



**Universität
Zürich**^{UZH}

Cosmic Variance in Milky Way-Like Galaxies at $z=0$

A thesis presented by
Peter Espenshade
to The Institute of Physics
under the supervision of
Prof. Dr. Jaiyul Yoo
in partial fulfillment of the
requirements for the degree of
Master of Science
in the subject of Physics

University of Zürich
January 2020

Abstract

This thesis is devoted to exploring the errors inherent in cosmological observations: the cosmic variance which is a combination of the error inherent in an observer's position, the direction they look, and their survey volume. The exploration and application of the cosmic variance in literature are typically lacking or limited. We offer direct results from a suite of dark matter-only simulations as well as theoretical models to analyze our results. We focus on Milky Way-based observations at $z = 0$ for a large variety of survey volumes. We compute the matter density, cosmic variance, and 2-point correlation for dark matter and halo species. We found that the mean and variance can be replicated analytically to varying degrees of success. A future work might consider a variable redshift as well as smaller survey volumes and their effect on cosmological observations. Additionally, one might try to apply our findings to the local Hubble tension as well as the missing baryon problem.

Contents

1	Introduction	1
2	Theoretical cosmology framework	3
2.1	The flat, smooth, and isotropic Universe	3
2.2	Cosmological perturbation theory	5
2.3	Statistics for inhomogeneities	6
2.4	Gaussianity and high-order correlators	7
2.5	Linear evolution of perturbations	9
2.6	Bias and dark matter halos	12
2.7	Spherical collapse and the halo mass function	13
3	N-body simulations	16
3.1	Simulation methods	16
3.2	Initial conditions	18
3.3	Dark matter-only simulations	19
3.4	Simulation approximations	20
4	Simulation measurements	22
4.1	Halo finding	22
4.2	Simplifying n-dimensional integrals with probability distributions	23
4.3	The halo model	24
4.4	Window functions	26
4.5	Densities over spherical sub-volumes	27
4.6	Computation of a conditioned density and variance	28
5	The statistics of number counts	33
5.1	Conditional counting and correlation	33
5.2	The mean and variance of a conditional number density	35
6	Results	39
6.1	Halo mass function	39
6.2	2-point correlation and linear bias	40
6.3	Power spectrum	41
6.4	The conditional density's mean and variance (full-sky)	41
6.5	The conditional density's mean and variance (variable solid angle)	44
7	Future direction and conclusion	53

Appendices	55
A Power spectrum manipulation	55
B Power spectrum and the transfer function	55
C Indexing $2^n \cdot R$ spherical sub-volumes	56
D A probability density function transformation	58
References	62

1 Introduction

Observations in cosmology can often span distances on order of millions of parsecs to even billions of parsecs. The telescopes and surveys used can have field-of-views that also span enormous ranges from fractions of an arcminute to a sizeable portion of the night sky. As a consequence, the volume surveyed varies wildly as a function of geometry and, as such, the number of celestial objects counted varies wildly as well. However, although we can control the direction and the depth of our observations, they are all made from the same position – our home in the Milky Way.

Our observation location is fixed in a gravity-well which means the local matter structure is much more dense than surrounding regions. And there are small-scale inhomogeneities that affect the gravity-well depending on which direction we look. So, the parameters we measure locally would vary if we could move around the Universe. We cannot, of course, directly measure this local parameter variation and we term this the cosmic variance. The cosmic variance quantifies how much measurements in one direction vary from another direction when made from the Earth.

Measuring the cosmic variance is the same as measuring the small-scale deviation from the cosmological principle. Note that when we say small-scale we mean that an observer’s survey volume is not deep enough and/or wide enough such that local structure biases our measurement. It is important to keep in mind that the cosmic variance can be the same as the sample variance. Sample variance is the variance due to the variation between samples when they are taken from a parent population. We can take our parent population to be the set of all observations that one can make from the Earth, then the sample and cosmic variance are the same.

In the realm of spherical cows, matter would be distributed according to a Poisson process. And the statistics of number counts would be greatly simplified. Statistics in this homogeneous and isotropic Universe would not rely so heavily on surveys volumes or their positions. In actuality, there exists three sources of error in cosmological observations: the position of our observer, the direction our observer looks, and the survey’s volume. We conveniently call these three effects the cosmic variance. Cosmic variance is important when the sample volume of a survey is comparable or smaller than the intrinsic clustering scale of the celestial object.

One use of the cosmic variance is in constraining the variance of the local expansion rate (the Hubble parameter). There is a well-known tension between recent measurements. [1] uses the local recession velocities of galaxies resulting in $H_0 = 73.8 \pm 2.4 \text{ km s}^{-1}\text{Mpc}^{-1}$. And [2] uses a Λ CDM model to fit to observations of the cosmic microwave background and baryon acoustic oscillations yielding $H_0 = 67.80 \pm 0.77 \text{ km s}^{-1}\text{Mpc}^{-1}$. This is a startlingly large discrepancy of 9%.

The cosmic variance for the Hubble tension was approximated in [3] and found to be 2.4% (for $z > 0.010$) or 1.3% (for $z > 0.023$). According to their result, this is not enough

to explain the Hubble tension and only strengthens it. However, their method is quite approximative. They assume a lognormal mass distribution of perturbations (supported by [4]) and only consider the effect of redshift on the variance. However, we expect there to be a non-negligible effect due to the survey solid angle and its direction which would increase their estimate of the variance. There is a more recent computation of the cosmic variance in [5] which also concludes that the Hubble tension is not yet solved. This thesis, however, will focus on computing the variance and addressing this tension is a goal of a future work.

Another local observation that could be affected by the cosmic variance are observations pertaining to the missing baryon problem. See [6] for a 2012 review of the missing baryon problem and how $\approx 30\%$ baryons were considered missing and assumed to be in the warm-hot intergalactic medium. The baryons are considered to be found by [7]. However, they do not consider the cosmic variance although their field-of-view along one side is $5'$, which is quite narrow, and their observations are at $z \approx 0.35$ and $z \approx 0.44$. We also will not address this problem in great detail but instead offer it as motivation to develop the cosmic variance.

Work on addressing and overcoming these errors has so far been limited and sparser than we would like. Many extragalactic surveys made today often neglect to even calculate the cosmic variance. And studies of the cosmic variance are limited by their survey geometries, sampling methods, and semi-analytic methods. See [8] and the references therein for typical methods in computing it.

Apart from computing the Milky Way cosmic variance, we are generally interested in exploring how other observables are affected by large-scale structure for a variable observer location. For example, the variance describes how the matter density varies between samples. So, we would like to compute this matter density as well. Another quantifier of cosmological structure is the 2-point correlation. This could tell us how likely we are to find clumping between one type of particle species and another. Generally we are interested in quantifying the relationship between a statistic and its conditional analog. The conditional analog is the original statistic when measurements are made from biased positions, looking in biased directions, and/or with a questionably small field-of-view (which is hard to avoid in many observations).

In this thesis, we begin by covering the background theoretical cosmology tools necessary to run simulations and describe the cosmic variance. We will then briefly review simulation techniques and carry out a suite of dark matter-only simulations to model our local Universe at redshift $z = 0$. We then consider how we can describe discrete data using our cosmology tools. And more importantly we will consider conditional statistics and develop methods and theory to describe them. Ultimately, we will analyze our data with the conditional statistics using both computationally difficult direct computations as well as theoretical predictions. We will compute the mean, variance, and 2-point correlation of matter density for a variable survey geometry as well as a variable observer location.

2 Theoretical cosmology framework

In this section we present a basic overview of the necessary theoretical cosmology tools to describe a flat, smooth, and isotropic Universe and how initial small inhomogeneities can evolve to form the nonlinear matter structure we see today: our perturbative Universe. We will focus on the theory necessary to run and describe the N-body simulations in section 3. See [9] for a great introductory text and [10] for more detailed descriptions.

2.1 The flat, smooth, and isotropic Universe

In spacetime, an observer's position is given by the 4-position $x^\mu = (x^0, x^i)$, where we work in natural units $c = \hbar = k_B = 1$. $x^0 = \eta$ is conformal time and x^i are spatial components. Latin indices i, j , etc. range from 1 to 3 and Greek indices α, β , etc. range from 0 to 4. The spacetime interval ds^2 can be computed if we know the corresponding metric $g_{\alpha\beta}$ through

$$ds^2 = g_{\alpha\beta} dx^\alpha dx^\beta, \quad (2.1)$$

where we are using Einstein notation. The Friedmann–Lemaître–Robertson–Walker (**FLRW**) **metric** describes a spatially homogeneous and isotropic universe. If we additionally assume no curvature, then the FLRW metric in Cartesian with a $(-, +, +, +)$ signature becomes

$$g_{\alpha\beta} = a^2(\eta) \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = a^2(\eta) \eta_{\alpha\beta}, \quad (2.2)$$

where $a(\eta)$ is the scale factor and $\eta_{\alpha\beta}$ is the Minkowski metric. We introduce the Hubble parameter

$$H = \frac{\dot{a}}{a}, \quad (2.3)$$

to quantify the Universe's expansion rate.

To solve explicitly for the evolution of the scale factor given the FLRW metric (i.e. in a homogeneous and isotropic Universe) we must recall Einstein's equations

$$G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} \mathcal{R} = 8\pi G T_{\mu\nu}, \quad (2.4)$$

where $G_{\mu\nu}$ is the Einstein tensor, $R_{\mu\nu}$ the Ricci tensor, \mathcal{R} the Ricci scalar, G the gravitational constant, and $T_{\mu\nu}$ the energy-momentum tensor. The terms on the left-hand side of the equation can be solved given the FLRW metric. If we consider a perfect fluid then

$$T_{\mu\nu} = (\rho + p) U_\mu U_\nu + p \eta_{\mu\nu} \quad (2.5)$$

where ρ and p are the fluid's density and pressure, respectively, and U is its 4-velocity. The **Friedmann equations** follow (without proof) from the Einstein equations for a perfect fluid in a FLRW Universe

$$H^2 = \frac{8\pi G}{3}\rho - \frac{\kappa}{a^2}, \quad (2.6)$$

$$\frac{\ddot{a}}{a} = \frac{-4\pi G}{3}(\rho + 3p). \quad (2.7)$$

where κ is the curvature (0 for flat, > 0 for closed and < 0 for open). The curvature of the Universe can be determined by comparing the energy density to a critical density

$$\rho_c \equiv \frac{3}{8\pi G}H^2. \quad (2.8)$$

This leads to our definition of a density parameter

$$\Omega \equiv \frac{\rho}{\rho_c}. \quad (2.9)$$

We can also compute a curvature density parameter if we define a curvature density

$$\rho_\kappa \equiv \frac{-3}{8\pi G}\frac{\kappa}{a^2}, \quad (2.10)$$

which lets us rewrite the Friedmann equation

$$1 = \Omega + \Omega_\kappa \quad (2.11)$$

and yields the equivalencies

$$\kappa = 0 \iff \rho = \rho_c \iff \Omega = 1 \iff \text{flat}, \quad (2.12)$$

$$\kappa > 0 \iff \rho > \rho_c \iff \Omega > 1 \iff \text{closed}, \quad (2.13)$$

$$\kappa < 0 \iff \rho < \rho_c \iff \Omega < 1 \iff \text{open}. \quad (2.14)$$

The flat FLRW metric can describe the early Universe before substantial structure has formed or our background Universe today upon which we consider perturbations to describe structure. Furthermore, when describing the current Universe at $z = 0$ we assume a Λ CDM model. Namely, we parameterize the Universe as flat and with vacuum energy Ω_Λ and dark matter Ω_{DM} parameters. If we combine energy conservation with an equation of state

$$p = w\rho \quad (2.15)$$

then it can be shown that the Friedmann equations yield the evolution of our energy components ρ

$$\rho_\Lambda \propto a^0, \quad (2.16)$$

$$\rho_{DM} \propto a^{-3}, \quad (2.17)$$

$$\rho_\kappa \propto a^{-2}, \quad (2.18)$$

$$\rho_r \propto a^{-4}, \quad (2.19)$$

$$(2.20)$$

(where r stands for radiation). These explicit forms for the scale factor allow us to solve for distances with equation 2.1 at $z = 0$. Let us consider a comoving frame where we are stationary with respect to the Universe's expansion. If we consider a timelike path, the FLRW metric tells us that $d\chi = d\eta$, where χ is the comoving distance. If we convert to redshift through $a = 1/(1+z)$ and plug in the explicit energy components of the Λ CDM Universe we can show

$$\chi(z) = \int_0^z \frac{dz'}{H_0 \sqrt{\Omega_{m,0}(1+z')^3 + \Omega_{\Lambda,0}}}, \quad (2.21)$$

where the 0 subscripts denote that these quantities are measured today.

Now we will change perspectives slightly and consider the motion of matter instead of the evolution of spacetime governed by the Friedmann equations. In particular, we treat matter as a pressureless fluid with macroscopic properties ρ , as before, as well as a velocity \mathbf{v} and gravitational potential Φ . We consider the **Euler equations** for mass and momentum conservation as well as the **Poisson equation** to describe this motion. These are, respectively,

$$\dot{\rho} + \nabla \cdot (\rho \mathbf{v}) = 0, \quad (2.22)$$

$$\dot{\mathbf{v}} + \mathbf{v} \cdot \nabla \mathbf{v} = -\nabla \Phi, \quad (2.23)$$

$$\nabla^2 \Phi = 4\pi G \rho. \quad (2.24)$$

However, these equations are given in proper coordinates \mathbf{r} which is not very useful if the role of the scale factor is non-negligible. To convert to the comoving frame (with coordinates \mathbf{x}) we must use the transformation

$$\mathbf{r} = a\mathbf{x}, \quad (2.25)$$

$$\nabla_{\mathbf{r}} = \frac{1}{a} \nabla_{\mathbf{x}}, \quad (2.26)$$

$$\frac{\partial}{\partial t_{\mathbf{r}}} = \frac{\partial}{\partial t_{\mathbf{x}}} - H\mathbf{x} \cdot \nabla_{\mathbf{x}}. \quad (2.27)$$

The subscripts can be thought of as representing the basis/coordinates our operator lives in or the variable that remains constant as the operator/derivative acts.

