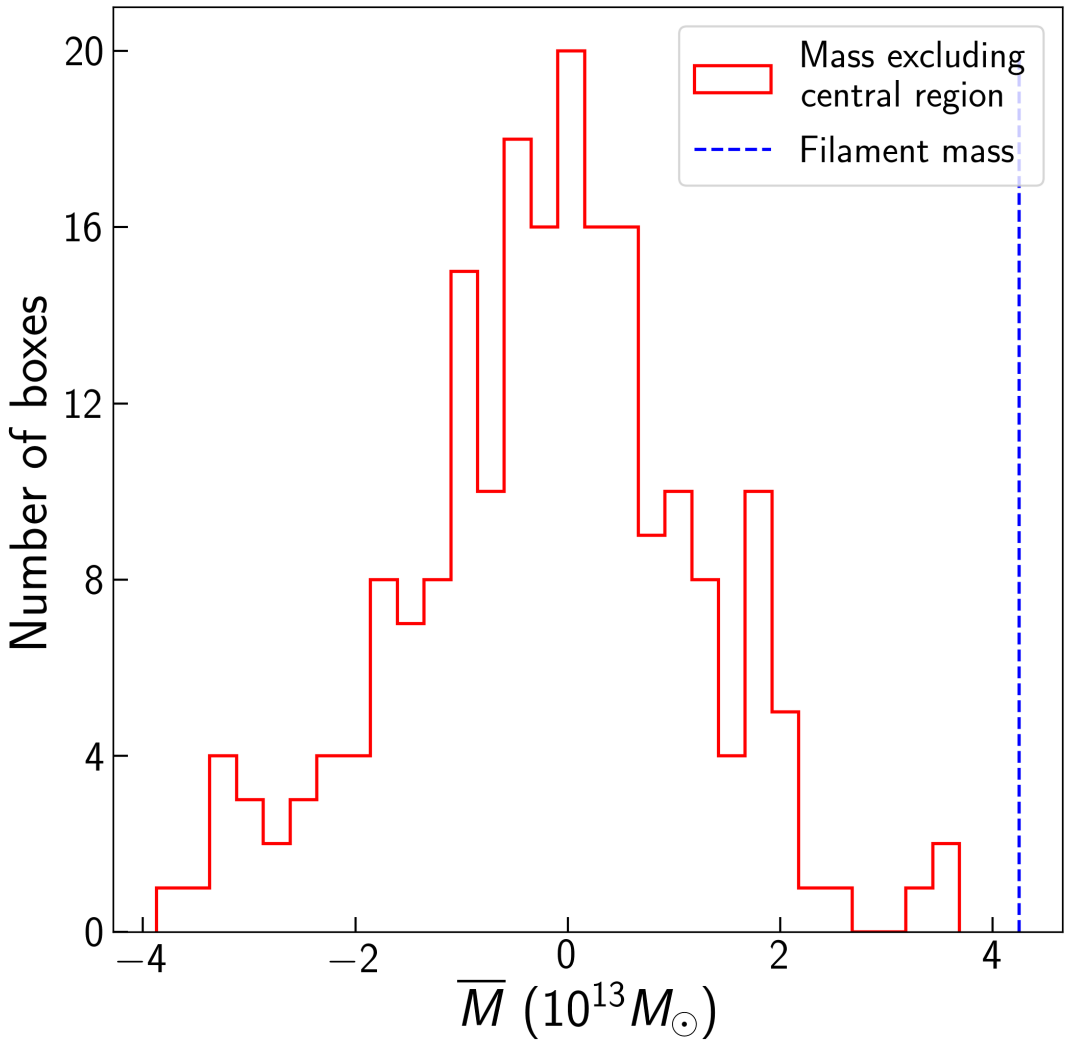

Figure 3.5: Histogram of the averaged masses within the same sized rectangular box. Specifically, boxes are placed on the excess convergence map over a region between ± 4 along the x axis and ± 4 along the y axis. A square centred at (0,0) with an area of 4 is excluded in order to avoid the central filament region.

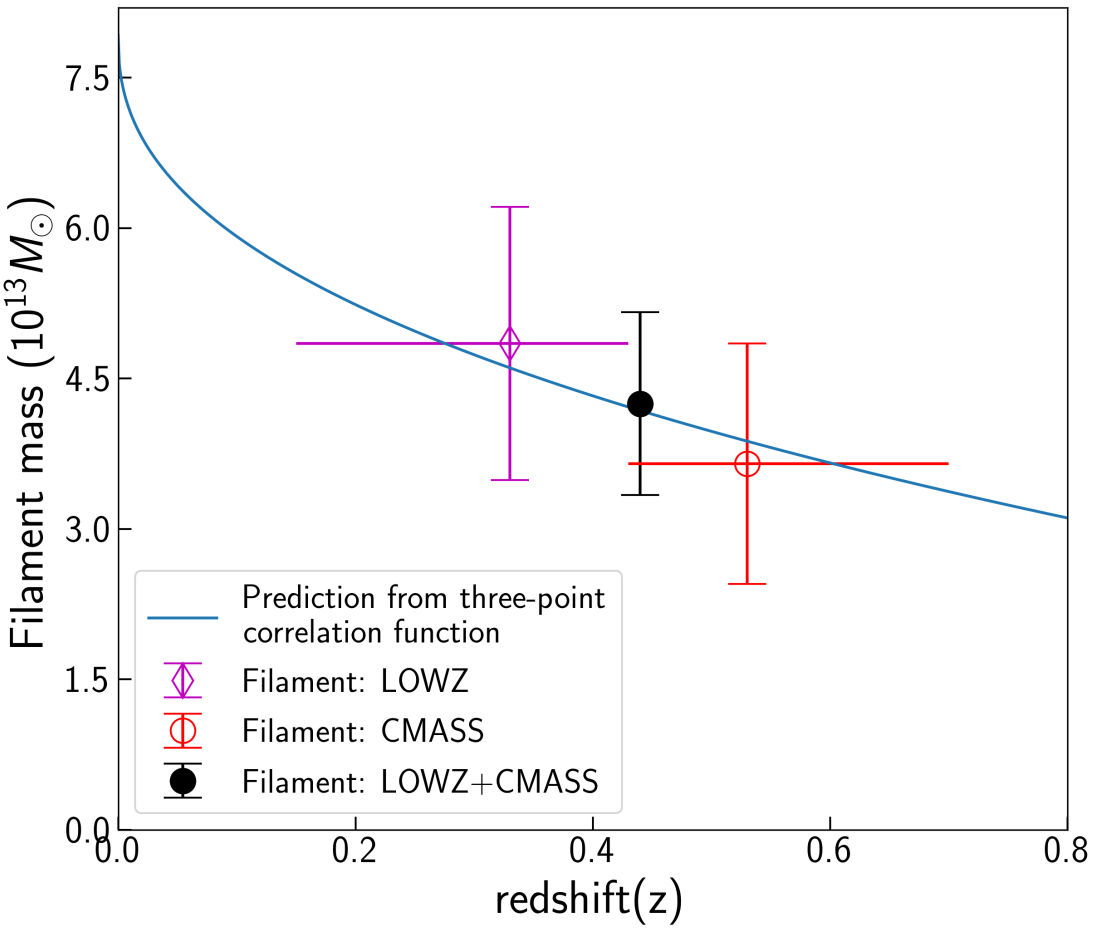

Figure 3.6: Filament mass as a function of redshift. The thin magenta diamond and red hollow circle with errorbars are the measurements from two independent samples LOWZ and CMASS. The solid black circle the average of the LOWZ and CMASS measurements. The blue line shows the scaling of the filament mass expected from three-point correlation function for a cosmic filament with 2D $R_{\rm sep}$= $6h^{-1}$–$10h^{-1}$ Mpc between two $10^{13}M_\odot$ haloes (see text for details).

boxes is shown in Figure 3.5, where the blue dashed line shows the measured mass of the filament within the rectangular box in the centre.

To measure the evolution of the filament mass with redshift, we perform the same mass measurement and uncertainty calculation for the LOWZ ($0.15 < z_{\rm LRG} < 0.43$) and CMASS ($0.43 < z_{\rm LRG} < 0.70$) subsamples separately. For the same filament box, we find $(4.85 \pm 1.36) \times 10^{13} M_{\odot}$ for the LOWZ sample and $(3.65 \pm 1.20) \times 10^{13} M_{\odot}$ for the CMASS sample. At face value, the observed mass of the filament appears to decrease with the increasing of redshift, although the difference is not statistically significant. Theoretically, the scaling of the filament mass with redshift can be predicted from evolution of the three-point correlation function which depends on the bispectrum whose evolution should scale like square of the linear growth factor (Clampitt et al., 2014; Epps & Hudson, 2017; Wang et al., 2019). The redshift scaling is shown in Figure 3.6, after fixing the normalization to agree with the mean filament mass at the mean filament redshift. Although the uncertainties are large, the data are consistent with the predicted redshift scaling.

Recently, Xia et al. (2019) claimed a 3.3 σ level detection of an anisotropic shear signal from the filament using only LOWZ LRG pairs, but combining shapes from KiDs+VIKING-450 Survey (KV450), the Red Cluster Sequence Lensing Survey (RCSLenS) and the CFHT-LenS. In their study, the selection criteria for filament candidates are slightly different from ours. They select all pairs with 2D projected separation $3h^{-1}{\rm Mpc} \leqslant R_{\rm sep} \leqslant 5h^{-1}{\rm Mpc}$ (instead of $6h^{-1}{\rm Mpc} \leqslant R_{\rm sep} \leqslant 10h^{-1}{\rm Mpc}$ as adopted by this study). Using the same method as presented above, we re-run the analysis to measure the mass of the filament from our data using their LRG separation criteria. We find $M_{3-5} = 4.35 \pm 1.35 \times 10^{13} M_{\odot}$. The mass measured from their study is $M_{\rm Xia+} = 6.7 \pm 3.1 \times 10^{13} M_{\odot}$, but this mass is obtained by fitting a power law density profile to the filaments. Colberg et al. (2005) studied filaments in N-body simulations and found that, in cylindrical coordinates aligned with the filament axis, the density profile has a core and then falls like r^{-2} at large radii. However, Mead et al. (2010) modelled the *projected* convergence map of the filaments with the same functional form:

$$\kappa(r) = \frac{\kappa_0}{1 + \left(\frac{r}{r_c}\right)^2}, \tag{3.3}$$

and Xia et al. (2019) used this model in their work. Their filament mass is defined by integrating the profile in the direction perpendicular to the filament from 0 to infinity, and along the filament over a range of length $R_{\rm sep}$ yielding a mass of $\pi r_c R_{\rm sep} \mathcal{F}_c$. In our analysis, the integration is conducted in a projected rectangle, where the height of the rectangle is $0.30 R_{\rm sep}$ and the length is $0.86 R_{\rm sep}$. We use their best-fit parameters of the density profile (presented in column "All" in their Table 3) and calculate the expected mass ratio between

their measurement and ours as

$$\frac{M_{3-5}}{M_{\rm Xia+}} = \frac{2 \times 0.86 \int_0^{0.15 R_{\rm sep}} \frac{\mathcal{F}_c}{1+(\frac{r}{r_c})^2} dr}{2 \times \int_0^{\infty} \frac{\mathcal{F}_c}{1+(\frac{r}{r_c})^2} dr} = 0.54, \tag{3.4}$$

where the value of $R_{\rm sep}$ is adopted as $4h^{-1}$Mpc. Thus given the differences in definition, we expect our mass to be 54% of theirs and we measure a ratio of 0.65, and so we conclude these are consistent.

3.3 Maps of Light and Stellar Mass in the Filament

In this section, we turn to the luminosity and stellar mass of the filament, which is the main focus of this work. Our method for measuring the light and stellar mass maps are similar to the construction of lensing map: we will stack galaxy luminosity for physical and non-physical pairs and subtract them to isolate the excess light associated with the filament. However, there are some subtle differences between these two situations. In Section 3.3.1 and 3.3.2, we discuss the LRG pair selection and galaxies used in the construction of the light map. In Section 3.3.3, we demonstrate our 2D projected luminosity and stellar mass density map. In Section 3.3.4, we present our method of measuring the total light content of filaments and some discussions on the results.