2.2 Cosmological perturbation theory

To describe the matter structure we see today, we must consider perturbations on top of the flat FLRW metric. We will consider perturbations to our macroscopic fluid quantities from our Euler and Poisson equations. We use the parameterization

$$\rho = \bar{\rho}(1 + \delta), \quad (2.28)$$

$$\mathbf{v} = H\mathbf{r} + \mathbf{u}, \quad (2.29)$$

$$\Phi = \bar{\Phi} + \delta\Phi, \quad (2.30)$$

for matter perturbations, where our observer is in an external reference frame. We have decomposed the velocity in terms of the Hubble flow with radial coordinate \mathbf{r} and peculiar velocity \mathbf{u} . If we transform to the comoving frame (equations 2.25 and 2.26) and substitute this parameterization into the Euler equations, we can derive the **nonlinear fluid equations** describing the pressureless matter perturbation

$$\dot{\delta} + \frac{1}{a} \nabla \cdot [(1 + \delta)\mathbf{u}] = 0, \quad (2.31)$$

$$\dot{\mathbf{u}} + H\mathbf{u} + \frac{1}{a} \mathbf{u} \cdot \nabla \mathbf{u} = -\frac{1}{a} \nabla \delta\Phi, \quad (2.32)$$

$$\nabla^2 \delta\Phi = 4\pi G \bar{\rho} a^2 \delta. \quad (2.33)$$

Note that there is often an implicit time dependence, but we will not consider this until section 2.5.

2.3 Statistics for inhomogeneities

Given a distribution of inhomogeneities characterized by $\delta(\mathbf{x})$ we want to develop useful and simpler descriptions of how matter behaves. For example, we may want to know the width of the distribution of inhomogeneities or if this distribution is skewed.

One such useful characteristic is the **2-point (auto)correlation function**

$$\xi(r) \equiv \langle \delta(\mathbf{x})\delta(\mathbf{x} + \mathbf{r}) \rangle, \quad (2.34)$$

where homogeneity and isotropy of $\delta(\mathbf{x})$ implies that $\xi(\mathbf{r})$ only depends on the magnitude r . The 2-point correlation measures the probability of matter clumping above a random/background distribution.

We can similarly define the (dimensionless) variance in terms of the perturbation or the correlation function

$$\sigma^2 \equiv \langle \delta(\mathbf{x})^2 \rangle = \xi(0) \quad (2.35)$$

In a flat Universe, we can take the Fourier transform (FT) of our perturbations. Our Fourier convention is

$$\delta(\mathbf{x}) = \int \frac{d^3k}{(2\pi)^3} e^{i\mathbf{k}\cdot\mathbf{x}} \delta(\mathbf{k}), \quad (2.36)$$

$$\delta(\mathbf{k}) = \int d^3x e^{-i\mathbf{k}\cdot\mathbf{x}} \delta(\mathbf{x}), \quad (2.37)$$

where \mathbf{k} is a comoving wavevector and $\delta(\mathbf{k})$ is generally complex. We define the FT of the correlation as our matter **power spectrum** (this is the Wiener-Khinchin theorem). From this it follows (see appendix section A) that

$$\langle \delta(\mathbf{k})\delta(\mathbf{k}') \rangle = (2\pi)^3 \delta^3(\mathbf{k} + \mathbf{k}') P(\mathbf{k}), \quad (2.38)$$

$$\langle |\delta(\mathbf{k})|^2 \rangle = (2\pi)^3 \delta^3(0) P(\mathbf{k}), \quad (2.39)$$

$$\sigma^2 = \int d \ln k \frac{k^3 P(k)}{2\pi^2}, \quad (2.40)$$

where $P(k)$, like $\xi(r)$, only depends on the magnitude k because of the isotropy and homogeneity of perturbations and δ^3 is the 3D Dirac delta function. It is clear from the above that the power spectrum has units of k^{-3} which leads to our next definition: the dimensionless power spectrum

$$\Delta^2(k) \equiv \frac{k^3 P(k)}{2\pi^2}. \quad (2.41)$$

$\Delta^2(k)$ gives us the variance of the matter perturbations in a bin $\ln k$. The following relations follow from the FT (and homogeneity and isotropy of δ)

$$\xi(r) = \int d \ln k \Delta^2(k) j_0(kr), \quad (2.42)$$

$$P(k) = 4\pi \int dr \xi(r) r^2 j_0(kr), \quad (2.43)$$

where j_0 is a spherical Bessel function of the first kind.

We will not go into detail here but an important prediction from inflation yields the primordial parametrization

$$\Delta_p^2(k) \equiv A_s \left(\frac{k}{k_0} \right)^{n_s - 1}, \quad (2.44)$$

where A_s is the scalar amplitude, n_s the spectral index, and k_0 the pivot scale. This can be equivalently written in terms of a primordial power spectrum

$$P_p(k) = A_s \left(\frac{k}{k_0} \right)^{n_s - 1} \frac{2\pi^2}{k^3}. \quad (2.45)$$

The primordial perturbations are thus scale-independent.

2.4 Gaussianity and high-order correlators

Gaussian distributions are characterized fully by their mean and variance. Thus, if our perturbations are Gaussian, we would not need higher-order moments to characterize them. And on large enough scales, the assumption of Gaussianity is reasonable. However, we must briefly consider the case where we have highly non-linear perturbations. This is useful to describe matter perturbations on the order 10s of mega-parsecs or smaller.

The 0th-order moment is the mean and we usually define $\langle \rho \rangle \equiv \bar{\rho}$ such that

$$\langle \delta \rangle = 0. \quad (2.46)$$

If we recall the definition of the covariance

$$\text{cov}[\mathbf{x}_1, \mathbf{x}_2] \equiv \langle \delta(\mathbf{x}_1) \delta(\mathbf{x}_2) \rangle - \langle \delta(\mathbf{x}_1) \rangle \langle \delta(\mathbf{x}_2) \rangle \quad (2.47)$$

then we see how 2.46 implies $\xi(\mathbf{x}_1, \mathbf{x}_2) = \text{cov}[\mathbf{x}_1, \mathbf{x}_2]$. In addition to the 2-point correlation function, we can define higher-point correlations as higher-order moments of perturbations. The **n -point correlation** is defined as the ensemble average of a product of n perturbations

$$\xi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n) \equiv \langle \delta(\mathbf{x}_1) \delta(\mathbf{x}_2) \cdots \delta(\mathbf{x}_n) \rangle. \quad (2.48)$$

Note that higher-order correlators do not necessarily describe higher-order perturbations (i.e. highly non-linear perturbations). For Gaussian fields, it is a well-known property that (assuming 2.46) all n -point correlation functions (i.e. central moments) are either 0 or a function of the variance. That is

$$\xi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{2n+1}) = 0, \quad (2.49)$$

$$\xi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{2n}) = \sigma^{2n} (2n - 1)!!. \quad (2.50)$$

The first couple of higher-order moments are

$$\langle \delta(\mathbf{x}_1) \delta(\mathbf{x}_2) \delta(\mathbf{x}_3) \rangle = 0, \quad (2.51)$$

$$\langle \delta(\mathbf{x}_1) \dots \delta(\mathbf{x}_4) \rangle = 3\sigma^4, \quad (2.52)$$

$$\langle \delta(\mathbf{x}_1) \dots \delta(\mathbf{x}_5) \rangle = 0, \quad (2.53)$$

$$\langle \delta(\mathbf{x}_1) \dots \delta(\mathbf{x}_6) \rangle = 15\sigma^6. \quad (2.54)$$

Additionally, an ensemble average is called connected if no function (or combination of functions) we are averaging can be separated out. For example, an average is called connected c if

$$\langle \delta(\mathbf{x}_1) \delta(\mathbf{x}_2) \dots \delta(\mathbf{x}_n) \rangle_c \neq \langle \delta(\mathbf{x}_1) \rangle \langle \delta(\mathbf{x}_2) \dots \delta(\mathbf{x}_n) \rangle \quad (2.55)$$

or for any other combination of the n functions. See [11] and [12] for useful diagrams and related descriptions. Note that connected moments are also called cumulants. This leads us to defining a **reduced** (or irreducible) **n -point correlation** function using the connected part of the average

$$\xi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n)_c \equiv \langle \delta(\mathbf{x}_1) \delta(\mathbf{x}_2) \cdots \delta(\mathbf{x}_n) \rangle_c. \quad (2.56)$$

Quite usefully, Isserlis' theorem gives us the relationship between the ensemble average and its connected part for central moments which can be written recursively

$$\xi(\mathbf{x}_1, \dots, \mathbf{x}_n) = \xi(\mathbf{x}_1, \dots, \mathbf{x}_n)_c + \sum_{k=1}^{n-1} \binom{n-1}{k-1} \xi(\mathbf{x}_1, \dots, \mathbf{x}_{n-k}) \xi(\mathbf{x}_1, \dots, \mathbf{x}_k)_c \quad (2.57)$$

(which is also called Wick's theorem). As an example, we compute the first four correlation functions in terms of their connected parts

$$\langle \delta(\mathbf{x}_1) \rangle = \langle \delta(\mathbf{x}_1) \rangle_c = 0, \quad (2.58)$$

$$\langle \delta(\mathbf{x}_1)\delta(\mathbf{x}_2) \rangle = \langle \delta(\mathbf{x}_1) \rangle_c \langle \delta(\mathbf{x}_2) \rangle_c + \langle \delta(\mathbf{x}_1)\delta(\mathbf{x}_2) \rangle_c = \langle \delta(\mathbf{x}_1)\delta(\mathbf{x}_2) \rangle_c, \quad (2.59)$$

$$\begin{aligned} \langle \delta(\mathbf{x}_1)\delta(\mathbf{x}_2)\delta(\mathbf{x}_3) \rangle &= \langle \delta(\mathbf{x}_1) \rangle_c \langle \delta(\mathbf{x}_2) \rangle_c \langle \delta(\mathbf{x}_3) \rangle_c \\ &+ \langle \delta(\mathbf{x}_1) \rangle_c \langle \delta(\mathbf{x}_2)\delta(\mathbf{x}_3) \rangle_c \text{ (3 terms)} + \langle \delta(\mathbf{x}_1)\delta(\mathbf{x}_2)\delta(\mathbf{x}_3) \rangle_c \\ &= \langle \delta(\mathbf{x}_1)\delta(\mathbf{x}_2)\delta(\mathbf{x}_3) \rangle_c, \end{aligned} \quad (2.60)$$

$$\begin{aligned} \langle \delta(\mathbf{x}_1)\delta(\mathbf{x}_2)\delta(\mathbf{x}_3)\delta(\mathbf{x}_4) \rangle &= \langle \delta(\mathbf{x}_1) \rangle_c \langle \delta(\mathbf{x}_2) \rangle_c \langle \delta(\mathbf{x}_3) \rangle_c \langle \delta(\mathbf{x}_4) \rangle_c \\ &+ \langle \delta(\mathbf{x}_1) \rangle_c \langle \delta(\mathbf{x}_2) \rangle_c \langle \delta(\mathbf{x}_3)\delta(\mathbf{x}_4) \rangle_c \text{ (6 terms)} \\ &+ \langle \delta(\mathbf{x}_1)\delta(\mathbf{x}_2) \rangle_c \langle \delta(\mathbf{x}_3)\delta(\mathbf{x}_4) \rangle_c \text{ (3 terms)} \\ &+ \langle \delta(\mathbf{x}_1) \rangle_c \langle \delta(\mathbf{x}_2)\delta(\mathbf{x}_3)\delta(\mathbf{x}_4) \rangle_c \text{ (4 terms)} + \langle \delta(\mathbf{x}_1)\delta(\mathbf{x}_2)\delta(\mathbf{x}_3)\delta(\mathbf{x}_4) \rangle_c \\ &= \langle \delta(\mathbf{x}_1)\delta(\mathbf{x}_2) \rangle_c \langle \delta(\mathbf{x}_3)\delta(\mathbf{x}_4) \rangle_c \text{ (3 terms)} + \langle \delta(\mathbf{x}_1)\delta(\mathbf{x}_2)\delta(\mathbf{x}_3)\delta(\mathbf{x}_4) \rangle_c. \end{aligned} \quad (2.61)$$

Thus we can see that the first three connected moments are equivalent to their non-connected moments and higher-order moments have extra non-vanishing terms. Additionally, for Gaussian fields, the only non-vanishing connected correlation function is the 2-point connected correlation.

The 3-point correlation is commonly denoted ζ and its FT is called the bispectrum and denoted B .

2.5 Linear evolution of perturbations

Up until this point we have only considered static perturbations. We will now turn our attention to small, linear-order perturbations that evolve in time.

Continuing with our description of matter perturbations, we will denote their evolution by $\delta^{(1)}(\mathbf{x}, t)$. They are comparable to the background density such that $\delta^{(1)}(\mathbf{x}, t) \ll 1$ and fully nonlinear perturbations satisfy $\delta(\mathbf{x}, t) \gg 1$.

At the linear level, the matter perturbations can be separated into an initial spatial component and time evolving **growth rate** $D(t)$. In Fourier space we define the growth rate as

$$\delta^{(1)}(\mathbf{k}, t) = D(t)\delta^{(1)}(\mathbf{k}, t_0), \quad (2.62)$$

where t_0 is some initial time where the perturbations were small and $D(t_0) \equiv 1$. If we combine two of our nonlinear fluid equations 2.31 and 2.32 as well as our growth rate definition, it can be shown to first order

$$\ddot{D} + 2H\dot{D} - 4\pi G\bar{\rho}D = 0. \quad (2.63)$$

Each of our linear modes $\delta^{(1)}(\mathbf{k}, t_0)$ evolves independently according to this evolution equation.

We can parameterize the matter potential perturbation 2.30 by introducing a linear **transfer function** T

$$\delta\Phi(\mathbf{k}, t) = N(\mathbf{k})T(\mathbf{k}, t_i) \frac{D(t)}{a(t)} \frac{a(t_i)}{D(t_i)} \equiv \frac{D(t)}{a} \delta\Phi(\mathbf{k}, t_i), \quad (2.64)$$

where $N(k)$ is used for normalizing the transfer function and specifying initial conditions at some early time t_i . We can combine this transfer function definition with our Poisson equation 2.33 to see how the power spectrum 2.45 evolves (see appendix section B)

$$P(k, t) = T^2(k)D^2(t)P_p(k). \quad (2.65)$$

As we will see in section 3, the transfer function and growth rates are necessary functions we need to describe linear order structure and set up the initial conditions for N-body simulations.

Instead of describing the evolution of a perturbation, we can consider the comoving position of matter particles \mathbf{x} (Eulerian) and how this evolves in time. The Lagrangian formulation describes how an initial (Lagrangian) position \mathbf{x}_i develops under the influence of a displacement field Ψ to give us a final position

$$\mathbf{x}(\mathbf{x}_i, t) = \mathbf{x}_i + \Psi(\mathbf{x}_i, t). \quad (2.66)$$

Lagrangian perturbation theory (LPT) considers a perturbative expansion of the displacement field

$$\Psi(\mathbf{x}_i, t) = \Psi^{(1)}(\mathbf{x}_i, t) + \Psi^{(2)}(\mathbf{x}_i, t) + \dots \quad (2.67)$$

The **Zel'dovich approximation** is the first-order LPT

$$\Psi \approx \Psi^{(1)} \quad (2.68)$$

developed in [13] and reviewed in [14]. Going forward we suppress the first-order subscript. We can solve for Ψ by way of [10]. We start by taking the FT of our linearly growing potential perturbation 2.64 and matter perturbation 2.62

$$\delta\Phi(\mathbf{x}, t) = \frac{D(t)}{a} \delta\Phi(\mathbf{x}, t_i), \quad (2.69)$$

$$\delta(\mathbf{x}, t) = D(t)\delta(\mathbf{x}, t_i), \quad (2.70)$$

and plug them into the perturbative Poisson equation 2.33

$$\nabla^2 \delta\Phi(\mathbf{x}, t_i) = 4\pi G \bar{\rho} a^3 \delta(\mathbf{x}, t_i). \quad (2.71)$$

At linear-order, the Euler equation for perturbations 2.32 becomes

$$\dot{\mathbf{u}} + \frac{\dot{a}}{a} \mathbf{u} = -\frac{1}{a} \nabla \delta\Phi(\mathbf{x}, t). \quad (2.72)$$

In an EdS Universe, 2.69 implies $\delta\Phi$ is independent of a . We can then integrate the Euler equation to yield

$$\mathbf{u} = -\frac{\delta\Phi(\mathbf{x}_i, t)}{a} \int dt \frac{D}{a}. \quad (2.73)$$

Next we use our growth rate evolution equation 2.63 to show

$$\mathbf{u} = -\frac{1}{4\pi G\bar{\rho}a} \frac{\dot{D}}{D} \nabla\delta\Phi(\mathbf{x}, t_i) \quad (2.74)$$

which can be integrated at linear order (using $\mathbf{u} = a\dot{\mathbf{x}}$) to yield the displacement. We combine this with our Lagrangian formalism 2.66 to find the field

$$\Psi(\mathbf{x}_i, t) = \mathbf{x}(\mathbf{x}_i, t) - \mathbf{x}_i = -\frac{D(a)}{4\pi G\bar{\rho}a^3} \nabla\delta\Phi(\mathbf{x}_i, t_i) \quad (2.75)$$

The Zel'dovich approximation, equations 2.72 and 2.75, gives the displacement and peculiar velocities of particles if we know their initial positions. This approximation can be used to describe the large-scale growth of structure, even up to large displacements of $\delta \sim 1$. This approximation however breaks down when shell crossing occurs [4].

Apart from the linear-order matter perturbation evolution presented above, we can consider the **Boltzmann equation** for any species (including, e.g., radiation)

$$\frac{df}{dt} = C[f], \quad (2.76)$$

where f , the distribution function, describes the phase-space of a particle and C , the collision integral, describes how the phase space of this particle evolves under interactions with its environment. The particles not only collide with each and interact directly but they warp the metric which indirectly affects nearby particles. Therefore, we must now consider perturbations to the FLRW metric. We use the conformal Newtonian gauge

$$g_{00}(\mathbf{x}, t) = -1 - 2\Psi(\mathbf{x}, t), \quad (2.77)$$

$$g_{0i}(\mathbf{x}, t) = 0, \quad (2.78)$$

$$g_{ij}(\mathbf{x}, t) = a^2\delta_{ij}[1 + 2\Phi(\mathbf{x}, t)], \quad (2.79)$$

where Ψ and Φ are linear, scalar perturbations. Note that in the case that $\Psi = \Phi = 0$ the above reduces to the FLRW metric for a homogeneous, isotropic Universe. See [9] for the explicit Boltzmann equations to describe how photons, baryons, and dark matter evolve under perturbations to the gravitational field.

To complement the Boltzmann equations, we need to describe how the evolution of the distribution functions of particles affects the underlying gravitational field. To this end, recall Einstein's equations 2.4. The Ricci tensor can be written in terms of Christoffel symbols which can, in turn, be determined by the metric. And the Ricci scalar is the contraction of the Ricci tensor. So the LHS of the Einstein equations is geometrical. Lastly, the energy-momentum tensor is determined by the energy-content of the species of our Universe. It can be shown that combining the Einstein and Boltzmann equations yields a set of equations governing the linear evolution of species in phase-space. See [9], [15], or [16] for explicit descriptions.

2.6 Bias and dark matter halos

With our current understanding of the underlying matter field and the evolution of small perturbations, we now want to determine how this is related to large-scale structure. Since DM only interacts gravitationally, it clumps and forms structure before baryonic matter does. DM only interacts gravitationally, so in the early-Universe, when things were hotter and there was more radiation, DM could clump together and settle more easily than baryons which were kept apart by radiation pressure. Broadly speaking, small perturbations in DM evolve over time according to 2.63. Some regions of space becomes overdense enough to collapse and form halos (we consider the formation of halos in more detail in the following section). Once the Universe cools to allow baryonic matter to form structure as well it will be attracted to the web of DM and DM halos. Most (if not all) galaxies have a massive underlying DM halo upon which luminous matter settles. So there is a clear correlation between DM halos and baryon matter positions and densities. This motivates the idea of using DM halos as a tracer for luminous matter perturbations. And we can additionally approximate the halo distribution as tracing the overall DM distribution.

The tracer between DM halos and DM is what we will consider here. A straightforward, but not too common way of computing it is

$$\delta_{Halos}(\mathbf{x}) = \sum_{n=1}^{\infty} b_n \cdot \delta_{Matter}^n(\mathbf{x}), \quad (2.80)$$

where b_n are the **bias** parameters. For small perturbations, a constant linear bias is a sufficient approximation

$$\delta_{Halos} \approx b \cdot \delta_{DM}. \quad (2.81)$$

The linear bias allows us to easily compute other statistical quantities for halos once we have DM descriptors. For example, it is easy to show from the definitions 2.28, 2.34, and 2.35 that at the linear scale

$$\rho_{Halos}(\mathbf{x}) \approx \bar{\rho} [1 + b \cdot \delta_{DM}(\mathbf{x})], \quad (2.82)$$

$$\sigma_{Halos}^2 \approx b^2 \cdot \sigma_{DM}^2, \quad (2.83)$$

$$\xi_{Halos}(r) \approx b^2 \cdot \xi_{DM}(r). \quad (2.84)$$

And in Fourier-space we have 2.37

$$\delta_{Halos}(\mathbf{k}) \approx b \cdot \int d^3x e^{-i\mathbf{k}\cdot\mathbf{x}} \delta_{DM}(\mathbf{x}) \quad (2.85)$$

which, from 2.39, yields

$$P_{Halos}(k) \approx b^2 \cdot P_{DM}(k). \quad (2.86)$$

Higher-order approximations of the bias are useful, especially on the size-scale we will be considering when analyzing the results of our simulations of section 3. However, we will leave this to a future work.

Although we will not study it in detail because the size scale we consider is usually larger, there exists a rich number of halo density profiles on the scale of the halo virial radius. A classic profile based on N-body simulations is the Navarro–Frenk–White (**NFW**) profile of [17]

$$\frac{\rho(r)}{\rho_c} = \frac{\delta_c}{(r/r_s)(1 + r/r_s)^2}, \quad (2.87)$$

where r is the radius of this spherically symmetric distribution, r_s a scale radius, ρ_c the critical density 2.8, and δ_c a characteristic contrast density. These parameters vary from halo to halo. This profile is easy to manipulate mathematically and commonly used although it does have some issues (e.g. a divergent central density and divergent total mass if no maximal r is chosen).

Alternatively, a more modern fit based on simulation data is given in [18].

2.7 Spherical collapse and the halo mass function

As described in section 2.5, the Zel’dovich approximation can be used to describe how gravitational perturbations evolve and thus lead to collapse and the formation of halos. This approach mentions nothing about the size and shape of halos but instead describes the underlying motion of matter. Alternatively, we can consider a spherical perturbation and ignore the dynamics of matter but use Friedman’s equations instead. This is the **spherical collapse approach** first developed in [19] and [20]. For this we consider a local region of the Universe that is closed and filled with pressureless matter. The evolution of this sphere is described by the Friedmann equation 2.6 which admits a parametric solution in terms of a development angle θ

$$\tilde{a} = \tilde{a}_t \left(\frac{1 - \cos \theta}{2} \right), \quad (2.88)$$

$$t = t_t \left(\frac{\theta - \sin \theta}{\pi} \right), \quad (2.89)$$

$$\theta \propto \eta, \quad \theta \in [0, 2\pi], \quad (2.90)$$

where we use a tilde to differentiate between the local and global frame and the t subscript refers to the turn-around value. Turn-around occurs at the maximal $\theta = \pi$ and collapse is the minimal $\theta = 2\pi$. The overdensity can be solved for in terms of the development angle (see e.g. [21] for more details)

$$\frac{\tilde{\rho}}{\rho} = \frac{9}{2} \frac{(\theta - \sin \theta)^2}{(1 - \cos \theta)^3}. \quad (2.91)$$

The overdensity at virialization is approximately

$$\frac{\tilde{\rho}}{\rho} \approx 178. \quad (2.92)$$

This is often denoted Δ_{178} and the approximation Δ_{200} is often used. This value will prove useful when analyzing our simulation data to identify halos (see section 4.1). If we expand the density perturbation to linear order and extrapolate to today using 2.62 we find the virial density

$$\delta_c \approx 1.686, \quad (2.93)$$

which we call the critical density.

We can now take our critical density value predicted by the spherical collapse model and consider a distribution of densities to make conclusions about the population of halos as a whole. We proceed with **Press-Schechter theory** from [22]. Also see [23] for more details and extensions of this model. We assume our perturbations follow a Gaussian distribution and equations 2.35 and 2.46. Then the probability density function of our perturbations is simply

$$p(\delta, M) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{-\delta^2}{2\sigma^2}\right), \quad (2.94)$$

where $p(\delta, M)$ is a probability of observing mass whose fractional deviation from the ensemble average is $\delta \pm \Delta\delta/2$. Then the probability of matter of mass $> M$ already collapsing (and forming halos) is

$$P(M) = \int_{\delta_c}^{\infty} p(\delta, M) d\delta = \frac{1}{2} \operatorname{erfc}\left(\frac{\delta_c}{\sqrt{2}\sigma}\right), \quad (2.95)$$

where erfc is the complementary error function. $P(M)$ is sometimes referred to as a collapse fraction $f(M)$ in literature. Unfortunately, the collapse fraction is 1/2 in the limit $\sigma \rightarrow \infty$ so we multiply by an ad hoc factor of 2 (we will soon learn why this happened). We can differentiate this probability to find the differential number density of halos in a mass bin $M \pm dm/2$

$$\frac{dn(M)}{dM} \equiv -2 \frac{\bar{\rho}}{M} \frac{dP}{dM} = \sqrt{\frac{2}{\pi}} \frac{\bar{\rho}}{M} \delta_c \frac{d\sigma^{-1}}{dM} \exp\left(\frac{-\delta_c^2}{2\sigma^2}\right). \quad (2.96)$$

which is the **halo mass function**. Since we know the critical density from the spherical collapse model we can solve for $dn(M)/dM$ numerically. We just need σ^2 from observations or simulations. σ^2 can be thought of as a function of r or M if we use a spherical mass smoothing

$$M = \bar{\rho} \frac{4}{3} \pi R^3. \quad (2.97)$$

It is insightful to express the halo mass function in log-space in terms of

$$\nu \equiv \frac{\delta_c}{\sigma(M)}, \quad (2.98)$$

then

$$\frac{dn(M)}{d \ln M} = \sqrt{\frac{2}{\pi}} \frac{\bar{\rho}}{M} \frac{d \ln \sigma^{-1}}{d \ln M} \nu \exp\left(\frac{-\nu^2}{2}\right). \quad (2.99)$$

It is now clear that there is an exponential high mass cut-off M_* when $\nu = 1$ then

$$M_* \approx 10^{13} h^{-1} M_\odot \quad \text{today}, \quad (2.100)$$

and a troublesome low mass divergence

$$\frac{dn}{d \ln M} \propto M^{-1} \quad (2.101)$$

which appears to exceed the number of low mass halo observations (see section 6.1 for more). Press-Schechter formalism does a reasonable job of describing how this initial distribution of particles is not necessary to infer the distribution of masses at a later time. Unlike the Zel'dovich approximation, we consider a mass probability distribution instead of a spatial function.

For a more standard and accurate halo mass functions see [24] and [25]. Additionally, [26] offers a fitting formula using large sets of simulation data.

The extended Press-Schechter formalism can explain the ad hoc factor of 2 from above. We can treat matter fluctuations as undergoing random walks. For a path accumulating matter and approaching the critical density there is a mirror path that is overdense and losing matter. The probabilities are equal, and this supplies the missing 2 factor.

3 N-body simulations

We wish to generate samples of DM particles positions to represent our local Universe today. And ultimately, we will use and extend the equations described in the previous section to describe this data. But, before we can do any of this, we must quickly describe some simulation methods and set up our initial conditions.

For our initial conditions and N-body simulations, we often picked parameters to match those given in standard precision cosmology simulations (see e.g. [27]) to the extent that our computing resources allowed.

Whenever applicable for generating our initial conditions and running our simulations we adopted the Λ CDM model that best fits the Planck 2018 results [28]. We list the used values in table 1.

3.1 Simulation methods

Rather than attempting to solve the Boltzmann-Einstein equations of motion (from section 2.5) exactly it is easier to consider a discrete set of N particles. A simple but usually impractical simulation method is a direct sum. This involves computing the gravitational field exactly resulting in N^2 computations. This method is so slow, however, that using more “approximate” methods will ultimately give us higher resolution in the same amount of run-time.

Mesh methods can be used to reduce the number of necessary computations. This involves taking the spatial domain of the particles and imposing a grid. The end result groups particles into bins with a general goal of using larger bins for low density regions and smaller bins for high density regions (or more generally we use smaller bins where there is a large gradient for some function we want to model like a velocity). Structured grids involve tessellating space into rectangles or a curvilinear structure that is easily stored in arrays. These structures can be efficiently computed and stored, however they are inflexible and insensitive to the relationships among neighbor particles. Alternatively an unstructured grid can be used that tessellates space using an irregular pattern of simple shapes. This grid is harder to store efficiently but offers better resolution to describe the underlying density field. And lastly a hybrid grid that combines the benefits of both structured and unstructured grids can be used. To solve our differential equations using a mesh, we can use a finite-element method (FEM) or similar methods. This involves using our boundary values on the mesh and approximating the Boltzmann-Einstein equations over the cell. These equations are then combined to describe the entire mesh and variational methods are used to try and minimize the error of this approximation.