3.3.1 LRG pair selection for the light map

In addition to the $z_{\rm sep}$ and 2D $R_{\rm sep}$ selection cuts mentioned in Section 3.1.3.2, to enhance the actual filament signal and suppress the background noise, for the light and stellar mass maps, we select all the LRGs within region $120^\circ \leqslant$ RA $\leqslant 230^\circ$, $10^\circ \leqslant$ Dec $\leqslant 60^\circ$, rather than restricting ourselves only to the CFHTLenS area. From the whole BOSS redshift coverage, there are 448,314 LRG pairs with $< z_{\rm pair} > = 0.44$ and $< R_{\rm sep} > = 8.04\ h^{-1}$ Mpc. For the LOWZ sample, we find 50,917 pairs with $< z_{\rm pair} > = 0.33$ and $< R_{\rm sep} > = 8.03\ h^{-1}$ Mpc. The CMASS sample has 390,748 pairs with $< z_{\rm pair} > = 0.53$ and $< R_{\rm sep} > = 8.04\ h^{-1}$ Mpc.

The lensing map was constructed by taking an ensemble average of the galaxy shapes over all LRG pairs. Notice that the number of galaxies entering into the average does not bias the result, although it does affect the noise. Consequently, the weak lensing mass map is not affected by incompleteness in the source catalogue. In contrast, for the light

map, it is necessary to count precisely the number of galaxies per LRG pair located in each area of the map. Furthermore this counting is done in rescaled-coordinates, and a given unit area will subtend a larger area on the sky for an LRG-pair at lower redshift or if $R_{\rm sep}$ is larger. This means that in order that the physical and non-physical light maps are comparable, the redshift and 2D $R_{\rm sep}$ distribution of two LRG-pair catalogues need to be as close as possible. In general, the non-physical pairs are randomly distributed, so the count of the non-physical pairs should increase linearly as a function of R separation, but this is not the case for the physical LRG pairs (which are clustered). Therefore, there is a higher fraction of non-physical pairs at large R separation bins than those of physical pairs, which would lead to differences in the mean galaxy number density between the physical and non-physical projected maps. To correct this, we apply a one-to-one match between physical and non-physical pairs based on their redshift and 2D projected separation by finding all non-physical pairs, where the difference Δ between the physical and non-physical pairs, limited to $\Delta z < 0.01$ and $\Delta R_{\rm sep} < 0.2$, and only keeping the one with the smallest $\Delta R_{\rm sep}$. In this way, we generate a catalogue of non-physical pairs which has approximately the same redshift and $R_{\rm sep}$ distributions as physical pairs, ensuring that the mean foreground/background level projected in the same filament box is comparable.

3.3.2 Photometric catalogue for light map

To produce a stacked light map, we use galaxies from the SDSS catalogue. Specifically, the photometric catalogue was constructed by cross-matching GalaxyTag and PhotoZ in SDSS DR14 (Beck et al., 2016). Galaxies that are marked with photometry flag CLEAN = 1, bestFitTemplateID > 0 and $z_{\rm photo} > 0$ are selected. The first criterion ensures the reliability of the photometric measurements, and the last two criteria are used to remove bad measurements. We include all galaxies with r-band cmodel magnitude brighter than 21.5. Note that, for LRG pairs located close to the boundaries on the sky of the photometric catalogue, photometric galaxies are only available on one side of the pair but lacking on the other side. To avoid this edge effect, the sky coverage of the source galaxies is chosen to be slightly larger (~ 8 degree) than the coverage of LRG spectroscopic catalogue. This yields a sample of 29,703,867 galaxies in the r band.

In order to investigate the colour of the galaxies in the filament, we also conduct the same analysis in the g-band, and the selection criteria for the photometric data are identical except for the choice of apparent magnitude limit which is set to 22.5[1]. These magnitude limits corresponds to approximately 50% completeness for each band.

[1]The choice of r and g band magnitude cut are more or less tentative. However, while computing the total luminosity, our final results are not sensitive to the choice of magnitude cut.

To compute the luminosity of a galaxy, we use the SDSS PHOTOZ table which contains $z_{\rm photo}$, spectral type, and K-corrections for all primary objects flagged as galaxies. The photometric redshifts are determined using the "empirical method" as discussed in Beck et al. (2016). The K-corrections are based on the best-fitting spectral template from Dobos et al. (2012) for each galaxy at its photometric redshift.

The luminosity is then

$$L_{^{0.1}n}/L_{\odot,^{0.1}n} = 10^{(M_{\odot,^{0.1}n} - M_{^{0.1}n})/2.5}\,, \tag{3.5}$$

where

$$M_{^{0.1}n} = m_n - {\rm DM}(z_{\rm photo}, \Omega_m = 0.3, \Omega_\Lambda = 0.7, h = 0.7) - K_{^{0.1}n}\,, \tag{3.6}$$

${\rm DM} = 5\log_{10}(D_{\rm lum}/10\,{\rm pc})$ and where n refers either to the $^{0.1}r$- or the $^{0.1}g$-band. We adopt these shifted bands to make the later comparison with SDSS literature results simpler.

Having calculated the luminosity, the stellar mass of a galaxy can be obtained from the stellar-mass-to-light ratio, $M_{\rm stellar}/L_{^{0.1}r}$. As mentioned previously, for the galaxies in the PHOTOZ Table, each galaxy is matched with a best-fitting template from which physical parameters, such as $M_{\rm stellar}/L_{^{0.1}r}$, are also estimated (Dobos et al., 2012). There are 53,453 galaxies (approximately 0.2% of the total sample) with calculated luminosity or stellar mass greater than $1\times10^{12}L_\odot$ or $1\times10^{12}M_\odot$ for both r and g band. We assume that most of these are due to catastrophic failures of the photometric redshifts and discard them from the samples.

3.3.3 Light and stellar mass map

We then stack the luminosities (or stellar masses) of SDSS galaxies around $\sim$ 450,000 LRG pairs, and subtract non-physical pairs from the physical pairs to produce the observed excess luminosity and stellar mass density per LRG pair shown in Figure 3.7 and 3.8. The left columns show the 2D projected map using the whole LRG sample (top), CMASS sample (middle) and LOWZ sample (bottom), where the rectangular box delineates the size of projected filament region. The values of $\Delta\Sigma_L$ and $\Delta\Sigma_{M_*}$ computed from the central filament box are $(0.68\pm0.04)\times10^{11}L_{\odot,^{0.1}r}$ and $(1.91\pm0.09)\times10^{11}M_\odot$ for LOWZ+CMASS catalogue, $(0.58\pm0.04)\times10^{11}L_{\odot,^{0.1}r}$ and $(1.49\pm0.09)\times10^{11}M_\odot$ for CMASS catalogue, and $(1.54\pm0.16)\times10^{11}L_{\odot,^{0.1}r}$ and $(4.00\pm0.41)\times10^{11}M_\odot$ for LOWZ catalogue, where the errorbars are evaluated from the covariance matrix. To further show the significance of our measurements, on the right panels of each plot, we plot the averaged excess luminosity and stellar mass density as a function of the x coordinate: for each x, we averaged

$\Delta\Sigma_L(\Delta\Sigma_{M_{\rm stellar}})$ along the y-axis. There is a highly significant detection of light and stellar mass in the filament region.

These maps show the "dumbbell" structure expected from the three-point correlation function, as calculated in Clampitt et al. (2014) and Epps & Hudson (2017): the excess light peaks at the locations of the LRG haloes, with a filament connecting them. Furthermore, the filament extends beyond the two LRG haloes, albeit at lower density. These maps are significantly less noisy than the WL mass maps due primarily to the much larger number of LRG pairs used.

3.3.4 Total luminosity and stellar mass in the filaments

The simple analysis discussed in the previous section, while useful for highlighting the spatial structure of the filament, has disadvantages for estimating the total light in the filament, due both to the flux limit of the SDSS r-band data, as well possible outliers in the photometric redshifts. We now discuss an alternative method that is less sensitive to these issues.

In order to measure the total light, within the 2D projected filament box, we fit a Schechter function to the excess galaxy counts (per LRG pair) as a function of apparent magnitude. This allows us to compute the total luminosity by integrating over the best-fitting Schechter function. Unlike the luminosity/stellar mass map, excess galaxy number counts do not depend on photometric redshifts. We follow the same selection criteria and corrections for the construction of physical and non-physical LRG pairs catalogue. The number of galaxies is then divided into 100 equal-sized apparent magnitude bins from 15 to 21.5 (15-22.5 for ${}^{0.1}g$ band), and subtraction gives us the observed excess number as a function of magnitude.. Then the best-fit parameterization of the Schechter function can be obtained by comparing the observation with forward-modelled predicted counts. In each apparent magnitude bin, the predicted galaxy number for a given LRG pair can be computed via a Schechter function (Schechter, 1976) by

$$N(m,z) = \phi_\star \int_{L_{\rm min}(z)}^{L_{\rm max}(z)} \left(\frac{L}{L_\star}\right)^\alpha e^{-L/L_\star} \frac{dL}{L_\star}\,, \tag{3.7}$$

where α is the slope of the luminosity function and $L_\star$ is the characteristic galaxy luminosity. The upper and lower luminosity limits in the integral can be computed with Equation (3.6) based on the distance modulus and K-corrections of the filament galaxies. Because the spectroscopic redshifts of the LRG pairs are known, we assume the redshifts of the galaxies in the filaments are identical to that of their host LRG pair. For the K-correction,

we make a histogram of the excess number of galaxies in the filamentary box as a function of K-correction, then the weighted average of the distribution is our final choice of K-correction. Finally, we average over the redshifts of the LRG pairs.

Note that the normalization factor defined here is different from the usual pre-factor in the Schechter function. To make the normalization factor more physical, we further define $n_\star$ as the ratio between the total integrated luminosity and $L_\star$:

$$n_\star = \frac{L_{\rm tot}}{L_\star}, \tag{3.8}$$

where $L_{\rm tot}$ is the total luminosity and index from the MCMC fits (as discussed below). $L_\star$ is calculated by using the best-fit $M_\star$ from MCMC.

Comparing the model to the observation, there are three different parameters to be determined: the normalization factor, $\phi_\star$, the slope, α, and the characteristic galaxy absolute magnitude, $M_\star$. To find the best-fitting set of parameters, we define the following χ^2:

$$\chi^2(\phi_\star, \alpha, M_\star) = [N_{\rm obs}(m) - N_{\rm mod}(m)]^T C^{-1} [N_{\rm obs}(m) - N_{\rm mod}(m)], \tag{3.9}$$

where $N_{\rm obs}(m)$ is the actual observation and $N_{\rm mod}(m)$ is the expected number per pair in each apparent magnitude bin integrated using Equation (3.7) and C is the sample covariance matrix. To compute the covariance matrix, for each LRG pair, we record the excess number of galaxies as a function of apparent magnitude in the central rectangular region. We then use the best-fitting parameters to compute the total light by integrating the Schechter function over luminosity from 0 to ∞. We expect to account for the missing light with this approach.

To obtain the best-fit parameterizations of the Schechter function, we use the Markov chain Monte Carlo (MCMC) method . More specifically, we use the online Python package EMCEE (Foreman-Mackey et al., 2013) with the likelihood function defined in Equation (3.9). For the fit, we assumed flat priors for all three parameters: $0.01 < \phi_\star < 5.0$, $-24.0 < M_\star < -19$, $-2.0 < \alpha < -0.01$ (when $\alpha \leq -2.0$, the luminosity integration diverges). The sampler was run with 30 walkers, each with 500 steps. We discard the first 150 burn-in steps, which leaves a MCMC chain with a shape of 10500 samples.

Figure 3.9 shows the observed excess number per LRG pair in the filament box as a function of apparent magnitude. Error bars in the observations are taken from the diagonal elements in the covariance matrix. The best-fit Schechter function and its uncertainty, calculated by using 50 randomly chosen MCMC samples, are also shown. The observed number at bright magnitudes fluctuates around zero, consistent with the behaviour of luminosity function. Error bars in the CMASS measurements are smaller than the error in

LOWZ sample because there are more pairs in the CMASS sample, although the LOWZ sample spans a greater range in magnitude. The best-fit parameters for the LOWZ and CMASS samples are summarised in Table 3.3[2]. The nominal values and uncertainties of the three fitted parameters are approximated by the average and standard deviation calculated from MCMC samples. Similarly, we compute the total luminosity and its uncertainties using the correlated parameters from the MCMC samples. The total stellar mass, $M_{\rm stellar,total}$, is computed by multiplying the total luminosity with the observed $M_{\rm stellar,observed}/L_{\rm observed}$ from the 2D map.

Notice that, for the CMASS sample, the degeneracy of the three parameters is significant. This is due to the flux limit of the SDSS imaging and the depth of the CMASS LRGS, by which only the bright part of Schechter function can be observed. This degeneracy could be broken with a deeper dataset, or more simply, using low-redshift sample (LOWZ).