A simple mesh method is the particle mesh (PM) in which the particles in each mesh cell are assigned a density profile. How this density is assigned varies. The simplest assignment is the cloud-in-cell (CIC) method which results in constant density cubes.

ΛCDM parameters	
H_0 [km s ⁻¹ Mpc ⁻¹]	67.32
$\Omega_{\Lambda,0}$	0.6842
$\Omega_{b,0}$	0.04939
$\Omega_{DM,0}$	0.2664
σ_8	0.8120
n_s	0.96605
k_0 [Mpc ⁻¹]	0.05
N-body parameters	
n_{mesh}	512
L_{box} [h ⁻¹ Mpc]	300
$L_{\text{softening}}$ [h ⁻¹ kpc] ..	16

Table 1: A summary of the parameters used in running our simulations. We assume values from the ΛCDM model that best fits the Planck 2018 results [28].

The potential and thus the forces for the PM cells are then computed and applied to the particles over a time-step to evolve the system. This method is considered outdated and can be partially remedied by the particle-particle-particle-mesh (P³M) method. Additionally, we have **tree methods** which start with a cubic mesh. Cubic cells which have more particles than some threshold are divided into smaller cubic cells. This is recursively repeated until all cells have a number count below the threshold. Ultimately only a small number of neighbor cells interactions are computed to evolve the system in time. Distant cells can also be treated collectively. A popular mesh method in astrophysics is adaptive mesh refinement (AMR). This implements a mesh with a variable bin size to resolve the dynamics of particularly sensitive regions (e.g. see the AMR-based structure formation code RAMSES [29]).

Apart from mesh methods we have **meshfree methods** which are defined in the reference frame of the particles as a whole and describe local interactions between neighbors. Because of the lack of a mesh, these methods are generally harder to store but are more suited to describe problems with complex boundaries. A popular meshfree method in astrophysics is smoothed-particle hydrodynamics (**SPH**) developed in [30] and [31]. Here macroscopic fluid quantities are averaged using window functions (see section 4.4 for a discussion) and the equations of motion follow from the Euler and Poisson equations 2.22-2.24.

The mesh grid methods are often called Eulerian when applied to fluid flows because they are stationary with respect to the fluid’s movement. And meshfree methods are

called Lagrangian for fluid flows because they are specified within the frame of reference of the fluid. The derivative with respect to time in this frame is the Lagrangian or material derivative

$$\frac{Df}{Dt} \equiv \frac{\partial f}{\partial t} + \mathbf{u} \cdot \nabla f, \quad (3.1)$$

where $f = f(\mathbf{x}, t)$ is a tensor, $\mathbf{u} = (\mathbf{x}, t)$ is the flow velocity, and ∇ is the covariant derivative.

Because we are discretizing our gravitational continuum into point particles, we end up with diverging gravitational forces and potentials at the particles' centers. This leads to the introduction of a **softening length** ϵ which modifies the gravitational force on small scales $\lesssim \epsilon$ to avoid divergence. A commonly used model is Plummer softening which modifies the gravitational potential of particles

$$\Phi(r) = -\frac{Gm}{r} \quad \rightarrow \quad \Phi_{Plummer}(r) = -\frac{Gm}{\sqrt{r^2 + \epsilon^2}}. \quad (3.2)$$

However, the error associated with Plummer softening can be further reduced if the softened potentials are instead replaced with windowed density functions (i.e. density kernels of finite extent) [32].

3.2 Initial conditions

Computing our initial conditions involves approximating the matter power spectrum $P(k, t)$ at an early enough redshift that a low order approximation is sufficient. The slightly nonlinear power spectrum is then fed into an N-body solver which generates a particle mesh with the same power spectrum. The N-body code then numerically evolves the particles into the nonlinear regime at $z = 0$ where galaxies and other structures are common.

We begin by generating the linear matter transfer function $T(k, z)$ (we will get to how we do this shortly). This allows us to compute the linear power spectrum according to 2.65. We just have to compute the primordial power spectrum 2.45 with values listed in table 1 and the growth rate, whose initial condition is $D(t_i) = 1$ and evolves linearly according to 2.63. $T(k, z)$ is generated using the Cosmic Linear Anisotropy Solving System (CLASS) [33]. CLASS numerically solves the Einstein-Boltzmann equations which we briefly described in section 2.5. In other words, CLASS takes a set of initial conditions in the early Universes describing matter and radiation and evolves them to linear order to generate large-scale observables (in particular the linear matter transfer function). CLASS has been shown to be accurate along with other Boltzmann solvers to the 0.1% level [34].

At an early enough time, a linear-order approximation of our power spectrum is sufficient to run our N-body solver. However, as will see, running a simulation to $z = 0$, where nonlinear structures such as galaxies are common, requires a better than linear approximation.

We next take our linear transfer function and correct for second-order terms. To this end, we take our linear transfer function and use the 2nd-order LPT initial conditions code (2LPTic) [35]. In section 2.5 we described LPT, in particular the 1st-order LPT Zel'dovich approximation. Very briefly, 2LPT approximates the Lagrangian displacement field 2.67 to second order

$$\Psi(\mathbf{x}_i, t) \approx \Psi^{(1)}(\mathbf{x}_i, t) + \Psi^{(2)}(\mathbf{x}_i, t), \quad (3.3)$$

which, if we can solve for explicitly, allows us to find the displacement and velocity of our particles by way of 2.66. The 2LPTic authors have shown that the matter power spectrum and DM halo mass function are better estimated if we initiate N-body simulations with their 2LPT based $T(k, z = 49)$ instead of a linear-order Zel'Dovich approximation of $T(k, z = 49)$.

3.3 Dark matter-only simulations

We ran DM-only simulations using the N-body solver GADGET-2 [36]. GADGET utilizes an SPH method to represent fluids as well as a tree method for gravitational forces. Since we are running DM-only simulations, GADGET only needs to track the gravitational interactions. This allows us to run larger simulations than we would be able to if we included baryons. As mentioned above, we initialized GADGET at $z = 49$ using a 2nd-order transfer function from 2LPTic which itself used a linear-order transfer function from CLASS. We ran simulations and outputted DM particle data such as positions and velocities at $z = 0$.

We ran 10 simulations, each with $n_{\text{mesh}}^3 = 512^3$ particles in periodic boxes $L_{\text{box}}^3 = (300 \text{ h}^{-1}\text{Mpc})^3$ yielding a mass resolution

$$m_{\text{res}} \approx 1.18 \times 10^{10} \text{h}^{-1} \text{M}_{\odot}. \quad (3.4)$$

We used a force softening of $L_{\text{softening}} = 16 \text{ h}^{-1}\text{kpc}$. To initialize our 10 simulations, we used different random seeds for 2LPTic.

Since we want to model the local Universe, we want galaxy sizes to be comparable to the Milky Way (MW). Based on [37] and the references therein we have chosen the range

$$M_h = 0.5 - 5.0 \times 10^{12} \text{h}^{-1} \text{M}_{\odot} \quad (3.5)$$

to represent MW halo virial masses. We chose m_{res} to be small enough so that MW halos in our simulation were made out of enough particles to offer high-enough resolution on the scale of a typical virial radius. As such our halos are composed of at least 40 particles. We touch on this more in section 4.1. The mean inter-particle distance is $L_{\text{box}}/n_{\text{mesh}} \approx 586 \text{ h}^{-1}\text{kpc}$. We chose $L_{\text{softening}}$ to be a few percentage of this to ensure that our halo virial radii are much larger than our softening length. We discuss halos more in the following section 4.

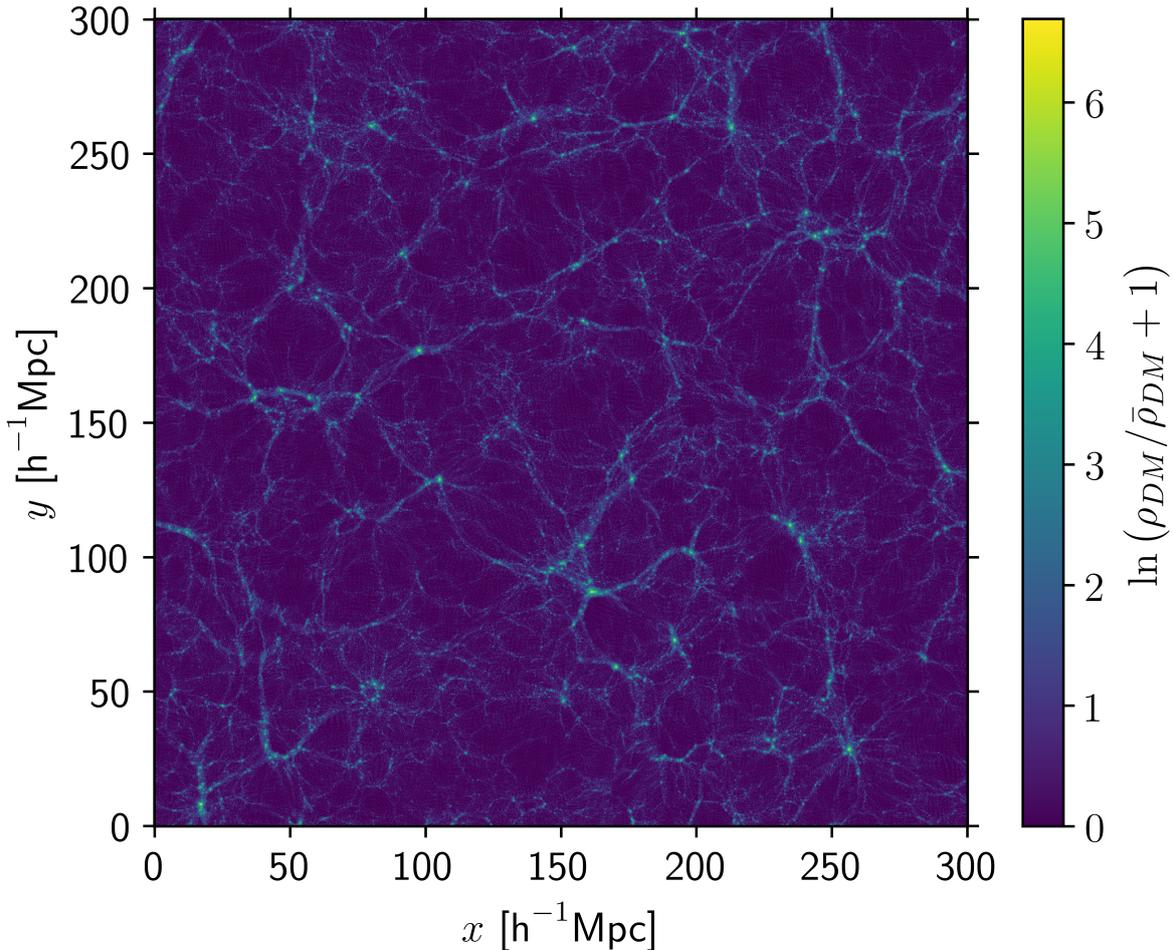


Figure 1: The dark matter density ρ_{DM} from one of our simulations as viewed from above. The density is averaged along the axis into the page over $10h^{-1}\text{Mpc}$ and is measured at $z = 0$. We normalize by the background DM density $\bar{\rho}_{DM}$.

See figure 1 for the output of one of our GADGET simulations. This figure shows the DM density at $z = 0$ and figure 2 provides a zoom-in.

3.4 Simulation approximations

One obvious approximation we are using in our simulations is the lack of baryonic matter. The large-scale structure of our Universe is well described by DM-only simulations where a mapping between DM and baryons can be considered (bias equation 2.80). However, processes such as galaxy, star, and planet formation which heavily depend on baryonic processes would not be well modeled by our simulation data. If we limit our analysis of our simulation data to large enough scales and late enough times we can safely ignore baryonic effects.

Additionally, we only save GADGET snapshots at $z = 0$ in a comoving box of simultaneity. So any statistics we compute will be making the approximation that the

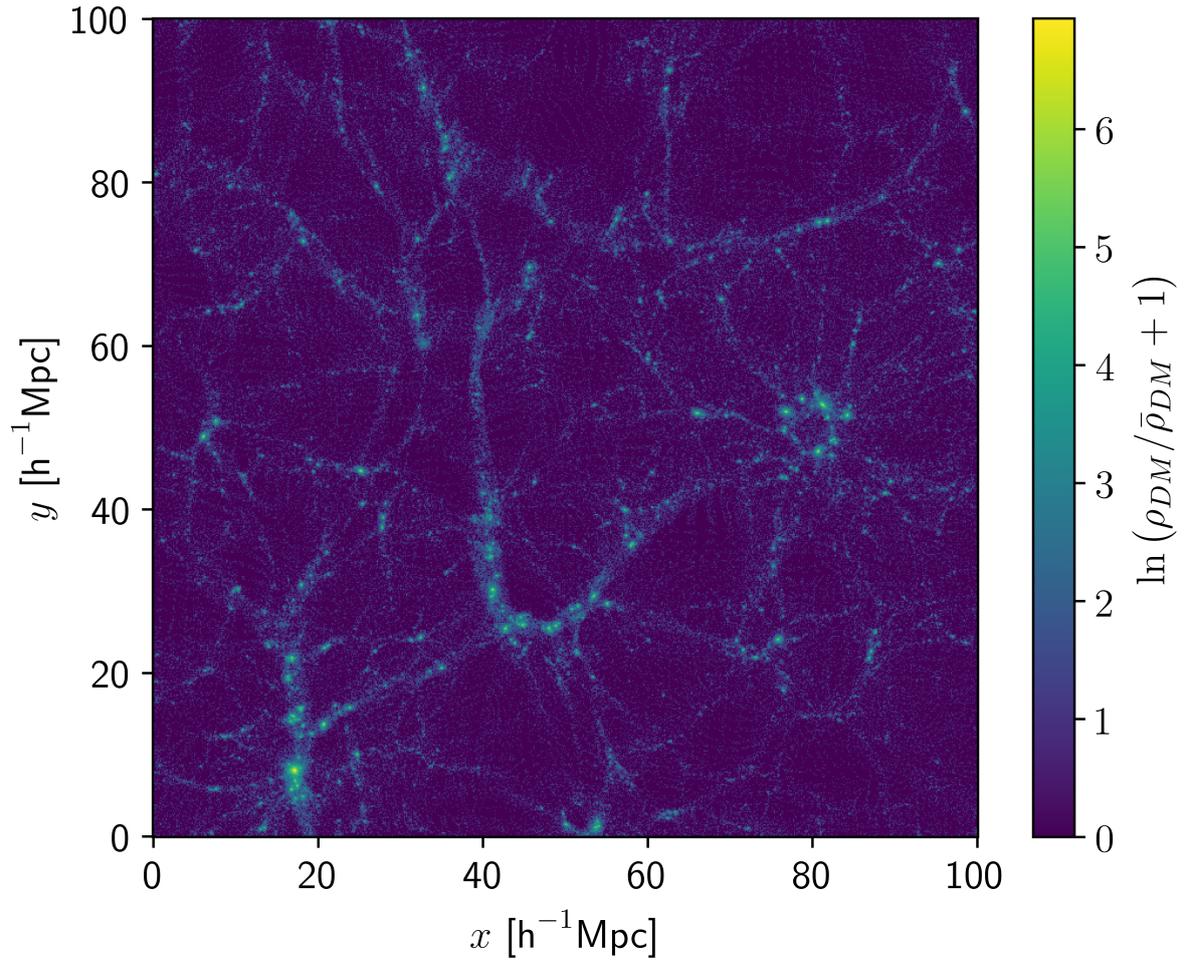


Figure 2: The zoomed-in dark matter density for a simulation. Compare this with the bottom-left corner of figure 1.

time it takes information to travel is not significant. However, our box side-length is only $300h^{-1}\text{Mpc}$ and we only wish to model local observations.

4 Simulation measurements

At this point we have considered cosmology theory and we have DM simulation data. We will now focus on some aspects of dealing with discrete data that will be useful to analyze our data.

As we will see, we typically treat ensemble averages $\langle \dots \rangle$ as a spatial average of some field. The ensemble average is an average over many random realizations of this field. However, we only have access to one random realization of this field. So, the **Ergodic hypothesis** postulates that ensemble averages are equal to spatial averages if we use enough statistically independent realizations within the same field realization. This has been proven for Gaussian fields, however we will assume it also holds at small scales slightly below the threshold of Gaussianity. This will require that we use a large enough sample such that spatial correlations can be successfully neglected.

4.1 Halo finding

As discussed in sections 2.6 and 2.7 there is a quantitative link between our DM particles and halos. We will briefly discuss the halo-finding methods we performed on our discrete data set.

The two foundational methods of halo finding are the spherical-overdensity (SO, i.e. the spherical collapse model we used for Press-Schechter theory) and the friends of friends (**FOF**) **algorithm** introduced in [38]. The SO method involves looking for peaks in the density field upon which a spherical density profile is centered. The halo is assumed to extend from this peak out to a density cut-off (a top-down approach). The FOF algorithm involves searching for all particle pairs whose distance are below a threshold, called the linking length. The groups of linked particles are called halos in this scheme (a bottom-up approach). However, this definition does not allow halos to overlap. So if the “actual” halos in a simulation have any overlap, the FOF algorithm will erroneously declare them a single halo.

The halo finder we use on our data is the Rockstar code developed in [39] which additionally uses velocity information to identify halos to higher precision. This code first uses a rapid variant of the FOF method to identify overdense regions which are then analyzed separately. The overdense regions are then analyzed in 6D position-velocity phase space. A phase-space linking length is used to identify substructure in this region. This substructure-identification is repeated until all substructure is found. Finally halo seeds are placed at each substructure and the nearest particles in phase-space are assigned hierarchically.

When using Rockstar to identify halos, we somewhat arbitrarily chose the definition of halos as a SO whose virial radius is 200 times the background density. This definition ensures that the virial radii of halos is larger than the force softening of our GADGET simulations. On average our MW-like halos 3.5 have a virial radius of $\approx 250h^{-1}\text{kpc}$.

Additionally, we choose to include unbound particles in our halos because they can make up a non-negligible mass fraction of the host halo at $z = 0$. This effect is explored in [40]. We chose our Rockstar softening length to be the same as our GADGET one, and we left the other parameters, e.g. the linking length of 0.28, at their default values.

For an in-depth comparison of halo finding methods see [41]. They show that Rockstar does an excellent job of recovering halo properties and that even the SO and FOF methods alone can determine halo properties to about 10% accuracy.

4.2 Simplifying n-dimensional integrals with probability distributions

Before continuing with our discussion of halos, we would like to develop a method to help solve integrals. Occasionally we are challenged with solving higher-dimensional integrals that offer no clear analytic solution (e.g. equation 5.26). If we can find a relationship between the underlying variables whose probability distribution we know, we can often reduce the dimension of the integral. In more concrete terms, consider an n-dimensional function $f(x_1, \dots, x_n)$. For simplicity we will work in Cartesian coordinates but our discussion can be easily applied to other systems. We impose the limits $x_{1,min} \leq x_1 \leq x_{1,max}$, $\Delta x_1 \equiv x_{1,max} - x_{1,min}$, and similarly for the other variables. Our ultimate goal is to solve some general n-dimensional integral

$$\int_{\Delta x_1} dx_1 \cdots \int_{\Delta x_n} dx_n f(x_1, \dots, x_n). \quad (4.1)$$

If we think of f as a random variable that takes on a value $f(x_1, \dots, x_n)$ for a specified x_1, \dots, x_n , the values occur with probabilities $\mathbb{P}(x_1), \dots, \mathbb{P}(x_n)$ which are all uniform distributions. That is $\mathbb{P}(x_1) = 1/\Delta x_1$ and similarly for the other variables. Then solving our above integral is equivalent to computing the expectation value of f

$$\begin{aligned} \mathbb{E}[f(x_1, \dots, x_n)] &= \frac{\int dx_1 \cdots \int dx_n f(x_1, \dots, x_n)}{\Delta x_1 \dots \Delta x_n} \\ &= \int dx_1 \cdots \int dx_n \mathbb{P}(x_1 \cap \dots \cap x_n) f(x_1, \dots, x_n). \end{aligned} \quad (4.2)$$

Our goal is to reduce the n-dimensional integral of $\mathbb{E}[f(x_1, \dots, x_n)]$ to a lower dimension by looking for relationships among the variables x_1, \dots, x_n whose probability distribution is known. The simplest case is if we know the distribution of the f values without needing to know the values x_1, \dots, x_n , then

$$\mathbb{E}[f] = \int df \mathbb{P}(f) f. \quad (4.3)$$

However, it is more common that we only know how a few underlying variables are related and distributed. Consider some 2-dimensional relation $g(x_1, x_2)$ such that $f(x_1, \dots, x_n) =$

$f(g(x_1, x_2), x_3, \dots, x_n)$. Using this function reduces the dimension of the integral needed by 1

$$\mathbb{E}[f(g, x_3, \dots, x_n)] = \int dg \int dx_3 \cdots \int dx_n \mathbb{P}(g \cap x_3 \cap \dots \cap x_n) f(g, x_3, \dots, x_n). \quad (4.4)$$

It might also be that we have multiple relationships among variables. These relationships can possibly depend on all n variables and even depend on the same variables as other functions. Lastly consider a function $g(x_1, \dots, x_n)$ such that $f(x_1, \dots, x_n) = f(g(x_1, \dots, x_n))$. This expectation value is often called the law of the unconscious statistician

$$\mathbb{E}[f(g)] = \int dg \mathbb{P}(g) f, \quad (4.5)$$

which can be useful if we know the distribution of g and not f .

One commonly used transformation in astrophysics involves the distribution of halo masses in space. Suppose we have some scalar field $f(x, y, z)$ describing the masses m of halos. We can solve for $\mathbb{E}[f]$ if we can directly compute the 3-dimensional integral

$$\frac{\int_{\Delta x} \int_{\Delta y} \int_{\Delta z} f(x, y, z) dx dy dz}{\Delta x \Delta y \Delta z}, \quad (4.6)$$

or if we know the probability distribution of f . Alternatively, recall the halo mass function of section 2.7. With the correct normalization the halo mass function is the probability density function of halo masses $\mathbb{P}(m) \propto n(m)$. Although we might not have an explicit form for $m(x, y, z)$, if we know the relationship $f(m)$, then we can apply equation 4.5

$$\mathbb{E}[f(m)] \propto \int dm n(m) f. \quad (4.7)$$

4.3 The halo model

The 2-point DM correlation function can be well approximated using only a halo density profile, halo mass function, and halo-halo correlation function. The assumptions and formalism presented below are known as the halo model and was first touched upon in [42]. The continuous Fourier-space functions were developed in [43]. See [44] and [45] for detailed descriptions and derivations.

The **halo model** begins by assuming DM halos are spherically symmetric with a density profile $\rho(r|M)$ determined by their mass M and distance from center r . We superimpose a grid on the halos with cell volumes ΔV_i . The cells are small enough such that the number of halos in a cell \mathcal{N}_i (the occupation number) is either 0 or 1. We can compute the density at position \mathbf{x} if we sum the density contribution from all cells

$$\rho(\mathbf{x}) = \sum_i \mathcal{N}_i \rho(\mathbf{x} - \mathbf{x}_i | M_i) \quad (4.8)$$

We solve for the DM-DM (pp) correlation function using this framework and equations 2.28 and 2.34

$$\xi_{pp}(r) = \frac{1}{\bar{\rho}^2} \langle \rho(\mathbf{x})\rho(\mathbf{x} + \mathbf{r}) \rangle - 1, \quad (4.9)$$

$$\begin{aligned} \langle \rho(\mathbf{x})\rho(\mathbf{x} + \mathbf{r}) \rangle &= \left\langle \sum_i \mathcal{N}_i \rho(\mathbf{x} - \mathbf{x}_i | M_i) \sum_j \mathcal{N}_j \rho(\mathbf{x} + \mathbf{r} - \mathbf{x}_j | M_j) \right\rangle \\ &= \sum_i \sum_j \left\langle \mathcal{N}_i \mathcal{N}_j \rho(\mathbf{x} - \mathbf{x}_i | M_i) \rho(\mathbf{x} + \mathbf{r} - \mathbf{x}_j | M_j) \right\rangle. \end{aligned} \quad (4.10)$$

We split the this 2-point statistic into the so-called 1-halo terms ($i = j$) and 2-halo terms ($i \neq j$)

$$\langle \rho(\mathbf{x})\rho(\mathbf{x} + \mathbf{r}) \rangle_{1h} = \sum_i \left\langle \mathcal{N}_i \rho(\mathbf{x} - \mathbf{x}_i | M_i) \rho(\mathbf{x} + \mathbf{r} - \mathbf{x}_i | M_i) \right\rangle, \quad (4.11)$$

$$\langle \rho(\mathbf{x})\rho(\mathbf{x} + \mathbf{r}) \rangle_{2h} = \sum_i \sum_{j \neq i} \left\langle \mathcal{N}_i \mathcal{N}_j \rho(\mathbf{x} - \mathbf{x}_i | M_i) \rho(\mathbf{x} + \mathbf{r} - \mathbf{x}_j | M_j) \right\rangle, \quad (4.12)$$

where we used $\mathcal{N}_i^2 = \mathcal{N}_i$. The 1-halo terms correspond to contributions to the correlation function when both contributions come from the same halo and the 2-halo term describes when the contributions come from separate halos. Instead of taking an ensemble average of a function of mass by integrating over space, we can use the halo mass function as a probability distribution function (see equation 4.7 and the surrounding discussion). We apply this to the 1-halo term

$$\begin{aligned} \langle \rho(\mathbf{x})\rho(\mathbf{x} + \mathbf{r}) \rangle_{1h} &= \sum_i \int dM \frac{dn}{dM} \Delta V_i \rho(\mathbf{x} - \mathbf{x}_i | M) \rho(\mathbf{x} + \mathbf{r} - \mathbf{x}_i | M), \\ &= \int dM \frac{dn}{dM} \int d^3 y \rho(\mathbf{x} - \mathbf{y} | M) \rho(\mathbf{x} + \mathbf{r} - \mathbf{y} | M), \end{aligned} \quad (4.13)$$

where the last term is a convolution integral and we used equation 2.46. And for the 2-halo term

$$\begin{aligned} \langle \rho(\mathbf{x})\rho(\mathbf{x} + \mathbf{r}) \rangle_{2h} &= \sum_i \sum_{j \neq i} \int dM_1 \frac{dn_1}{dM_1} \int dM_2 \frac{dn_2}{dM_2} \Delta V_i \Delta V_j \\ &\quad \times [1 + \xi_{hh}(|\mathbf{x}_i - \mathbf{x}_j|, M_1, M_2)] \rho(\mathbf{x} - \mathbf{x}_i | M_1) \rho(\mathbf{x} + \mathbf{r} - \mathbf{x}_j | M_2), \\ &\approx \bar{\rho}^2 + \int dM_1 \frac{dn_1}{dM_1} b(M_1) \int dM_2 \frac{dn_2}{dM_2} b(M_2) \int d^3 y_1 \int d^3 y_2 \\ &\quad \times \xi_{pp}^{lin}(|\mathbf{y}_1 - \mathbf{y}_2|, M_1, M_2) \rho(\mathbf{x} - \mathbf{y}_1 | M_1) \rho(\mathbf{x} + \mathbf{r} - \mathbf{y}_2 | M_2), \end{aligned} \quad (4.14)$$

which again contains a convolution integral and we used the linear bias equation 2.84.

Because both 1-halo and 2-halo terms contain convolutions, it is often easier to FT transform, compute the 1h and 2h power spectra (which have products now instead of convolutions) and then transform back. If we use the notation

$$\xi_{pp}(r) = \xi_{pp}^{1h}(r) + \xi_{pp}^{2h}(r), \quad (4.15)$$

$$P_{pp}(k) = P_{pp}^{1h}(k) + P_{pp}^{2h}(k), \quad (4.16)$$

then

$$\xi_{pp}^{1h}(r) = \frac{1}{\bar{\rho}^2} \int dM \frac{dn}{dM} \int d^3y \rho(\mathbf{x} - \mathbf{y}|M) \rho(\mathbf{x} + \mathbf{r} - \mathbf{y}|M), \quad (4.17)$$

$$\begin{aligned} \xi_{pp}^{2h}(r) &= \frac{1}{\bar{\rho}^2} \int dM_1 \frac{dn_1}{dM_1} b(M_1) \int dM_2 \frac{dn_2}{dM_2} b(M_2) \int d^3y_1 \int d^3y_2 \\ &\quad \times \xi_{pp}^{lin}(|\mathbf{y}_1 - \mathbf{y}_2|, M_1, M_2) \rho(\mathbf{x} - \mathbf{y}_1|M_1) \rho(\mathbf{x} + \mathbf{r} - \mathbf{y}_2|M_2), \end{aligned} \quad (4.18)$$

$$P_{pp}^{1h}(k) = \frac{1}{\bar{\rho}^2} \int dM \frac{dn}{dM} |\rho(\mathbf{k}|M)|^2, \quad (4.19)$$

$$P_{pp}^{2h}(k) = P_{pp}^{lin}(k) \left[\frac{1}{\bar{\rho}} \int dM \frac{dn}{dM} b(M) \rho(\mathbf{k}|M) \right]^2. \quad (4.20)$$

Thus, we have arrived at the prediction for the DM correlation function using the halo model. We just have to supply a halo density function (e.g. the NFW profile 2.87) and models for halo clustering and abundance (e.g. the Press-Schechter halo mass function 2.99 and models for the linear bias and power spectrum which we can find, as we will see, using our simulation data).