Finally, we consider the stellar populations of filament galaxies, as estimated via their mean colour, as determined from the luminosities in different bands, as follows:

$$^{0.1}(g-r) = -2.5\log_{10}\left(\frac{L_{^{0.1}g,\rm total}}{L_{^{0.1}r,\rm total}}\right) + M_{\odot,^{0.1}g} - M_{\odot,^{0.1}r}, \tag{3.10}$$

where the luminosity is measured in solar units, the magnitude of the Sun are $M_{\odot,^{0.1}g}$ = 5.45 and $M_{\odot,^{0.1}r}$ = 4.76 (Blanton et al., 2003), and uncertainties in the colour are propagated from the total luminosity errors. The results for different sample catalogues are summarised in the last column in Table 3.2.

3.3.5 Discussion

With the measurements of $M_{\rm total}$, $M_{\rm stellar,total}$ and $L_{\rm total}$ obtained in a consistent way, we calculate the total-mass-to-light ratio and the stellar-mass-to-total-mass ratio. The results are in Table 3.2, and shown in Figure 3.11. Within the uncertainties there is no significant evolution of either ratio from $z = 0.33$ to $z = 0.53$.

First, we compare our results with the universal average mass-to-light ratio from Loveday et al. (2015). Using GAMA-II galaxy sample, they estimated the comoving luminosity density evolution in the $^{0.1}r$ band over the redshift range $0.002 < z < 0.65$. Based on their results, the calculated M/L evolution as a function of redshift is shown in Figure 3.11, where

[2]Note that in the table, we exclude the Schechter fit results with combined LOWZ and CMASS sample as the $L_{\rm total}$ values for case LOWZ+CMASS are computed by averaging LOWZ and CMASS measurements.

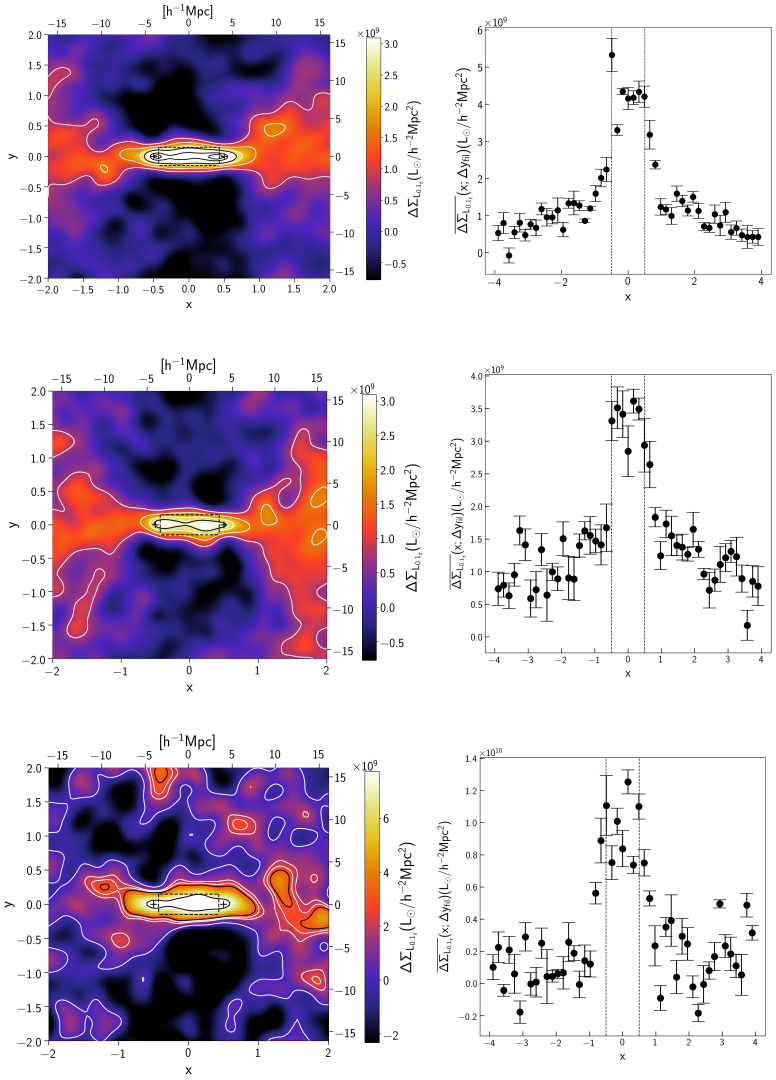

Figure 3.7: The excess 2D stacked Σ_L map asfter the subtraction using the whole LRG sample (top), CMASS sample (middle) and LOWZ sample (bottom). A Gaussian filter of smoothing scale 0.90 h^{-1}Mpc (0.10 in units of x, y) has been applied to the image to suppress small-scale noise. LRG halos in the standardized coordinate system are highlighted as "+" signs. For each case, the average excess luminosity density along the x axis is also shown on the right, where the average at each point of X is taken along $\Delta y = 2.45$ Mpc h^{-1} (from -0.15$R_{\rm sep}$ to 0.15 $R_{\rm sep}$).

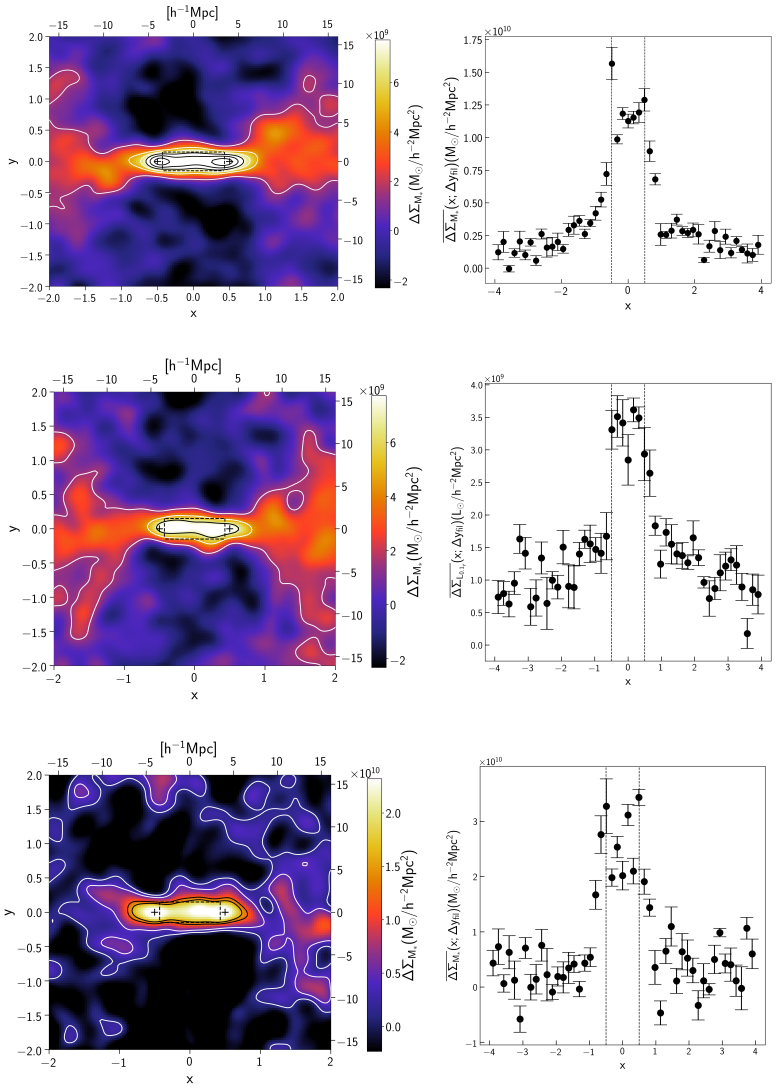

Figure 3.8: As in Figure 3.7, but for stellar mass.

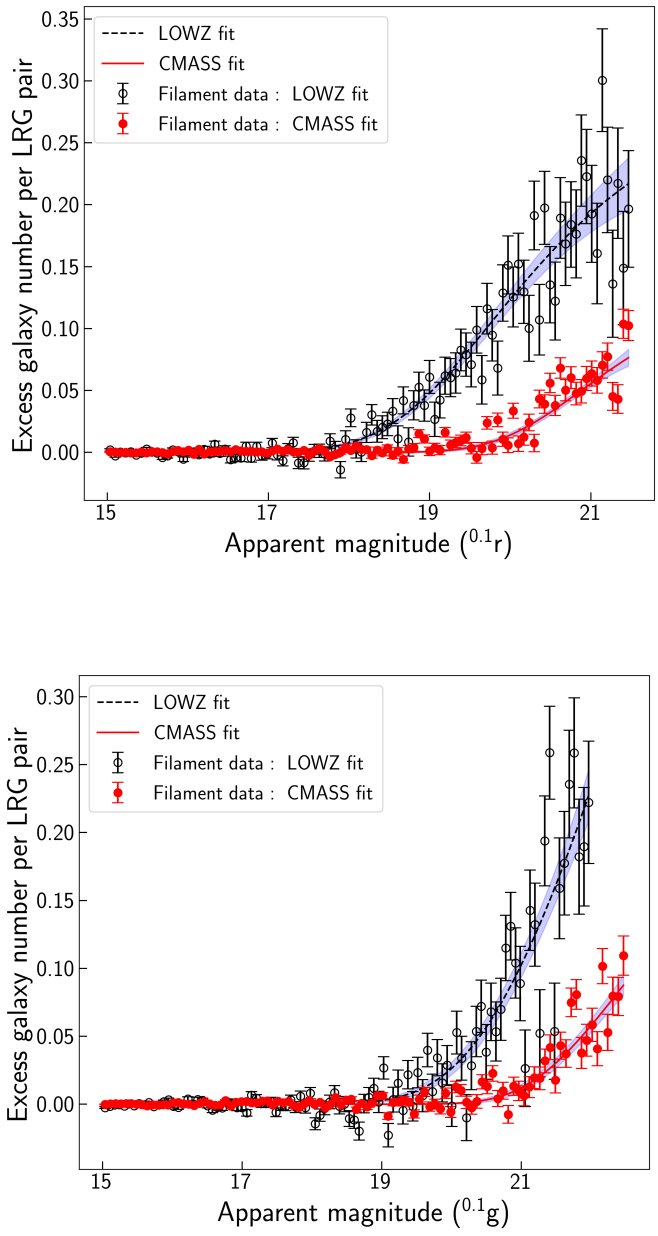

Figure 3.9: Observed excess number of galaxy per LRG pair as a function of apparent magnitude. The best-fit Schechter function for LOWZ and CMASS sample for different bands are over-plotted on the same figure. The best-fit line and its standard deviation are computed from 50 MCMC samples.

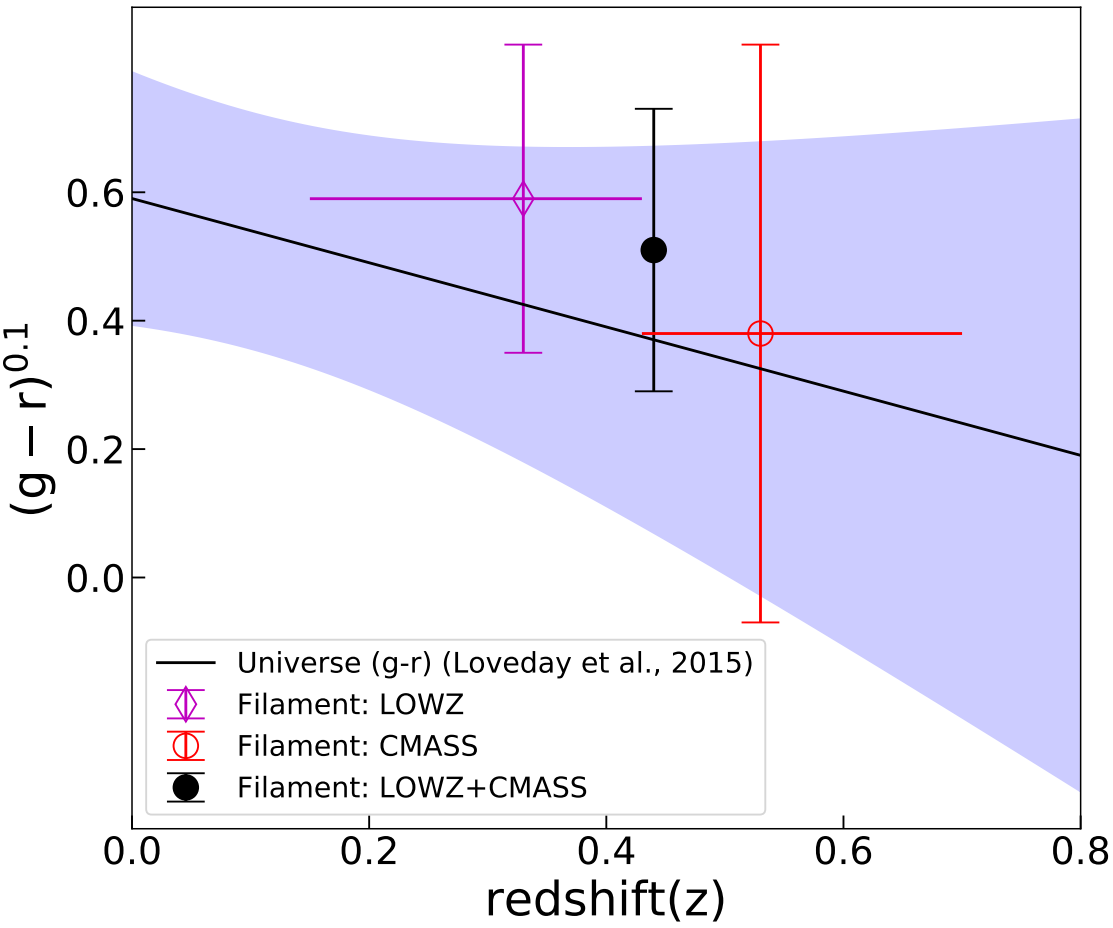

Figure 3.10: Filament color as a function of redshift. The black solid line with 1-σ shaded region shows the universe (g-r) value as a function of redshift predicted by Loveday et al. (2012). Magenta thin diamond and red hollow circle with errorbars are the colour values of filaments measured from LOWZ and CMASS. Black point with black errorbar shows the results from the combined samples.