More recent models of the nonlinear power spectrum are based on large cosmological simulations in addition to the halo model. One such example is the **halofit model** [46] which was later revised in [47]. We will not describe these models here, but we use the revised version of the halofit model for the nonlinear power spectrum in our results section 6.

4.4 Window functions

It is worth noting that computing densities for the simulations involves averaging over a finite volume. So, the concept of a density defined at a single point is more of a mathematical concept than of practical use. We introduce what are called window functions to act as a smoothing functions that convert our discrete density functions into continuous ones.

The simplest case when computing densities is to consider a spherical volume inside which particles are included and outside of which they are ignored. This weighting is known as the **spherical top-hat window function**

$$W(x, < r) = \frac{1}{\frac{4}{3}\pi R^3} \begin{cases} 1 & x < r \\ 0 & x \geq r \end{cases}. \quad (4.21)$$

With this definition, the window function results in a smoothed density

$$\rho(\mathbf{x}, < r) \equiv \int d^3x' W(|\mathbf{x} - \mathbf{x}'|, < r) \rho(\mathbf{x}'). \quad (4.22)$$

We will denote statistical functions computed with the Window function using $< r$. With our parameterization 2.28 and since $\int d^3x W(x, < r) = 1$, it is easy to find the smoothed

density perturbation

$$\delta(\mathbf{x}, < r) = \int d^3x' W(|\mathbf{x} - \mathbf{x}'|, < r) \delta(\mathbf{x}'). \quad (4.23)$$

The smoothed mean and perturbation are convolutions. Therefore, their FTs are simple products

$$\rho(\mathbf{k}, < r) = \rho(\mathbf{k}) W(k, < r), \quad (4.24)$$

$$\delta(\mathbf{k}, < r) = \delta(\mathbf{k}) W(k, < r), \quad (4.25)$$

where the FT of the top-hat window function can be written in terms of a spherical Bessel function of the first kind

$$W(k, < r) = 3 \frac{j_1(kr)}{kr}. \quad (4.26)$$

The correlation function can be similarly adapted yielding a convolution. And the smoothed power spectrum is simply

$$P_W(k, < r) = W^2(k, < r) P(k). \quad (4.27)$$

As an aside, if the power spectrum is computed there is a discrete form of the FT (the DFT) to transform to the correlation function. The fast version of this (the FFT) is readily available in coding libraries and is widely used. Generally, when computing a discrete sum in Fourier-space with periodic boundaries, we are constrained to using integer multiples of the fundamental frequency $2\pi/L_{\text{box}}$. Our sampling rate is $\nu = 2\pi n_{\text{mesh}}/L_{\text{box}}$. Sampling effects become significant near the Nyquist frequency $\nu/2$ [48] which is the highest frequency that can be fully reconstructed.

The smoothed power spectrum, by equation 2.40, produces the variance

$$\sigma^2(< r) = \int d \ln k \Delta^2(k) W^2(k, < r). \quad (4.28)$$

The parameter σ_8 is merely $\sigma^2(< r)$ for DM at $r = 8h^{-1}\text{Mpc}$. Note that the variance can also be computed directly but more difficultly from the smoothed density perturbation using its definition 2.35.

4.5 Densities over spherical sub-volumes

To make our measurements more applicable to telescope observations, we will consider densities computed over smaller volumes. Now imagine placing an observer with a telescope at the center of a sphere. For simplicity the telescope's field of view (FOV) is the sphere's solid angle. So if we cut the sphere into hemispheres, the telescope now has two disjoint FOVs each with a solid angle of 2π steradians. In general, we construct smaller volumes by taking the original spherical field FOV and cutting along planes of constant ϕ and θ , the spherical coordinates. The solid angles of the resulting spherical sections

thus represent the now smaller FOV of our observer’s telescope, which can be oriented along any of these disjoint sections.

A FOV constructed with this method, which is explicitly described in appendix section C, results in arbitrarily small solid angles. If we call the algorithmic step of this cutting process n , then at step n we have 2^n unique FOVs all with equal solid angles of $4\pi/2^n$. Thus our observer can orient his telescope in 2^n directions and observe unique FOVs all of size $4\pi/2^n$. If stars were randomly distributed, then the observer would see an equal number of stars in each FOV. However, for large-scale structure with especially small FOVs, the density of matter observed varies a lot and the variance over cosmological scales becomes important. This is the **cosmic variance**, which we will explore in the results section 6.

We depict the cutting algorithm in figure 3 for $n = 0$ to $n = 3$ and we depict the telescope-observer set-up in figure 4 for $n = 8$.

As an example, consider deep pencil surveys. These surveys are only considering a $\lesssim 1$ square degree patch of sky. This is only $\approx 1/41,000^{\text{th}}$ the surface of a sphere. The first algorithmic step resulting in a surface at least this small is $n = 16$. In the results section we only consider up to $n = 11$ because of the computational difficulty of computing higher n . $n = 11$ corresponds to ≈ 20 square degrees.

4.6 Computation of a conditioned density and variance

In previous sections we considered continuous functions and in 4.4 we considered a transformation to discrete data. But in addition to this we might not always be interested in “integrating” over all space. For example, we might want to compute the smoothed density 4.22 over a limited domain. We will come back to why we want to do this shortly after introducing some helpful tools.

We are interested in measuring statistics about random positions, DM particles, and MW-like halo centers 3.5. We will denote these species as

$$R : \text{random positions}, \tag{4.29}$$

$$p : \text{dark matter particles}, \tag{4.30}$$

$$h : \text{MW-like halo centers}. \tag{4.31}$$

Furthermore, we will use the following variables as subscripts

$$1, 2 \in \{R, p, h\}. \tag{4.32}$$

Now suppose X is a quantity we wish to measure like a density or velocity from a data set or simulation. We write

$$X_2^s(\mathbf{x}) \tag{4.33}$$

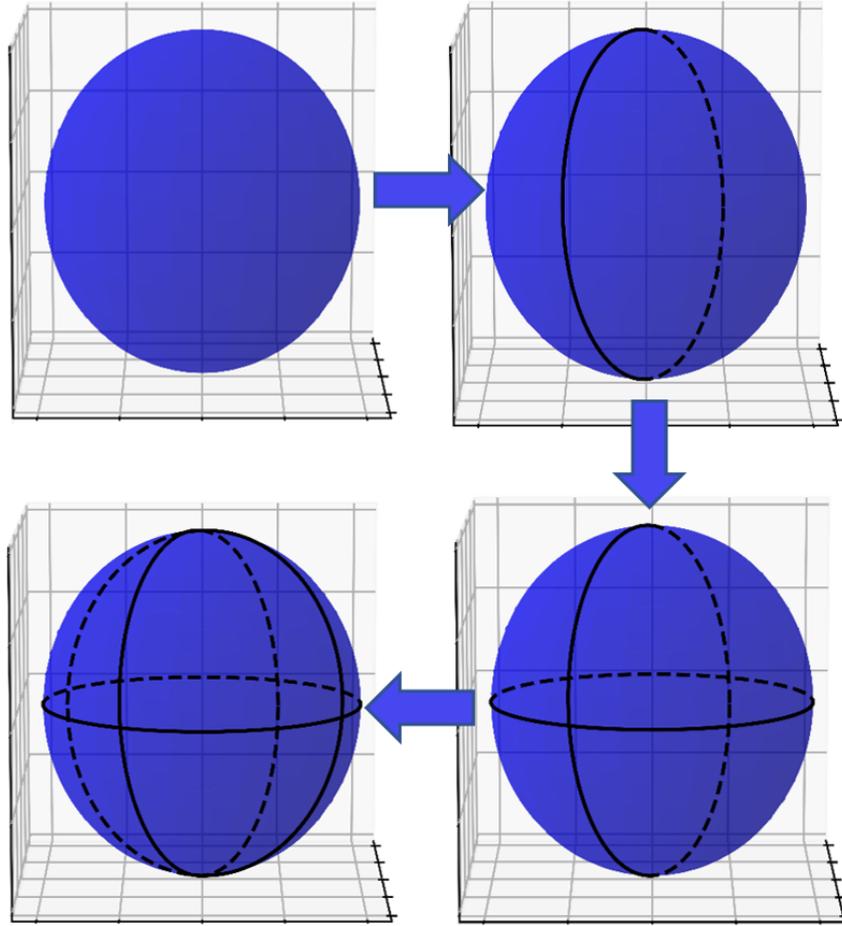


Figure 3: A visualization of the sphere cutting algorithm for the first 4 configurations resulting in 2^0 , 2^1 , 2^2 , and 2^3 spherical subsections. We consider up to the step resulting in 2^{11} subsections when analyzing our data. These are the window functions we average over instead of only using a spherical window.

to denote a measurement in simulation s of a particle species 2 at a position \mathbf{x} . We can average this quantity over all space and compute a mean and variance. But if instead we only average over select regions of space, then we are introducing bias (in the statistical sense and not of section 2.6). To avoid confusion, we sometimes refer to this as a conditional estimator as well as a biased estimator. For example, we might want to measure $X_2^s(\mathbf{x})$ but only on particles of species 1. This leads us to defining a **conditional mean and variance**

$$X_{2|1}^s \equiv \frac{1}{N_1} \sum_{\mathbf{x}_1}^{N_1} X_2^s(\mathbf{x}_1), \quad (4.34)$$

$$(\sigma_{2|1}^s)^2 \equiv \frac{1}{N_1 - 1} \sum_{\mathbf{x}_1}^{N_1} [X_2^s(\mathbf{x}_1) - X_{2|1}^s]^2, \quad (4.35)$$

where we are making N_1 measurements. Note that if N_1 is the number of type 1 particles (the entire population instead of a sample), then the $(N_1 - 1)^{-1}$ factor for the variance

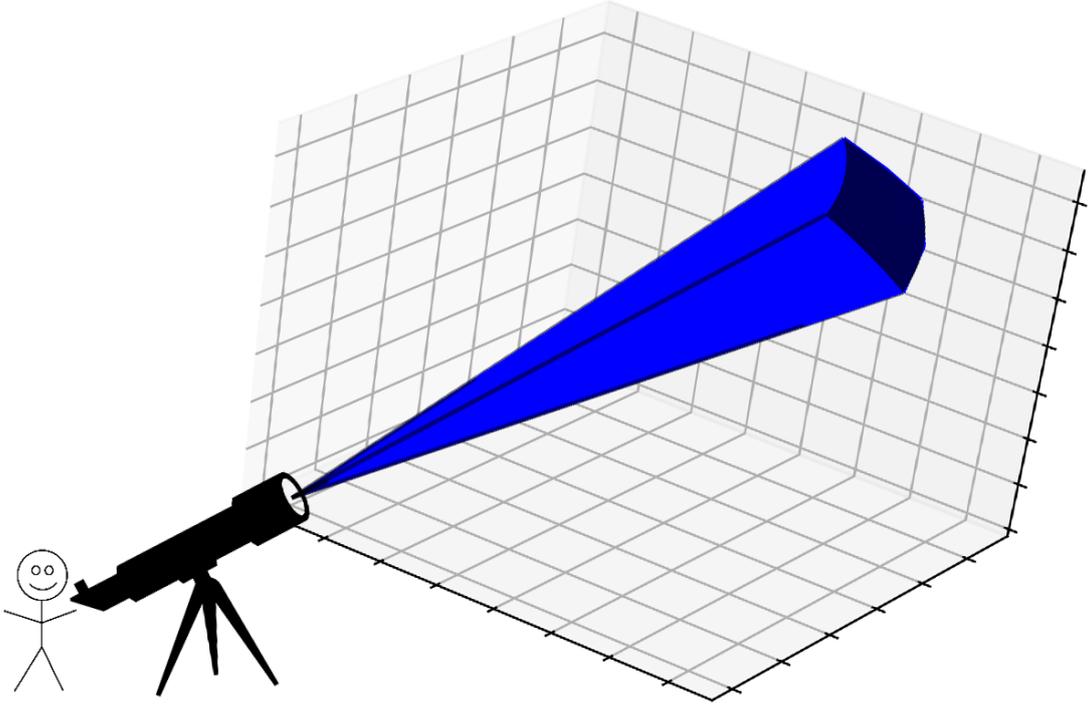


Figure 4: A visualization of our smoothing function which is made to mimic a telescope’s FOV for large n , where the spherical section takes up $4\pi/2^n$ steradians at its end. In this case $n = 8$.

becomes $(N_1)^{-1}$. It is sometimes useful to instead use the expanded form of the variance

$$\begin{aligned} (\sigma_{2|1}^s)^2 &= \frac{1}{N_1 - 1} \sum_{\mathbf{x}_1}^{N_1} \left[(X_2^s(\mathbf{x}_1))^2 + (X_{2|1}^s)^2 - 2X_2^s(\mathbf{x}_1)X_{2|1}^s \right], \\ &= \frac{1}{N_1 - 1} \sum_{\mathbf{x}_1}^{N_1} [X_2^s(\mathbf{x}_1)]^2 - \frac{N_1}{N_1 - 1} (X_{2|1}^s)^2. \end{aligned} \quad (4.36)$$

Also, it is not hard to convert to a **dimensionless variance**

$$(\sigma_{2|1}^s)^2 \rightarrow (\sigma_{2|1}^s)^2 / (X_{2|1}^s)^2. \quad (4.37)$$

If we are able to make our measurements from random positions, that is $1 = R$, then we call our measurements unbiased (or unconditional). However, since we are particularly interested in measuring how matter clusters around MW-like galaxies, we wish to measure quantities such as e.g. the average density of dark matter around MW-like halos $\rho_{p|h}$ or the average density of MW-like halos measured from other MW-like halos $\rho_{h|h}$. This is what was meant at the beginning of the section when we said that restricting the domain over which we “integrate” over can be useful.