Sample name	$\overline{z}_{\rm pair}$	M (1)	$M_{\rm stellar}$ (2)	$L_{^{0.1}g}$ (3)	$L_{^{0.1}r}$ (4)
$6h^{-1}$Mpc $\leqslant R_{\rm sep} \leqslant 10h^{-1}$Mpc					
LOWZ	0.33	4.85±1.36	4.08±0.77	1.70±0.32	1.57±0.19
CMASS	0.53	3.65±1.20	2.16±0.65	1.11±0.33	0.84±0.28
LOWZ+CMASS	0.44	4.25±0.91	3.12±0.50	1.41±0.23	1.21±0.15
$3h^{-1}$Mpc $\leqslant R_{\rm sep} \leqslant 5h^{-1}$Mpc					
LOWZ	0.30	4.35±1.35	1.78±0.74	0.56±0.23	0.73±0.26
CMASS	0.50	2.60±1.18	1.69±0.43	0.65±0.12	0.62±0.11
LOWZ+CMASS	0.40	3.48±0.90	1.74±0.43	0.61±0.13	0.68±0.14

Table 3.1: Summary for mass and luminosity measurements of filaments for three different catalogues, where M, $M_{\rm stellar}$, $L_{^{0.1}g}$, $L_{^{0.1}r}$ for LOWZ + CMASS sample are computed by taking the simple average between LOWZ and CMASS values. Note: (1): total mass from weak lensing. The unit is $10^{13}M_{\odot}$ and the uncertainty is mainly dominated by the shape noise. (2): total filament stellar mass in unit of $10^{11}M_{\odot}$. It is calculated by using the observed stellar-mass-to-light ratio from maps and the total light from MCMC. (3),(4): total luminosity in $^{0.1}g$ and $^{0.1}r$ band. The units are $10^{11}L_{\odot,^{0.1}g}$ and $10^{11}L_{\odot,^{0.1}r}$, $^{0.1}r$. For comparison, the last three rows include the results determined by using the Xia et al. (2019) selection criteria of filaments.

Sample name	$\overline{z}_{\rm pair}$	$M_{\rm stellar}/M$	$M/L_{^{0.1}r}$	$^{0.1}(g-r)$
$6h^{-1}$Mpc $\leqslant R_{\rm sep} \leqslant 10h^{-1}$Mpc				
LOWZ	0.33	$0.84 \pm 0.26\%$	309±94	0.59±0.24
CMASS	0.53	$0.59 \pm 0.26\%$	435±189	0.38±0.45
LOWZ+CMASS	0.44	$0.73 \pm 0.20\%$	351±87	0.51±0.22
$3h^{-1}$Mpc $\leqslant R_{\rm sep} \leqslant 5h^{-1}$Mpc				
LOWZ	0.30	$0.41 \pm 0.21\%$	596±282	0.97±0.59
CMASS	0.50	$0.65 \pm 0.34\%$	419±204	0.63±0.28
LOWZ+CMASS	0.40	$0.50 \pm 0.18\%$	512±169	0.80±0.32

Table 3.2: Summary for M/L, M_*/M and $^{0.1}(g-r)$ measurements of filaments for three different catalogues. For comparison, the last three rows include the results determined by using the Xia et al. (2019) selection criteria of filaments.

Sample name	$^{0.1}r$ band fit			$^{0.1}g$ band fit		
	n_*	α	$M_{*,^{0.1}r}$	n_*	α	$M_{*,^{0.1}g}$
$6-10h^{-1}$Mpc						
LOWZ	4.24±0.77	-1.05±0.09	-21.67±0.14	4.72±1.48	-1.38±0.17	-20.96±0.26
CMASS	2.33±1.15	-0.83±0.41	-21.63±0.38	2.78±1.27	-1.20±0.36	-21.08±0.39
$3-5h^{-1}$Mpc						
LOWZ	3.17±1.40	-0.72±0.27	-21.17±0.29	3.73±1.83	-0.81±0.52	-20.00±0.32
CMASS	2.30±0.48	-1.00 (fixed)	-21.34±0.12	4.06±0.79	-1.00 (fixed)	-20.09±0.10

Table 3.3: Table that summarizes the best-fit parameters of Schechter function for different samples in $^{0.1}r$ and $^{0.1}g$ band. Best fitted values and their corresponding uncertainties are determined from MCMC algorithm.

we have adopted a comoving matter density, for $\Omega_{m,0} = 0.3$, of $\rho_{\rm m} = 4.09 \times 10^{10} M_{\odot}{\rm Mpc}^{-3}$. At $z = 0.33$, their model predicts $M/L = 208 \pm 26$, and at $z = 0.53$ the ratio is 143±18. Our result is slightly higher than the universal average suggesting filaments are darker than the Universe as a whole.

The $M_{\rm stellar}/M$ as a function of redshift is illustrated in Figure 3.12. Based on the results from Moustakas et al. (2013), we can see that the universal average $M_{\rm stellar}/M$ does not significantly vary with redshift. For comparison, we also over-plot the result obtained by Bahcall & Kulier (2014), where they measured a constant $M_{\rm stellar}/M$ value = 0.9±0.1% at cosmic scales larger than 300 $h^{-1}kpc$ away from the BCG center. The stellar mass fraction values obtained in our study are consistent with the values found in both of these papers.

Our M/L measured in $^{0.1}r$ band ($\lambda_{^{0.1}r} \sim 5665$ Å) is close to the measurements obtained from Bahcall & Kulier (2014), where they found that $M/L_{^{0.25}i}$ ($\lambda_{^{0.25}i} = 6100$ Å) is around 400-500 at $R_{\rm clustercentric} \sim 7\ h^{-1}$ Mpc. Our M/L result is higher than the result obtained from clusters (Yang et al., 2007), where they measured a $M/L_{19.5}$ value of ~ 270 for a 10^{13} $M_{\odot}$ halo. This suggests that filament is a "darker" environment compared with galaxy groups.

With regards to the colour of the stellar populations of filament galaxies, due to the large uncertainties, we find no statistically significant evolution over the redshift range considered. We can compare our results with the universal-average colour values obtained from Blanton et al. (2003), where they measure the $^{0.1}(g-r)$ value by fitting the luminosity function at $z = 0.1$ using SDSS galaxies. In their study, the $^{0.1}(g-r)$ value obtained at $z = 0.1$ is 0.73 ± 0.04. In order estimate the evolution of the colour as a function of redshift,

we use the LF evolution parameters of Loveday et al. (2012). The colour, $^{0.1}\widetilde{(g-r)}$[3], of the universe changes from 0.43 at $z = 0.33$, 0.37 at $z = 0.44$, to 0.33 at $z = 0.53$. Using the best-fit parametric luminosity function from Loveday et al. (2012), the colour as a function of redshift is shown in Figure 3.10, where we also over-plot the colour measurements from filaments obtained by this study. The colours of filament galaxies are consistent with the universal average colour, given the uncertainties.

In summary, the picture that emerges is that, within the uncertainties, the mass-to-light ratios, stellar-to-total mass fraction and the colours are statistically consistent with the universal averages. One interpretation of this is that both the stellar mass fractions as a function of halo mass, and the halo abundances in the filament environment are close to the universal averages. Due to the large error bars, it is hard to put strong constraints on the redshift dependence of these quantities.

[3]To obtain an estimate on how the colour changes with redshift, we ignore all uncertainties in the parameter fits and only use the nominal values.

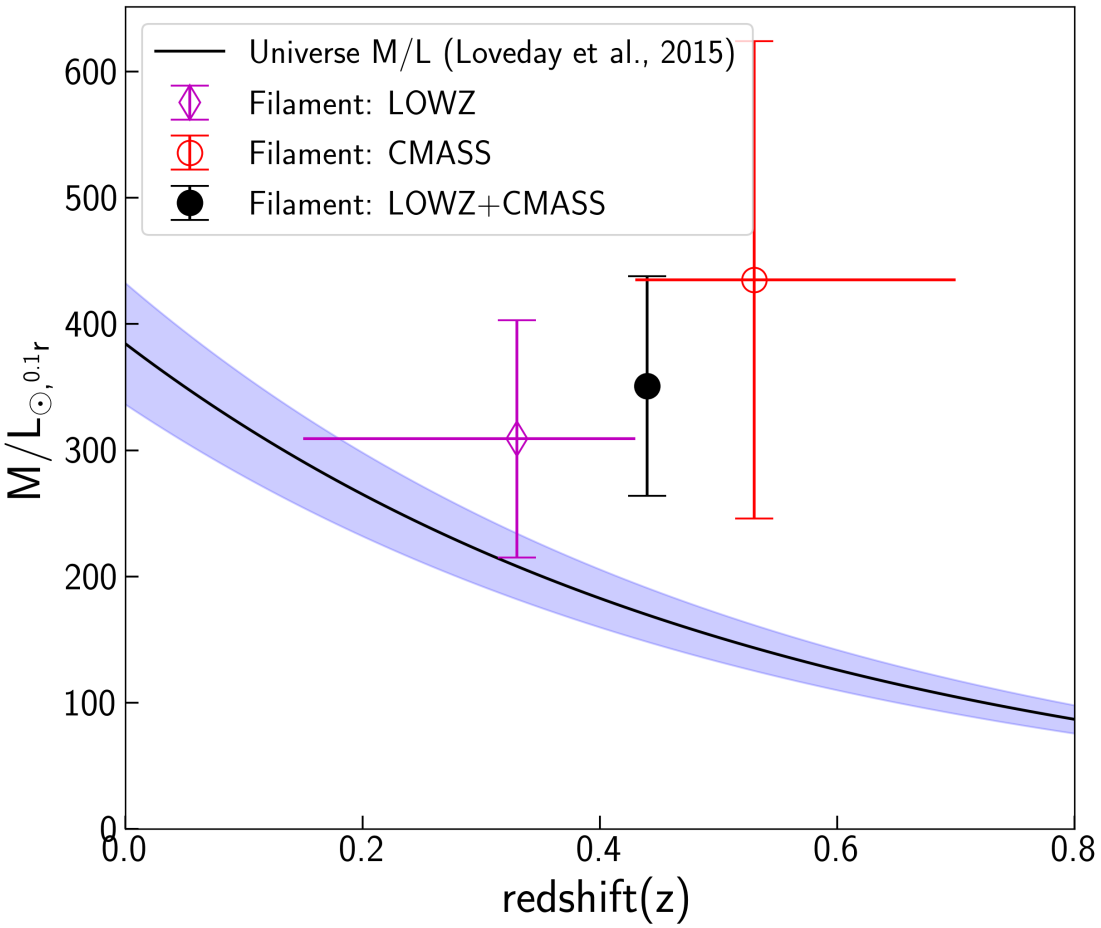

Figure 3.11: The filament M/L as a function of redshift. The black solid line with 1-σ shaded region shows the universe M/L value as a function of redshift predicted by Loveday et al. (2015). The thin magenta diamond and hollow red circle with error bars are the filament measurements from two independent samples (LOWZ and CMASS), where the horizontal error bars indicate the redshift coverage. The black circle with black errorbar shows the results from averaging LOWZ and CMASS.