Since our mean and variance thus far are simulation dependent, it would be useful to come up with a representative description of these as a whole. We begin by computing

the mean and variance of the simulation mean [4.34](#)

$$X_{2|1} \equiv \frac{1}{N_s} \sum_s^{N_s} X_{2|1}^s, \quad (4.38)$$

$$\sigma_{2|1}^2 \equiv \frac{1}{N_s - 1} \sum_s^{N_s} (X_{2|1}^s - X_{2|1})^2, \quad (4.39)$$

which we will refer to as the **simulation-averaged mean** and (and its error) and N_s is the number of simulations. This variance is also known as the **standard error** and can be approximated using

$$\sigma_{2|1}^2 \approx \frac{(\sigma_{2|1}^s)^2}{N_s}. \quad (4.40)$$

These are useful quantities to gauge the convergence of our measurements. For example, if the number of data points we sample N_1 is large enough, then $\sigma_{2|1}^2$ may be small (conversely the number of simulations we ran could be large enough) and we might conclude that our measurements of $X_{2|1}$ have converged. Next we compute the mean and variance of the simulation variance [4.35](#)

$$\bar{\sigma}_{2|1}^2 \equiv \frac{1}{N_s} \sum_s^{N_s} (\sigma_{2|1}^s)^2, \quad (4.41)$$

$$\sigma_{\sigma^2}^2 \equiv \frac{1}{N_s - 1} \sum_s^{N_s} [(\sigma_{2|1}^s)^2 - \bar{\sigma}_{2|1}^2]^2, \quad (4.42)$$

which we will refer to as the **simulation-averaged variance** (and its error).

If we are interested in computing quantities like the matter density for a discrete data set, we must use a smoothing function. If we measure a species 2 from the biased position 1 and smooth this over a sphere of radius r (we use the spherical window function [4.21](#)), we write

$$X_2^s(\mathbf{x}_1, < r). \quad (4.43)$$

If we switch back to continuous functions, then this is completely analogous to equation [4.22](#)

$$X_2^s(\mathbf{x}_1, < r) = \int d^3x'_1 W(|\mathbf{x}_1 - \mathbf{x}'_1|, < r) X_2^s(\mathbf{x}'_1), \quad (4.44)$$

We can compute statistics on this quantity by simply substituting this into the previous equations. For example, the mean of this windowed, conditional measurement

$$X_{2|1}^s(< r) \quad (4.45)$$

can be computed by substituting $X_2^s(\mathbf{x}_1, < r)$ in place of $X_2^s(\mathbf{x}_1)$ in equation [4.34](#) and now the sum is over spherical windows.

To avoid confusion with other notation we should note that, for example, the cross-correlation function $\xi_{12}(r)$ is computed using an ensemble average and uses a fixed r (instead of being averaged over $< r$ like with a window function). So the correlation function measures fluctuations $\delta_1(\mathbf{x})\delta_2(\mathbf{x} + \mathbf{r})$ where \mathbf{x} is a random position. Relating this to the notation of this section

$$\xi_{12}(r) = \xi_{12|R}(r). \quad (4.46)$$

That is, the correlation is “conditioned” on random positions (i.e. it is not conditioned).

As an example, the background density of our DM simulation data is simply $\bar{\rho}_p = m_{\text{res}} n_{\text{mesh}}^3 / L_{\text{box}}^3 \approx 5.89 \times 10^{10} \text{ M}_{\odot} / \text{h}^{-3} \text{ Mpc}^3$. This is equal to $\bar{\rho}_{p|1}(< r)$ for large-enough scales r such that the species 1 we are conditioning on does not affect our measurement.

We will readily apply the notion of conditional means and variances in section 5 as well as our results section 6.

5 The statistics of number counts

5.1 Conditional counting and correlation

Practically speaking, directly computing the 2-point correlation function 2.34 is difficult. It requires computing density perturbation on a grid as well as the perturbation on another grid a distance r away. For N points in either grid, this would take $\mathcal{O}(N^2)$ time to complete. Instead, we will consider a ‘‘pair-counting’’ estimator. Eventually we will extend the formalism of pair-counting to compute a conditional density and conditional variance to offer descriptions of how the position of an observer affects what they measure.

In this section we will use the notation and definitions of section 4.6, particularly the notation about making measurements from biased positions (note that when we say ‘‘bias’’ here we mean it in the sense of a conditional probability and not the bias between perturbations of section 2.6). We begin by considering a grid of particles of type 1 (e.g. DM particles or MW-like halo centers 3.5 p or h , respectively). The occupation number of each cell at position \mathbf{x}_1 is $\mathcal{N}_1(\mathbf{x}_1)$. We assume that the volume of each cell ΔV_1 is small enough such that $\mathcal{N}_1(\mathbf{x}_1)$ is either 1 or 0. This occupation number can be written as a perturbation to the background density using equation 2.28

$$\mathcal{N}_1 = \bar{n}_1(1 + \delta_1)\Delta V_1, \quad (5.1)$$

where we dropped the position notation for simplicity. The background number density \bar{n}_1 and cell volume cancel each other

$$\bar{n}_1 = \frac{N_1}{V_1}, \quad (5.2)$$

$$\Delta V_1 = \frac{V_1}{N_1}, \quad (5.3)$$

where N_1 is the total number of particles and V_1 is the ensemble volume. Now we introduce a second particle species 2 (possibly the same type as species 1) whose cells are a distance $|\mathbf{r}|$ from species 1 at \mathbf{x}_1 . This species also has a canceling number density and cell volume

$$\bar{n}_2 = \frac{N_2^{shell}(r)}{V_2(r)}, \quad (5.4)$$

$$\Delta V_2 = \frac{V_2(r)}{N_2(r)}, \quad (5.5)$$

where $N_2^{shell}(r)$ is the expected number of particles in a spherical shell of radius r and $V_2(r)$ is the volume of this shell. The product of occupation numbers

$$N_2(\mathbf{x}_1, \mathbf{r}) \equiv \mathcal{N}_1\mathcal{N}_2 = (1 + \delta_1)(1 + \delta_2), \quad (5.6)$$

is either 1 or 0 depending on whether there is both a type 1 particle at \mathbf{x}_1 and a type 2 particle at $\mathbf{x}_1 + \mathbf{r}$. We can use this to count the total number of pairs.

Before continuing it is useful to differentiate between ensemble averages we will use for a general function $f(\mathbf{x}, \mathbf{x} + \mathbf{r})$ which we abbreviate as $f(\mathbf{x}, \mathbf{r})$. The first averages over translations \mathbf{x} for a fixed vector \mathbf{r}

$$f(\mathbf{r}) \equiv \langle f(\mathbf{x}, \mathbf{r}) \rangle_{\mathbf{x}} \equiv \frac{1}{V_1} \int dV_1 f(\mathbf{x}', \mathbf{r}), \quad (5.7)$$

the second averages over translations and rotations for a fixed magnitude r

$$f(r) \equiv \langle f(\mathbf{x}, \mathbf{r}) \rangle_{\mathbf{x}, r} \equiv \frac{1}{V_1 V_2(r)} \iint f(\mathbf{x}', \mathbf{r}') dV_1 dV_2(\mathbf{r}') = \frac{1}{V_2(r)} \int f(\mathbf{r}') dV_2(\mathbf{r}'), \quad (5.8)$$

and the third averages over r as well

$$\begin{aligned} f(< r) &\equiv \langle f(\mathbf{x}, \mathbf{r}) \rangle_{\mathbf{x}, < r} \equiv \frac{1}{V_1 V_2(< r)} \iint f(\mathbf{x}', \mathbf{r}') dV_1 dV_2(< \mathbf{r}') \\ &= \frac{1}{V_2(< r)} \int f(r') V_2(r') dr' = \frac{1}{V_2(< r)} \int f(\mathbf{r}') dV_2(< \mathbf{r}'), \end{aligned} \quad (5.9)$$

where $V_2(< r)$ is a sphere's volume of radius r . Additionally, we can compute a volume average over r without averaging over \mathbf{x} . This should sound familiar: it is the windowed function

$$f(\mathbf{x}, < r) = \frac{1}{V_2(r)} \int f(\mathbf{x}', \mathbf{r}') dV_2(< \mathbf{r}'). \quad (5.10)$$

We can take the rotational ensemble average of $N_2(\mathbf{x}_1, \mathbf{r})$ to find an alternative definition of the 2-point cross-correlation

$$N_{2|1}(r) \equiv \langle N_2(\mathbf{x}_1, \mathbf{r}) \rangle_{\mathbf{x}, r} = \langle 1 + \delta_1 \delta_2 \rangle_{\mathbf{x}, r} \quad (5.11)$$

$$\implies \xi_{12}(r) \equiv N_{2|1}(r) - 1, \quad (5.12)$$

where we used

$$\langle \delta \rangle_{\mathbf{x}, r} = \langle \delta \rangle_{\mathbf{x}, < r} = 0 \quad (5.13)$$

for both species, which follows by using the appropriate definition of the background density 2.46. $N_{2|1}(r) \cdot N_2^{shell}(r)$ is the average number of type 2 particles measured a distance r from the biased location of species 1. This is the continuous analog of our conditional mean equation 4.34. The correlation is a cross-correlation if $1 \neq 2$ and auto-correlation if $1 = 2$. From this definition we see that the correlation function is the excess number of pairs above the expected number of pairs for randomly distributed points (i.e. independently distributed pairs). For example, randomly distributed points are uncorrelated $\xi_{12}(r) = 0$ and $N_{2|1}(r) = 1$. Positively correlated species $\xi_{12}(r) > 0$ increase the likelihood of pair finding and negatively correlated species $\xi_{12}(r) < 0$ decrease this likelihood.

A commonly used notation to describe equation 5.12 (for the auto-correlation) is that of [49]

$$dP = \bar{n}_1^2 dV_1 dV_2 [1 + \xi(r)]. \quad (5.14)$$

This treats the total number of pairs as a probability dP by considering the *average* number of pairs found in cells of volumes dV_1 and dV_2 . When applied to observational data sets this is also sometimes expressed

$$DD(r)\Delta r = RR(r)\Delta r [1 + \xi(r)] \quad (5.15)$$

where DD is the number of data-data pairs and RR are approximate random-random pairs in a spherical shell of depth Δr .

5.2 The mean and variance of a conditional number density

Now consider converting our conditional number counts 5.11 to a number density

$$n_{2|1}(r) \equiv \bar{n}_2 \langle N_2(\mathbf{x}_1, \mathbf{r}) \rangle_{\mathbf{x}_1, \mathbf{r}}. \quad (5.16)$$

We can also compute the conditional number density within a spherical volume

$$n_{2|1}(< r) \equiv \bar{n}_2 \langle N_2(\mathbf{x}_1, \mathbf{r}) \rangle_{\mathbf{x}, < r} = \bar{n}_2 \langle 1 + \delta_1 \delta_2 \rangle_{\mathbf{x}, < r} \quad (5.17)$$

which leads us to defining a volume-averaged 2-point correlation

$$\xi_{12}(r) = \langle \delta_1 \delta_2 \rangle_{\mathbf{x}, r} \quad (\text{from equation 2.34}), \quad (5.18)$$

$$\xi_{12}(< r) \equiv \langle \delta_1 \delta_2 \rangle_{\mathbf{x}, < r}. \quad (5.19)$$

We can solve for the volume-averaged correlation using equation 2.42

$$\begin{aligned} \xi_{12}(< r) &= \frac{4\pi}{V_2(< r)} \int r'^2 \xi_{12}(r') dr' = \frac{3}{r^3} \int d \ln k \Delta_{12}^2(k) \int dr j_0(kr) r^2 \\ &= \int d \ln k \Delta_{12}^2(k) \frac{3j_1(kr)}{kr}. \end{aligned} \quad (5.20)$$

In conclusion, we have

$$n_{2|1}(< r) = \bar{n}_2 [1 + \xi_{12}(< r)], \quad (5.21)$$

which is analogous to the spherical surface equation 5.12.

Along with the conditional number density, we would like to compute a conditional variance $\sigma_{2|1}^2(< r)$. This has the same functional form as equation 4.36 with one caveat

$$\begin{aligned} \sigma_{2|1}^2(< r) &\equiv \frac{N_1}{N_1 - 1} \bar{n}_2^2 \langle \mathcal{N}_1 \mathcal{N}_A \mathcal{N}_B \rangle_{\mathbf{x}, < r} - \frac{N_1}{N_1 - 1} \bar{n}_2^2 \langle \mathcal{N}_1 \mathcal{N}_2 \rangle_{\mathbf{x}, < r}^2 \\ &= \frac{N_1}{N_1 - 1} \bar{n}_2^2 \langle \mathcal{N}_1 \mathcal{N}_A \mathcal{N}_B \rangle_{\mathbf{x}, < r} - \frac{N_1}{N_1 - 1} n_{2|1}^2(< r). \end{aligned} \quad (5.22)$$

The caveat is that in order to compute the number count squared of particles in a sphere we only square \mathcal{N}_2 and not \mathcal{N}_1 . This square results in cross-terms, so we differentiate

between different \mathcal{N}_2 with \mathcal{N}_A and \mathcal{N}_B . So, the average over r is now an average over two r values: r_A and r_B . We can solve for the first term by plugging in our occupation numbers

$$\begin{aligned}\langle \mathcal{N}_1 \mathcal{N}_A \mathcal{N}_B \rangle_{\mathbf{x}, < \mathbf{r}} &= \langle (1 + \delta_1)(1 + \delta_A)(1 + \delta_B) \rangle_{\mathbf{x}, < \mathbf{r}} \\ &= 1 + \langle \delta_1 \delta_A \rangle_{\mathbf{x}, < \mathbf{r}} + \langle \delta_1 \delta_B \rangle_{\mathbf{x}, < \mathbf{r}} + \langle \delta_A \delta_B \rangle_{\mathbf{x}, < \mathbf{r}} + \langle \delta_1 \delta_A \delta_B \rangle_{\mathbf{x}, < \mathbf{r}}.\end{aligned}\quad (5.23)$$

The first two averages are both the volume-averaged correlation function 5.19. We can solve for the next term by first considering a rotational average for fixed r

$$\xi_{AB}(r) \equiv \langle \delta_A \delta_B \rangle_{\mathbf{x}, r}.\quad (5.24)$$

We continue by plugging in the FT

$$\begin{aligned}\xi_{AB}(r) &= \iint \frac{d^3 k}{(2\pi)^3} \frac{d^3 k'}{(2\pi)^3} \langle \delta_2(\mathbf{k}) \delta_2(\mathbf{k}') \rangle = \int \frac{d^3 k}{(2\pi)^3} e^{i\mathbf{k} \cdot (\mathbf{r}_A - \mathbf{r}_B)} P_{22}(\mathbf{k}) \\ &= \int d \ln k \Delta_{22}^2(k) j_0(kr_{AB}).\end{aligned}\quad (5.25)$$