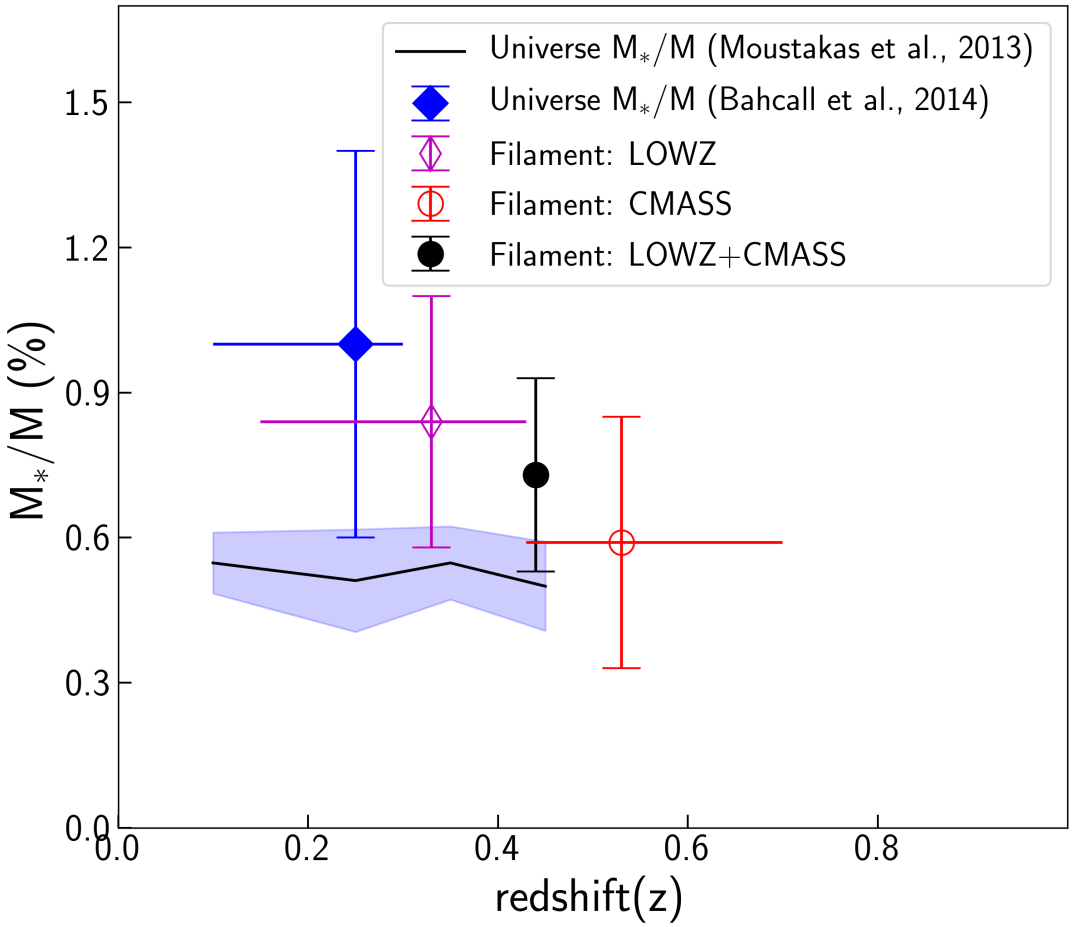

Figure 3.12: The filament M_*/M as a function of redshift. The black solid line with shaded region shows the universal average M_*/M value as a function of redshift from Moustakas et al. (2013), and the result from Bahcall & Kulier (2014) is also shown by the blue diamond..

Chapter 4

Conclusions

In this thesis, we present two studies related to probing dark matter from the Galaxy to the cosmic web. In the first study, we develop a novel method to constrain gravitational potentials from small-scale clustering in the space of action variables, and use it to provide precise constraints on the Milky Way dark matter halo potential within 9-15 kpc from Gaia DR2. In the second study, we investigate how dark the filaments in the cosmic web are by evaluating quantities such as mass-to-light ratios.

To constrain the Milky Way potential, we first derive the likelihood function for different host potentials, with the assumption that the stellar distribution in action space is a uniformly distributed background with correlated Gaussian fluctuations on small scales, showing that it can be written as an integral over the two-point correlation function evaluated in action space. The main advantage of our method is that, contrary to past studies, it does not require identification of any kind of compact structures or streams beforehand, assume the circular orbits, or any equilibrium state of the distribution. We first check the viability of our method in simulations of streams with different host potentials, showing that it recovers the normalization (slope) of the host potential with less than 1% (4%) systematic error (while stochastic errors shrink with the number of stars in the sample). We then apply our method to analyze two samples from Gaia DR2 over radial ranges of 9-11 kpc and 11.5-15 kpc, and studied the effect of selection cuts on the final results. Including all the known systematic errors, we find the parameters $(f_h, \alpha) = (0.391 \pm 0.009, 1.835 \pm 0.092)$ and $(0.351 \pm 0.012, 1.687 \pm 0.079)$, for 9-11 kpc and 11.5-15 kpc respectively, for the median and 68% CI uncertainty from the posterior distribution. For both simulations and real data, we can visually confirm that the potential that maximizes the likelihood function does indeed correspond to the largest two-point correlation function and most compact distribution in action space, which again demonstrates the reliability of our method.

We would like to clarify that the fraction of DM, f_h, characterizes the fraction of halo component contribution to the radial force *extrapolated to* the position of the Sun. Given the uncertainty, our final index constraints obtained at two radial bins are consistent within $\sim 1.2\sigma$. Based on the NFW prediction, the absolute value of index gets larger at outer radii. Therefore, a power law extrapolation back to the solar radius would lead to a larger local halo density. Therefore, the fact that we find a larger value of f_h in outer radii is consistent with the expectations from NFW (or any profile that gets steeper at larger radii).

To our knowledge, this is the first study that constrains the halo potential of the Milky Way using action space clustering with real data. While more work is needed to fully understand the systematic error of this method (as discussed in Section 2.4), its sheer statistical power is formidable as it scales with the number of all the stars in the sample, and with proper calibrations can provide exquisite constraints on dark matter potential. Further improvements (or checks) may come from identification of streams beforehand or other criteria to separate disk and halo components (as presented in e.g. Bonaca et al., 2017; Helmi et al., 2017; Myeong et al., 2018a; Necib et al., 2018). Additionally, in this study, we only varied the parameters in the local dark matter halo density profile but kept the stellar disk potential fixed. More robust constraints, left for future work, requires varying the parameters in the disk potential as well, and possibly include other probes of stellar density.

To investigate the M/L ratios of filaments, we first reproduce the mass map of stacked filament between two Luminous Red Galaxy pairs by following the same procedures outlined in Epps & Hudson (2017). Using weak gravitational lensing, we obtain a $\sim 4.5\sigma$ detection, with a mass estimate of $(4.25 \pm 0.91) \times 10^{13} M_{\odot}$. To isolate the filament, a catalogue of non-physical LRG (projected) pairs is constructed. We also conduct the analysis in two redshift-independent samples (LOWZ and CMASS) in order to capture any redshift dependence of filament mass, and we get $(4.85 \pm 1.36) \times 10^{13} M_{\odot}$ using LOWZ sample and $(3.65 \pm 1.20) \times 10^{13} M_{\odot}$ using CMASS sample. Using the same logic, we produce the 2D projected luminosity and stellar mass density map of stacked filaments. To compute the total luminosity, we fit a Schechter function to the observed excess number per LRG pair as a function of apparent magnitude. To estimate the stellar type of filament galaxies, luminosity analysis is conducted at SDSS g and r bands. Then we calculate the M/L ratios, M_*/M ratios and ${}^{0.1}(g-r)$ of filaments using different samples. We find $M/L = 309 \pm 94$, $M_*/M = 0.84 \pm 0.26\%$, ${}^{0.1}(g-r) = 0.59 \pm 0.24$ for LOWZ sample, and $M/L = 435 \pm 189$, $M_*/M = 0.59 \pm 0.26\%$, ${}^{0.1}(g-r) = 0.38 \pm 0.45$ for CMASS sample. If we combine both samples, we find $M/L = 351 \pm 87$, $M_*/M = 0.73 \pm 0.20\%$, ${}^{0.1}(g-r) = 0.51 \pm 0.22$. From this study, we demonstrate that the cosmic web, although dominated by the dark matter,

is not entirely dark. We then compare our measured M/L and M_*/M values of filaments with the same quantities measured at different cosmic environments, finding that they are comparable to the cosmic mean value.

The uncertainties remain large for this analysis, given current data. The advent of other surveys could do help to provide more accurate and precise measurements. For instance, the recently-completed Dark Energy Survey (DES: The Dark Energy Survey Collaboration, 2005) has promised to obtain the shape measurements and photo-zs of galaxies with $\sigma(z) \sim 0.07$ out to $z > 1$ over an area of 5,000 degree^2. Other ongoing ground-based surveys such as the Hyper Suprime-Cam Subaru Strategic Program (HSC-SSP: Aihara et al., 2018) and the Canada-France Imaging Survey (CFIS: Ibata et al., 2017) will provide ellipticities and photometry measurements in much wider and deeper region compared with CFHTLens. For example, HSC-Wide survey will take images 1 magnitude deeper within region nearly ten times larger than CFHTLens. The number density will reach approximately 30 per square arcminute (almost twice "denser" than CFHTLens), which will provide unprecedented accuracy to shape measurements for weak lensing analysis. Surveys such as SDSS-IV Extended Baryon Oscillation Spectroscopic Survey (eBOSS: Dawson et al., 2016) and the Dark Energy Spectroscopic Instrument survey (DESI: DESI Collaboration et al., 2016) will measure more galaxy redshifts, mapping large-scale structure in deeper and larger volumes. Compared with BOSS, there are at least 300,000 more LRGs can be accurately identified over a redshift range $0.6< z <1.0$ in eBOSS, which can help us map the large-scale structure distribution in a deeper and larger volume. Furthermore, some planned large space-based missions such as Euclid (Laureijs et al., 2011) and WFIRST (Spergel et al., 2015) are free from contamination and limitations of ground-based missions. These surveys are able to provide more precise and deeper spectroscopic and photometric redshift estimates. For WFIRST, the effective number density of galaxies can reach approximately 45 arcminute^{-2} for weak lensing analysis and it can improve the accuracy of photometric redshift measurements for hundreds of millions of galaxies out to $z \sim 3$. For Euclid, the effective number density of galaxies can be approximately 30 arcminute^{-2} and the photometric redshift measurements can reach a precision of $\sigma_z/(1+z) < 0.05$ out to $z \sim 2$. Therefore, these large surveys can be helpful to largely shrink down the uncertainties in both mass and luminosity measurements. With the increase in the accuracy and completeness of different surveys, we hope to provide more reliable and complete measurements of the cosmic web.

References

Afshordi N., Mohayaee R., Bertschinger E., 2009, PRD, 79, 083526

Aihara H., et al., 2018, PASJ, 70, S4

Bahcall N. A., Kulier A., 2014, MNRAS, 439, 2505

Baumgardt H., 2017, MNRAS, 464, 2174

Beck R., Dobos L., Budavári T., Szalay A. S., Csabai I., 2016, MNRAS, 460, 1371

Benítez N., 2000, APJ, 536, 571

Bernardeau F., 1994, APJ, 433, 1

Bernardeau F., Colombi S., Gaztañaga E., Scoccimarro R., 2002, PHYSREP, 367, 1

Binney J., 2012, MNRAS, 426, 1324

Binney J., McMillan P. J., 2016, MNRAS, 456, 1982

Blanton M. R., et al., 2003, APJ, 592, 819

Bolton A. S., et al., 2012, AJ, 144, 144

Bonaca A., Conroy C., Wetzel A., Hopkins P. F., Kereš D., 2017, APJ, 845, 101

Bond J. R., Kofman L., Pogosyan D., 1996, NAT, 380, 603

Bovy J., 2014, APJ, 795, 95

Bovy J., 2015, APJS, 216, 29

Bovy J., Rix H.-W., 2013, APJ, 779, 115

Bovy J., Bahmanyar A., Fritz T. K., Kallivayalil N., 2016, APJ, 833, 31

Buckley M. R., Hogg D. W., Price-Whelan A. M., 2019, arXiv e-prints, p. arXiv:1907.00987

Budavári T., Wild V., Szalay A. e. S., Dobos L., Yip C.-W., 2009, MNRAS, 394, 1496

Clampitt J., Miyatake H., Jain B., Takada M., 2014, arXiv e-prints, p. arXiv:1402.3302

Colberg J. M., Krughoff K. S., Connolly A. J., 2005, MNRAS, 359, 272

Colless M., et al., 2003, arXiv e-prints, pp astro–ph/0306581

Cropper M., et al., 2018, arXiv e-prints, p. arXiv:1804.09369

DESI Collaboration et al., 2016, arXiv e-prints, p. arXiv:1611.00036

Dawson K. S., et al., 2013, AJ, 145, 10

Dawson K. S., et al., 2016, AJ, 151, 44

Dietrich J. P., Werner N., Clowe D., Finoguenov A., Kitching T., Miller L., Simionescu A., 2012, NAT, 487, 202

Dobos L., Csabai I., Yip C.-W., Budavári T., Wild V., Szalay A. e. S., 2012, MNRAS, 420, 1217

Eilers A.-C., Hogg D. W., Rix H.-W., Ness M. K., 2019, APJ, 871, 120

Eisenstein D. J., et al., 2011, AJ, 142, 72

Epps S. D., Hudson M. J., 2017, MNRAS, 468, 2605

Erben T., et al., 2013, MNRAS, 433, 2545

Flynn C., Holmberg J., Portinari L., Fuchs B., Jahreiß H., 2006, MNRAS, 372, 1149

Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, PASP, 125, 306

Gaia Collaboration et al., 2016, AAP, 595, A1

Gaia Collaboration et al., 2018, AAP, 616, A1

Gott J. Richard I., Jurić M., Schlegel D., Hoyle F., Vogeley M., Tegmark M., Bahcall N., Brinkmann J., 2005, APJ, 624, 463

Gravity Collaboration et al., 2018, AAP, 615, L15

Gregory S. A., Thompson L. A., Tifft W. G., 1981, APJ, 243, 411

Gressman P. T., Strain R. M., 2010, arXiv e-prints, p. arXiv:1011.5441

Guzzo L., et al., 2014, AAP, 566, A108

Hahn O., Porciani C., Carollo C. M., Dekel A., 2007, MNRAS, 375, 489

Helmi A., Veljanoski J., Breddels M. A., Tian H., Sales L. V., 2017, AAP, 598, A58

Heymans C., et al., 2012, MNRAS, 427, 146

Higuchi Y., Oguri M., Shirasaki M., 2014, MNRAS, 441, 745

Higuchi Y., Oguri M., Tanaka M., Sakurai J., 2015, arXiv e-prints, p. arXiv:1503.06373

Hildebrandt H., et al., 2012, MNRAS, 421, 2355

Huchra J. P., et al., 2012, APJS, 199, 26

Hudson M. J., et al., 2015, MNRAS, 447, 298

Ibata R. A., et al., 2017, APJ, 848, 128

Jain B., Bertschinger E., 1994, APJ, 431, 495

Jauzac M., et al., 2012, MNRAS, 426, 3369

Kaiser N., Squires G., 1993, APJ, 404, 441

Kondo H., Miyatake H., Shirasaki M., Sugiyama N., Nishizawa A. J., 2019, arXiv e-prints, p. arXiv:1905.08991

Kullback S., Leibler R. A., 1951, Ann. Math. Statist., 22, 79

Laureijs R., et al., 2011, arXiv e-prints, p. arXiv:1110.3193

Libeskind N. I., et al., 2018, MNRAS, 473, 1195

Lietzen H., et al., 2016, AAP, 588, L4

Loveday J., et al., 2012, MNRAS, 420, 1239

Loveday J., et al., 2015, MNRAS, 451, 1540

Magorrian J., 2014, MNRAS, 437, 2230

Mandelbaum R., Seljak U., Cool R. J., Blanton M., Hirata C. M., Brinkmann J., 2006, MNRAS, 372, 758

Marinoni C., Hudson M. J., 2002, APJ, 569, 101

McMillan P. J., 2017, MNRAS, 465, 76

Mead J. M. G., King L. J., McCarthy I. G., 2010, MNRAS, 401, 2257

Miller L., et al., 2013, MNRAS, 429, 2858

Miyatake H., et al., 2015a, APJ, 806, 1

Miyatake H., et al., 2015b, APJ, 806, 1

Mo H., van den Bosch F. C., White S., 2010, Galaxy Formation and Evolution

Montero-Dorta A. D., Prada F., 2009, MNRAS, 399, 1106

Moustakas J., et al., 2013, APJ, 767, 50

Myeong G. C., Evans N. W., Belokurov V., Amorisco N. C., Koposov S. E., 2018a, MNRAS, 475, 1537

Myeong G. C., Evans N. W., Belokurov V., Sand ers J. L., Koposov S. E., 2018b, APJL, 863, L28

Narayan R., Bartelmann M., 1999, in Dekel A., Ostriker J. P., eds, Formation of Structure in the Universe. p. 360

Navarro J. F., Frenk C. S., White S. D. M., 1996, APJ, 462, 563

Necib L., Lisanti M., Belokurov V., 2018, arXiv e-prints, p. arXiv:1807.02519

Oort J. H., 1983, ARAA, 21, 373

Parejko J. K., et al., 2013, Monthly Notices of the Royal Astronomical Society, 429, 98

Parker L. C., Hudson M. J., Carlberg R. G., Hoekstra H., 2005, APJ, 634, 806

Peñarrubia J., Koposov S. E., Walker M. G., 2012, APJ, 760, 2

Peebles P. J. E., 1980, The large-scale structure of the universe

Planck Collaboration et al., 2016, AAP, 594, A13

Sanders J., 2012, MNRAS, 426, 128

Sanders J. L., Binney J., 2013, MNRAS, 433, 1826

Sanderson R. E., Helmi A., Hogg D. W., 2015, APJ, 801, 98

Sanderson R. E., Hartke J., Helmi A., 2017, APJ, 836, 234

Schechter P., 1976, APJ, 203, 297

Shectman S. A., Landy S. D., Oemler A., Tucker D. L., Lin H., Kirshner R. P., Schechter P. L., 1996, APJ, 470, 172

Sheldon E. S., et al., 2009, APJ, 703, 2232

Sparke L. S., Gallagher John S. I., 2007, Galaxies in the Universe: An Introduction

Spergel D., et al., 2015, arXiv e-prints, p. arXiv:1503.03757

Springel V., et al., 2005, NAT, 435, 629

Tanimura H., et al., 2019, MNRAS, 483, 223

The Dark Energy Survey Collaboration 2005, arXiv e-prints, pp astro–ph/0510346

Tortora C., Napolitano N. R., Romanowsky A. J., Capaccioli M., Covone G., 2009, MNRAS, 396, 1132

Vasiliev E., 2019, MNRAS, 484, 2832

Vogelsberger M., et al., 2014, MNRAS, 444, 1518

Wang Y., et al., 2019, PASA, 36, e015

Xia Q., et al., 2019, arXiv e-prints, p. arXiv:1909.05852

Yang X., Mo H. J., van den Bosch F. C., Pasquali A., Li C., Barden M., 2007, APJ, 671, 153

Yang T., Sarma Boruah S., Afshordi N., 2019, arXiv e-prints, p. arXiv:1908.02336

Zel'dovich Y. B., 1970, AAP, 500, 13

Zeldovich I. B., Einasto J., Shandarin S. F., 1982, NAT, 300, 407

de Graaff A., Cai Y.-C., Heymans C., Peacock J. A., 2019, AAP, 624, A48

de Lapparent V., Geller M. J., Huchra J. P., 1986, APJl, 302, L1

de Zeeuw T., 1985, MNRAS, 216, 273

APPENDICES

Appendix A

Derivation of the Likelihood Function

In this section, we discuss the derivation of the likelihood which we use to constrain the parameters of the Milky Way potential. Using Bayes' Theorem, we have,

$$\mathcal{P}(\Theta|\mathcal{D}) = \frac{\mathcal{P}(\mathcal{D}|\Theta)\mathcal{P}(\Theta)}{\mathcal{P}(\mathcal{D})} \tag{A.1}$$

where Θ are the parameters we want to constrain, while $\mathcal{D}$ is the data we have available. In particular, in the context of constraining the Milky Way potential which is our goal here, Θ represents the parameters of the potential, (f_h, α).

Let us divide the action space into M small bins such that the average number of stars *per bin* (over all of action space) is $\bar{n} \ll 1$. However, because of hierarchical structure formation, the expected number counts of stars in different bins will not be independent of each other. In order to capture this, we assume that the star count in bin a is a Poisson sampling of a mean, $\bar{n} + \bar{n}\chi_a$, where χ_a's are correlated random gaussian variables. Therefore, the probability of measuring star counts $\{\nu_a\}$ is given by:

$$\mathcal{P}(\{\nu_a\}, \{\chi_a\}) = \frac{\exp\left(-\frac{1}{2}\sum_{a,b}\chi_a\xi_{ab}^{-1}\chi_b\right)}{\sqrt{\det(\xi)}(2\pi)^{M/2}} \prod_a \frac{(\bar{n} + \bar{n}\chi_a)^{\nu_a}\exp(-\bar{n} - \bar{n}\chi_a)}{\nu_a!}. \tag{A.2}$$

From this definition, it follows that

$$\langle\chi_a\chi_b\rangle = \xi_{ab}, \tag{A.3}$$

is the covariance matrix of the random gaussian variables $\{\chi_a\}$. Since we do not directly observe χ_a's, we should marginalize over them. Therefore, in the limit of $M \to \infty$ (i.e.

when $\nu_a = 0$ or 1) the posterior (Equation A.1) is given by:

$$\mathcal{P}(f_h, \alpha | \{a_k\}) \propto \mathcal{P}_{\rm prior}(f_h, \alpha) \int \frac{\exp\left(-\frac{1}{2}\sum_{a,b}\chi_a \xi^{-1}_{ab}\chi_b\right)}{\sqrt{\det(\xi)}(2\pi)^{M/2}} \left\{\prod_k [1+\chi_{a_k}]\right\} \prod_a d\chi_a, \quad \text{(A.4)}$$

where a_k is the action-space bin in which the k-th star lies, and we assume $\sum_a \chi_a = 0$ over the entire action space. Using Wick's theorem, the above Gaussian integral can be expressed in terms of a sum of the product of 2-point functions over all possible pairings of stars. This yields the likelihood (defined as the ratio of posterior to prior):

$$\begin{aligned}\mathcal{L}(\text{data}|f_h, \alpha) \equiv \frac{\mathcal{P}(f_h, \alpha|\{a_k\})}{\mathcal{P}_{\rm prior}(f_h, \alpha)} &\propto \left\langle \prod_k [1+\chi_{a_k}] \right\rangle \propto \sum_{\text{pairings}} \prod_{\text{pairs}} [1+\xi_{\rm pair}] \\ &\propto \left\langle \exp\left[\sum_{\text{pairs}} \ln\left(1+\xi_{\rm pair}\right)\right] \right\rangle_{\text{pairings}}. \quad \text{(A.5)}\end{aligned}$$

Now, for a large number of pairs, we can expect that the exponent in this expression have smaller and smaller relative fluctuations around its mean for different possible pairings. This is often known as the *mean-field approximation* in statistical mechanics (where the sum represents the partition function and the exponent is proportional to the energy), and allows us to move the average inside the exponent:

$$\ln \mathcal{L}(\text{data}|f_h, \alpha) \approx \left\langle \sum_{\text{pairs}} \ln\left(1+\xi_{\rm pair}\right) \right\rangle_{\text{pairings}} \quad \text{(A.6)}$$

This equation defines our log-likelihood formula adopted in Equation 2.5 in the main text (and subsequent statistical analyses), where we further assume that the 2-point function ξ only depends on the normalized distance D (Equation 2.2) in the action space.

Appendix B

Sanity Check for Stream-Only Simulations

To check whether the maximum likelihood test does correspond to the most clustering in the action space for either simulations, we plot the 2D projection of stellar distribution and two-point correlation function varying the choice of potential parameters f_h and α used in the action computation. Here, we summarize the results. Figure B.2 shows how a 2D projection of stellar distribution in the action space varies with different choices of potential for one of the simulations. In these figures, f_h (α) is fixed, while α (f_h) is varying across its correct value. As expected, for both simulations, the most compact distributions occur when parameters approach the correct values for the simulation (middle panel in both figures). This is also verified in the behavior of the two-point correlation function in Figure B.3, where one of the parameters is fixed and the other one is varying.For both simulations, we see that the two-point correlation function is indeed maximized around the expected value, which proves the viability of our method.

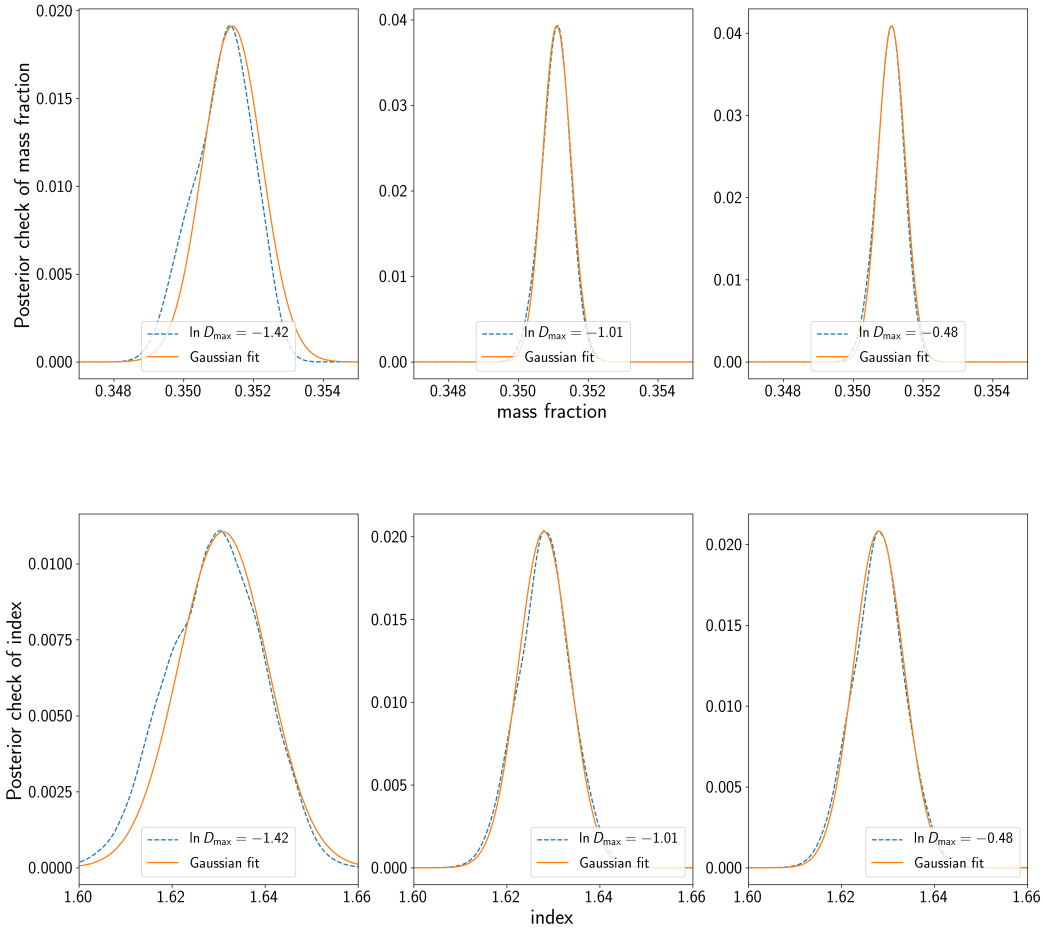

Figure B.1: The posterior distribution of f_h (upper panel) and α (lower panel) using the simulated power law potential with parameters $[f_h = 0.35,\ \alpha = 1.70]$. Distributions are computed at three different values of $D_{\max}$. For comparison, a gaussian function is also over-plotted on each case respectively.

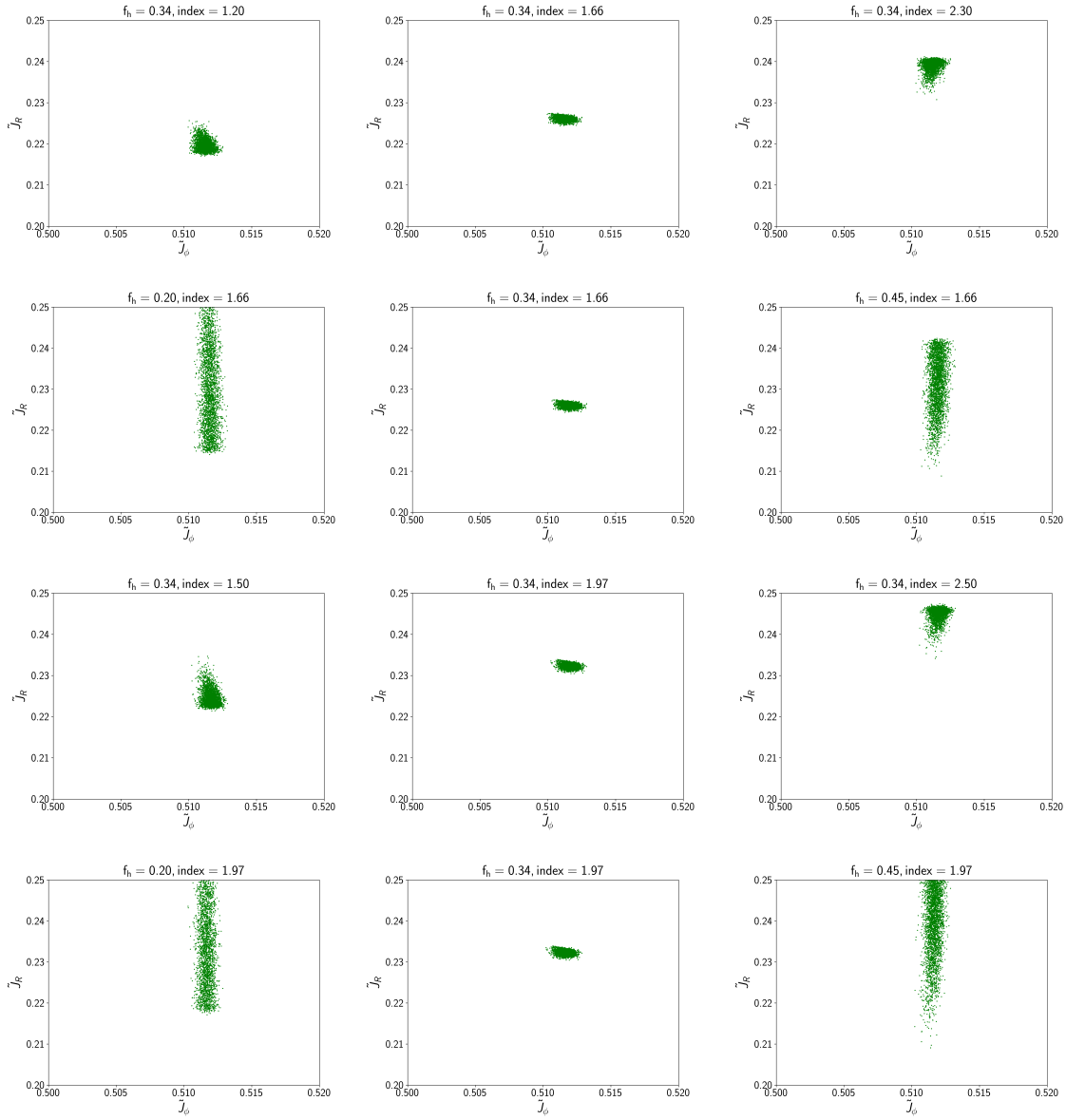

Figure B.2: Stellar distribution of one stream in $\tilde{J}_R$ – $\tilde{J}_\phi$ plane, where $\tilde{J}_R$ and $\tilde{J}_\phi$ are defined as J_R/σ_{J_R} and J_ϕ/σ_{J_ϕ}. We can see that the most clustering situation occurs when approaching the correct values ($[f_h = 0.35, \alpha = 1.70]$ for the first two rows and $[f_h = 0.35, \alpha = 2.00]$ for the last two).

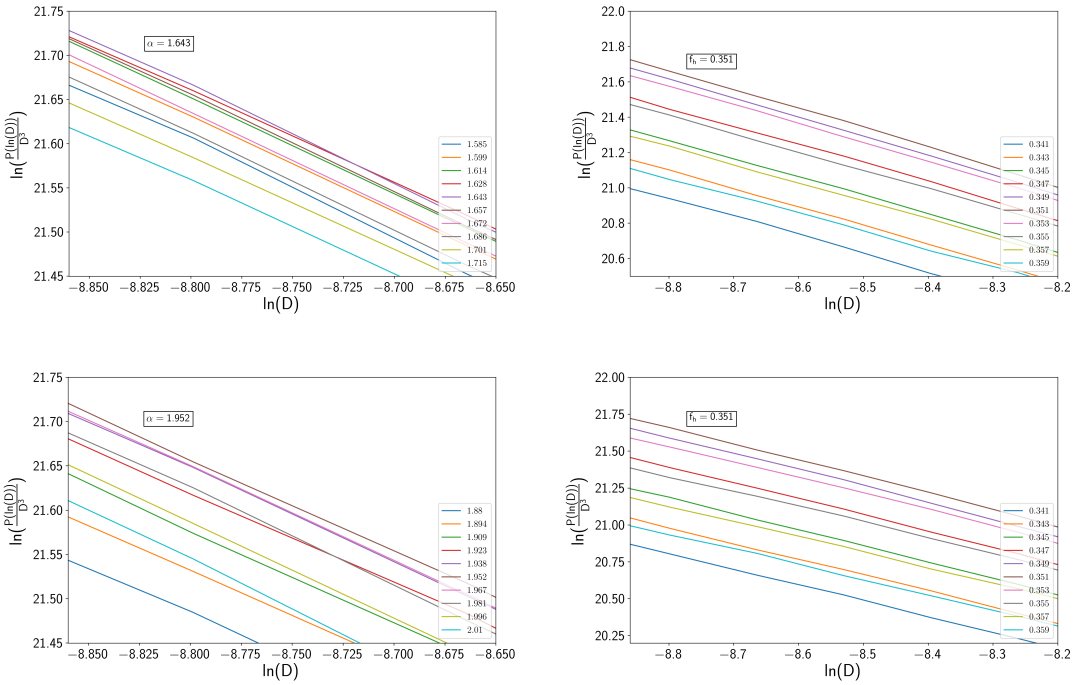

Figure B.3: Correlation function, $\frac{P(\ln D)}{D^3}$, as a function of $\ln D$. Top panel: $\frac{P(\ln D)}{D^3}$ vs $\ln D$ for case $[f_h = 0.35, \alpha = 1.70]$ with fixed f_h (α) on the left (right). Different colors indicate different parametrizations of potential. Bottom panel: $\frac{P(\ln D)}{D^3}$ vs $\ln D$ for case $[f_h = 0.35, \alpha = 2.00]$ with fixed f_h (α) on the left (right). As expected, while varying the choices of potential, the two-point correlation function is maximized around the correct value of parameter.

Appendix C

Simulation with the Inclusion of Background Stars

In Section 2.3.1, we conducted a simulation using a system that is entirely composed of stream stars. The discrepancies between the initial simulated host potentials and the parameters recovered by simulation are regarded as the systematic errors, and the systematic errors are further propagated to the real data analysis. However, as the assumption made in the likelihood derivation (Appendix A) is that the stellar distribution in action space is a uniform background plus gaussian fluctuations, in this section, we present the results obtained from another set of simulations with a realistic background.

To generate a uniform background in the action space, we take real observations from Gaia DR2 and then calculate their actions using a power law potential with [$f_h = 0.35$, $\alpha = 1.70$]. Then, we randomize the action distribution by adding a random gaussian scatter to each J_i. The scatter is generated from a gaussian distribution $N(0, \sigma^2_{\rm scatter})$, where $\sigma^2_{\rm scatter} = \frac{1}{3} < J_i^2 > D^2_{\rm max}$ with $\ln D_{\rm max}$ = -1 (equivalent to coarse-graining the action distribution using a gaussian filter of width $\ln D_{\rm max}$ = -1). Then, the randomized action variables are transformed back to the position and velocity in the cylindrical coordinates using the TorusMapper code developed by Binney & McMillan (2016), which is implemented using `galpy` package. Following these steps, a set of background stars, which are generated by randomizing the action distribution calculated from Gaia real data, can be produced. Therefore, we can then combine this background with three stream stars that are evolved in the same host potential for further analysis. Using the likelihood function defined in Equation 2.5, we will then test whether the recovered potential that corresponds to the most clustering distribution in action space is the same as the initial input.

Here, we include a wider prior (Equation 2.8) to capture whether there might be any double peak features that are presented in real data analysis. Figure C.1 shows the posterior distribution of two parameters. Although the posterior distribution of each parameter is dominated by a single peak around the correct value, at some values of $\ln D_{\rm max}$, the inclusion of the background stars seems to cause the distribution present a double peak features. In Figure C.3, we present the correlation function as a function of $\ln D$ and the error bar plots, where the error bars are determined by the quadratic fit (as introduced in Section 2.3.1) and from the posterior distribution (as introduced in Section 2.3.2), respectively. From the top panel, we can see that the two-point correlation is maximized when approaching the correct values. To be noted that there is a "dip" presented in the correlation function, and this is also reflected in the error bar plots: both of the f_h and α estimation do not drastically change with $\ln D_{\rm max}$ when $\ln D_{\rm max} \lesssim$ -2. This justifies our criteria for choosing the free parameter $\ln D_{\rm max}$. We want to choose a value of $\ln D_{\rm max}$ which gives the least uncertain measurements, and at $\ln D_{\rm max} \lesssim \ln D_{\rm max,optimum}$, the estimations of both parameters should be stabilized (not a strong function of $\ln D_{\rm max}$) and the constraints need to be all consistent with each other (within error bars). For this specific set of simulation, it should be $\ln D_{\rm max,optimum} \sim$ -2 (while in real data, we chose $\ln D_{\rm max,optimum} \sim$ -1). That is the point where our method is still valid and the constraints can be safely obtained with confidence. Also, compare the systematic errors obtained from this simulation with those got from Section 2.3.1, it seems like the inclusion of background stars improve the constraints, "bringing" the constraints recovered from likelihood function closer to the correct values (for example, at $\ln D_{\rm max} \sim$ -2, the constraint we get here is $f_h = 0.352 \pm 0.003$ $\alpha = 1.678 \pm 0.058$ from quadratic fit and $f_h = 0.352 \pm 0.003$ $\alpha = 1.657^{+0.066}_{-0.051}$ from median and posterior distribution, while for a stream-only simulation, we have $f_h = 0.352 \pm 0.001$ $\alpha = 1.634 \pm 0.014$ at same $\ln D_{\rm max}$ from quadratic fit only). Therefore, we here conclude that the systematic errors estimated from a stream-only simulation should be conservative, and we thus propagate those systematic errors to the real data analysis.

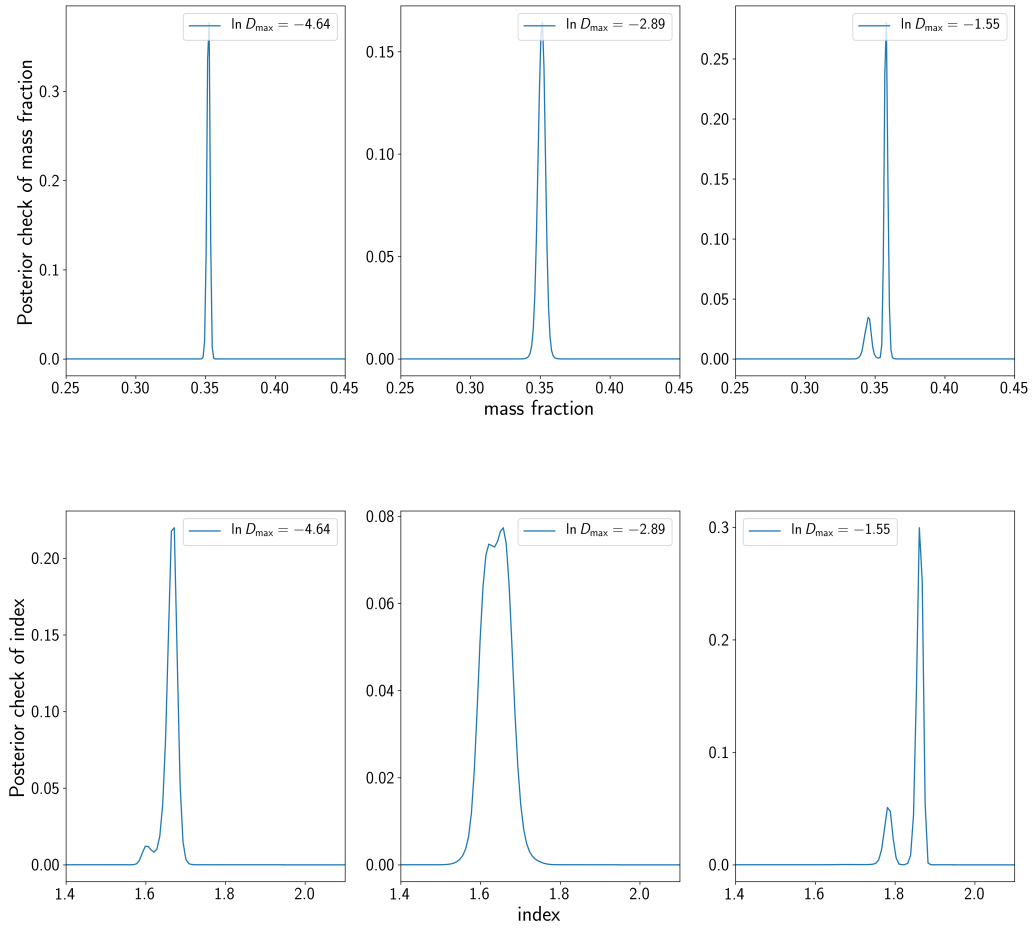

Figure C.1: The posterior distribution of f_h (upper panel) and α (lower panel) using for simulation including background stars. Distributions are computed at three different values of $D_{\max}$.

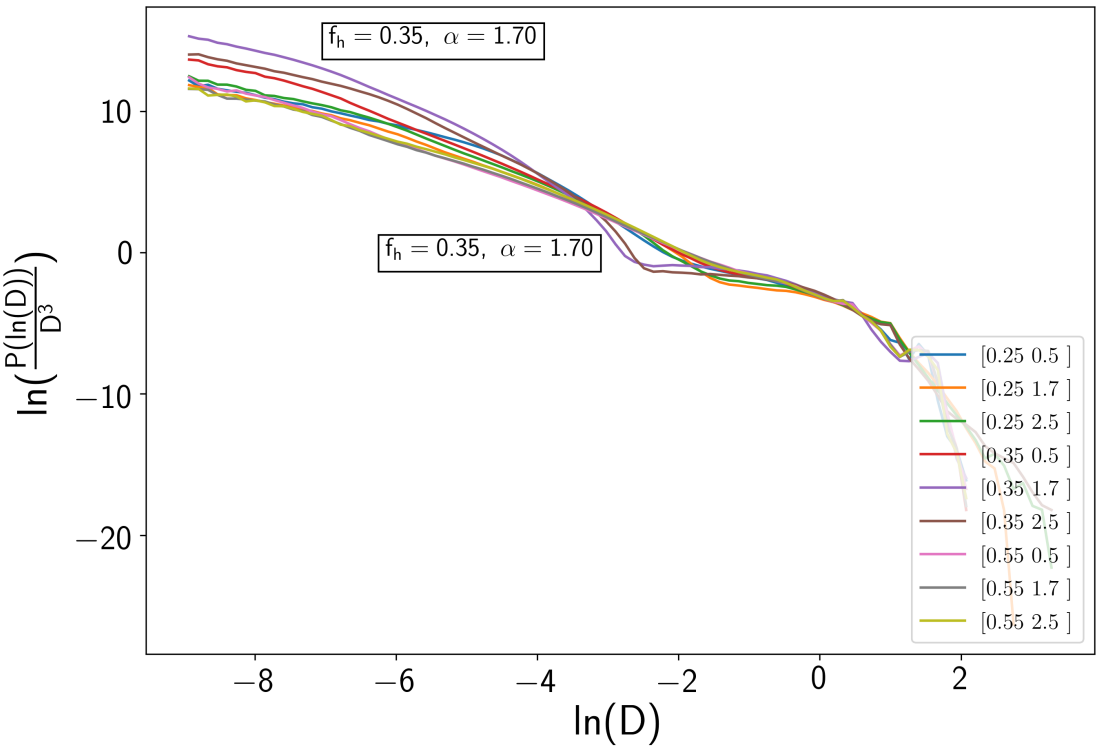

Figure C.2: $\frac{P(\ln D)}{D^3}$ as a function of $\ln D$ in action space. The two-point correlation function is indeed maximized around the expected values.

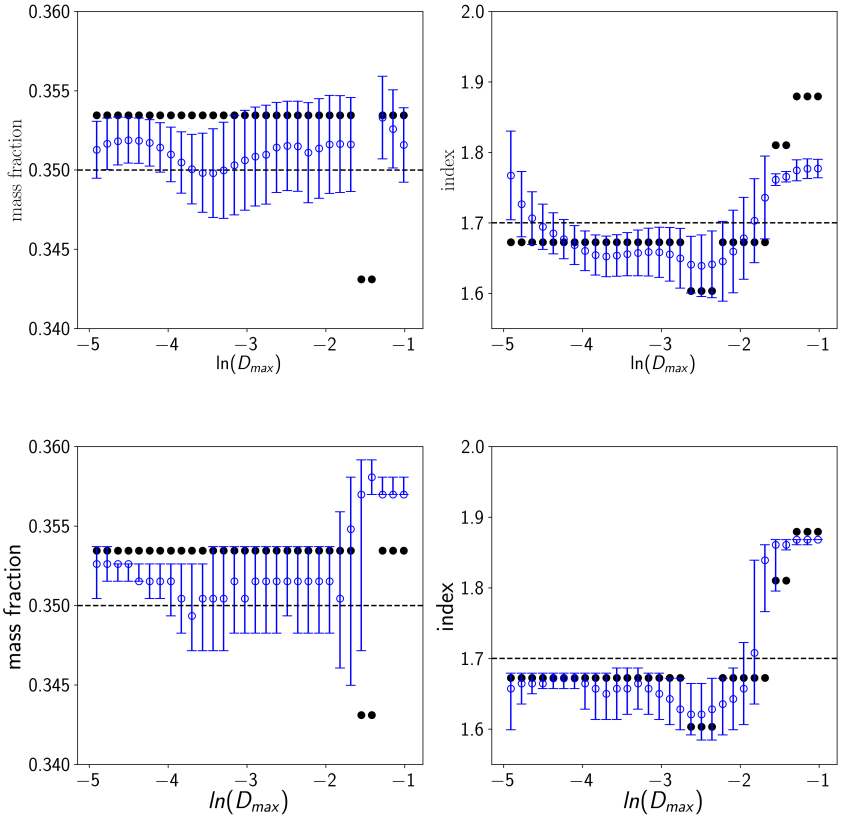

Figure C.3: Error bar plot (blue hollow points with error bars) as a function of $\ln D_{\max}$, where the error bars are determined from quadratic fit (upper panel) or by finding the 68%CI around the median (lower panel). Black solid points shows the maximum of likelihood. It can be seen that the estimates are stabilized until $\ln D_{\max}$ is smaller than $\sim$ -2, which is consistent with the trend showed in the correlation function plot.

Appendix D

Combining Measurements with Unknown Systematic Errors

Here, we discuss how to combine measurements x_i (of a single quantity x) that have independent known stochastic gaussian errors σ_i, as well as an unknown (but independent) systematic gaussian error $\sigma_{\rm sys}$. The joint likelihood is given by:

$$\mathcal{L}(x, \sigma_{\rm sys} | \{x_i, \sigma_i\}) = \prod_i \frac{\exp\left[-\frac{(x-x_i)^2}{2(\sigma_i^2+\sigma_{\rm sys}^2)}\right]}{\sqrt{2\pi(\sigma_i^2+\sigma_{\rm sys}^2)}}. \tag{D.1}$$

Now, assuming a flat prior on $\sigma_{\rm sys}$, up to some maximum $\sigma_{\rm sys,max}$, we can find the posterior on the parameter x:

$$P(x) \propto \int_0^{\sigma_{\rm sys,max}} d\sigma_{\rm sys} \prod_i \frac{\exp\left[-\frac{(x-x_i)^2}{2(\sigma_i^2+\sigma_{\rm sys}^2)}\right]}{\sqrt{2\pi(\sigma_i^2+\sigma_{\rm sys}^2)}}. \tag{D.2}$$

In practice, based on the difference in results we find with and without the cuts, we make conservative choices of $\sigma_{\rm sys,max} = 0.02$ and 0.2 for f_h and α determinations, respectively.