And the average over r is

$$\xi_{AB}(< r) \equiv \langle \delta_A \delta_B \rangle_{\mathbf{x}, < \mathbf{r}} = \int d \ln k \Delta_{22}^2(k) \frac{1}{V_2^2} \iint dV_A dV_B j_0(kr_{AB}).\quad (5.26)$$

The inner integral requires solving a six-dimensional integral. This can be solved analytically if we consider the probability density function f of randomly distributed distances r_{AB} in a sphere (see the discussion of section 4.2). Then

$$\frac{1}{V_2^2} \iint dV_A dV_B j_0(kr_{AB}) = \int dr_{AB} f(r_{AB}) j_0(kr_{AB}).\quad (5.27)$$

Fortunately, most of the work in determining $f(r_{AB})$ has been done already (e.g. see [50] and the references therein). See appendix section D for more details. The probability density is

$$f(r_{AB}) = \frac{6}{r} \left(\frac{r_{AB}}{2r} \right)^2 \left(1 - \frac{r_{AB}}{2r} \right)^2 \left(2 + \frac{r_{AB}}{2r} \right),\quad (5.28)$$

which yields

$$\begin{aligned}\xi_{AB}(< r) &= \int d \ln k \Delta_{22}^2(k) \left[\frac{3j_1(kr)}{kr} \right]^2 = \int d \ln k \Delta_{22}^2(k) W^2(k, < r) \\ \implies \xi_{AB}(< r) &= \sigma_{22}^2(< r),\end{aligned}\quad (5.29)$$

where W is a spherical window function and the last equality uses the windowed variance equation 4.28. This last equality is intuitively clear if we write the variance in terms of

a windowed perturbation

$$\delta_2(\mathbf{x}, < r) \equiv \frac{1}{V_2(< r)} \int \delta_2(\mathbf{x} + \mathbf{r}') dV_2(< \mathbf{r}'), \quad (5.30)$$

$$\begin{aligned} \sigma_{22}^2(< r) &= \langle \delta_2^2(\mathbf{x}, < r) \rangle_{\mathbf{x}} = \frac{1}{V_1 V_2^2(< r)} \int dV_1 \left[\int \delta_2(\mathbf{x} + \mathbf{r}') dV_2(\mathbf{r}') \right]^2 \\ &= \frac{1}{V_1 V_2^2(< r)} \iiint \delta_A \delta_B dV_1 dV_A dV_B = \xi_{AB}(< r). \end{aligned} \quad (5.31)$$

Next, we want to solve for the volume-averaged three-point correlation

$$\zeta_{1AB}(< r) \equiv \langle \delta_1 \delta_A \delta_B \rangle_{\mathbf{x}, < r}, \quad (5.32)$$

where

$$\zeta_{1AB}(r) \equiv \langle \delta_1 \delta_A \delta_B \rangle_{\mathbf{x}, r}. \quad (5.33)$$

Note that although we write $\zeta_{1AB}(r)$ the 3-point correlation is dependent on the three distances r_A , r_B , and r_{AB} . Hence $\zeta_{1AB}(r) = \zeta_{1AB}(r_A, r_B, r_{AB})$. Unlike with the 2-point correlation function's power spectrum, the bispectrum is generally less understood so we will not take the FT. Instead we adopt the hierarchical ansatz [51]

$$\zeta_{1AB}(r) \equiv Q \left[\xi_{1A}(r) \xi_{1B}(r) + \xi_{1A}(r) \xi_{AB}(r) + \xi_{1B}(r) \xi_{AB}(r) \right]. \quad (5.34)$$

We can take the coefficient $Q \approx 1.3$ as a very rough approximation. Alternatively, the behavior of Q using an alternative definition has been studied more recently (e.g. [52])

$$Q_{1AB}(r) \equiv \frac{\zeta_{1AB}(r)}{\xi_{1A}(r) \xi_{1B}(r) + \xi_{1A}(r) \xi_{AB}(r) + \xi_{1B}(r) \xi_{AB}(r)}. \quad (5.35)$$

For now, we proceed with this second definition. The volume-averaged 3-point correlation is thus

$$\zeta_{1AB}(< r) = \left\langle Q_{1AB}(r) \xi_{1A}(r) \xi_{1B}(r) \right\rangle_{\mathbf{x}, < r} + 2 \left\langle Q_{1AB}(r) \xi_{1A}(r) \xi_{AB}(r) \right\rangle_{\mathbf{x}, < r}, \quad (5.36)$$

where we have used the fact that $Q_{1AB}(r)$ is spherically symmetric. We can use a probability density function f again to find

$$\begin{aligned} \zeta_{1AB}(< r) &= \iiint f(r_A, r_B, r_{AB}) Q_{1AB}(r_A, r_B, r_{AB}) \\ &\quad \times \left[\xi_{1A}(r_A) \xi_{1B}(r_B) + 2 \xi_{1A}(r_A) \xi_{AB}(r_{AB}) \right] dr_A dr_B dr_{AB}. \end{aligned} \quad (5.37)$$

Although we know the distributions $f(r_A)$, $f(r_B)$, and $f(r_{AB})$ we do not know their joint distribution. But we have found that this can be approximated quite well using a simple numerical Monte Carlo method. And by using a probability distribution we have managed to decrease our six-dimensional integral to three dimensions.

In conclusion, our (dimensionless) variance on the number density of species 2 conditioned on species 1 is

$$\sigma_{2|1}^2(< r) = \frac{1 + 2\xi_{12}(< r) + \xi_{AB}(< r) + \zeta_{1AB}(< r)}{[1 + \xi_{12}(< r)]^2} - 1, \quad (5.38)$$

where the volume-averaged 2-point correlation functions (equations 5.20 and 5.29) are given in terms of the power spectrum and the volume-averaged 3-point correlation is given by the hierarchy ansatz (equation 5.37) which is limited by the accuracy of our Q model.

6 Results

We now present the results of our simulations which are described in section 3. We are ultimately concerned with computing statistics to describe local observations an astronomer might make on Earth and especially how the cosmic variance comes into play. To this end, our simulations are at $z = 0$ in a box of simultaneity of size $[300h^{-1}\text{Mpc}]^3$. And the particles we analyze are either DM particles or MW-like halo centers 3.5. Our statistics are determined by computing number counts in spheres and spherical subsections which are centered on a particle type. As depicted in section 4.5, a spherical section extending from a MW-like halo that counts DM particles can represent an Earth-based telescope observation assuming that the solid angle is the appropriate size. Our statistics will additionally allow us to analyze the clustering behavior of DM particles and MW-like halos.

In the succeeding sections, we will use the notation of section 4.6 and 5 when describing a measured quantity (e.g. $\xi_{hp}(< r)$ is the volume-averaged 2-point correlation of (MW-like) halo-DM pairs in a sphere of radius r and $\rho_{p|h}(< r)$ is the density of DM particles in spheres of radius r conditioned on MW halo centers). When analyzing our measurements, we will use the definitions of means and variances defined in 4.6.

For each of our 10 simulations, we sample 10,000 particles of p or h whenever applicable. We found that this sample size gave us reasonable convergence without being too computationally expensive. To gauge the convergence of this sample size we computed the standard error on the simulation-averaged mean 4.39. We do not plot this error when it is especially small.

The main results of our paper are the conditional matter density and variance of section 5.2 which we computed directly.

We will not publicly post the code used or the simulation data used to generate our results but please feel to contact us and we will gladly share it.

6.1 Halo mass function

We used the FOF algorithm and Rockstar code to perform our halo-finding. The details of these techniques are described in section 4.1. With our halo-centers we always used a lower cut-off of $5.0 \times 10^{11}h^{-1}M_{\odot}$ to ensure halos are made of at least ≈ 40 particles. We computed the halo mass function by binning a number count per mass per volume. Additionally, we compare our results with the simple Press-Schechter equation 2.99. As discussed earlier, the Press-Schechter model overestimates low-mass halos and underestimates high-mass halos with a turn-over of $M_* \approx 10^{13}h^{-1}M_{\odot}$. We find these characteristics as expected.

We plotted this comparison in 5 using the average of the halo mass function over all simulations and its standard error (equations 4.38 and 4.39). Because of the simplicity of FOF, its halo mass function was computed only as a check on our Rockstar output.

Hence, we only use our Rockstar halos in the succeeding sections. We additionally use an upper cut-off of $5.0 \times 10^{12} h^{-1} M_{\odot}$ to represent MW-like halos.

6.2 2-point correlation and linear bias

We computed the 2-point correlation function using our pair-counting estimator 5.12. We used our simulation data (which gives us DM p positions) as well as the Rockstar output (which gives us MW-like h positions) to compute the auto-correlations $\xi_{pp}(r)$ and $\xi_{hh}(r)$ as well as the cross-correlation $\xi_{hp}(r)$.

We computed the correlations per simulation then took the average and standard error according to equations 4.38 and 4.39. This is shown in figure 6. Although the correlation function in theory extends to $> 100 h^{-1} \text{Mpc}$, we do not have the resolution with our sample size to plot this. At large scales the variance becomes comparable to the correlation size so we would have to increase our sample size.

We can compare our correlation functions with those of [53]. Their simulations however, span a larger spatial range than ours do and as such they analyze both halos and less massive subhalos. We find however that our plots are in general agreement.

Lastly, it is often useful to convert between DM and MW-like halo perturbations δ_p and δ_h , respectively. Thus, we need the linear bias of section 2.6. We will use the correlation function to solve for this using equation 2.84. It is clear that at the linear level we have two ways to compute the bias

$$b_{hh}(r) \equiv \sqrt{\frac{\xi_{hh}(r)}{\xi_{pp}(r)}}, \quad (6.1)$$

$$b_{hp}(r) \equiv \frac{\xi_{hp}(r)}{\xi_{pp}(r)}. \quad (6.2)$$

We compute both linear biases for every simulation. We then average this bias over the simulations and compute its standard error (equations 4.38 and 4.39). We plot the simulation-averaged biases in figure 8.

We expect and found the biases to agree with each other on large, linear scales. The biases disagree on small scales which is unsurprising because they are only linear approximations. A high order approximation is expected to remedy this but is outside the scope of this work.

Since the correlation functions become comparable to their errors at large r the biases' errors also grow. Thus, we can only plot the bias up to $\approx 60 h^{-1} \text{Mpc}$ before its error dominates. In the following sections we need a linear bias model, so we averaged our biases over the range which is both linear and not dominated by error. That is we average $b_{hh}(r)$ and $b_{hp}(r)$ over $10 - 50 h^{-1} \text{Mpc}$ and call these b_{hh} and b_{hp} , respectively.

We can also compare our hh linear bias with that of [53]. Although they use a wider range of halo masses we have found $b_{hh}(r)$ in agreement. Additionally, we can compare

our averaged bias b_{hh} with that of [54]. They plot their bias in units of a nonlinear mass. For our MW-like halo mass range we found that our result is in agreement.

6.3 Power spectrum

Throughout the succeeding sections we make use of the revised halofit model for the nonlinear power spectrum [47]. On large scales, we can use the output of CLASS at $z = 0$ to compute the power spectrum (section 3.2). We can take the FFT of the power spectrum to obtain the 2-point correlation. Alternatively and easily, we can compute integral 2.43 which is what we opt for in later sections.

We plot the nonlinear and linear power spectra in figure 7 as well as the output from our GADGET simulation at $z = 0$ using 2.43. Our simulation-based computation of the power spectrum depends on the correlation function which, as we have seen in figure 6, is only computed over a fixed range of r values and has increasing error on large scales. Because of this, we expect our simulation-based computation of the power spectrum to be approximative. On scales $\gtrsim 1.5k$ nonlinear structure becomes significant and the linear spectra under-predicts perturbations.

6.4 The conditional density's mean and variance (full-sky)

In this and the following section we directly compute the conditional matter density and variance. We compute this for spherical window functions as well as spherical subsections ranging from $4\pi/2^0$ to $4\pi/2^{11}$ steradians (according to section 4.5). These direct computations are the main result we present in this thesis. We also compute the theoretical predictions of section 5.2 for comparison.

For a given simulation s , we start by computing the windowed density in a sphere $\rho_2^s(\mathbf{x}_1, < r)$, where 2 is the particle type we are measuring and species 1 is at \mathbf{x}_1 . This is given by the discrete form of equation 4.44

$$\rho_2^s(\mathbf{x}_1, < r) = \bar{m}_2 \cdot \frac{\# \text{ of type 2 particles within distance } r \text{ of } \mathbf{x}_1}{4\pi r^3/3} \quad (6.3)$$

where we exclude the central particle we are conditioning on (this is to better describe equation 5.6 which is 0 if there is not a pair of particles) and \bar{m}_2 is the average mass of particle 2 (see the discussion around equation 6.8 on why we do not end up needing this term). We then average this spherical density over species 1 according to equation 4.34 to compute the conditional density for a simulation

$$\rho_{2|1}^s(< r) = \frac{1}{N_1} \sum_{\mathbf{x}_1}^{N_1} \rho_2^s(\mathbf{x}_1, < r). \quad (6.4)$$

We also compute the conditional variance of this according to equation 4.35

$$[\sigma_{2|1}^s(< r)]^2 = \frac{1}{N_1 - 1} \sum_{\mathbf{x}_1}^{N_1} \left[\rho_{2|1}^s(\mathbf{x}_1, < r) - \rho_{2|1}^s(< r) \right]^2. \quad (6.5)$$

Next we average this conditional mean and variance according to equations 4.38 and 4.39 to give us the **simulation-averaged conditional density** and **simulation-averaged conditional variance**

$$\rho_{2|1}(< r) = \frac{1}{N_s} \sum_s^{N_s} \rho_{2|1}^s(< r), \quad (6.6)$$

$$\sigma_{2|1}^2(< r) = \frac{1}{N_s} \sum_s^{N_s} [\sigma_{2|1}^s(< r)]^2. \quad (6.7)$$

We plot these simulation-averaged quantities in figures 9 (DM conditioned on DM), 10 (DM conditioned on MW-like halos), and 11 (MW-like halos conditioned on MW-like halos). We normalize figures 9 and 10 by the background DM density $\bar{\rho}_p$. Since DM particles are all the same particle mass, the matter density is a number density times the particle mass. Therefore, the matter density ratio is equivalent to a number density ratio

$$\frac{\rho_{2|1}(< r)}{\bar{\rho}_2} = \frac{n_{2|1}(< r)}{\bar{n}_2}. \quad (6.8)$$

And for figure 11, although all halos are different masses, we assume equation 6.8 holds and normalize with the simulation-averaged number density of MW-like halos. Our two background number densities are

$$\bar{n}_p = 512^3 / [300 \text{ h}^{-1} \text{Mpc}]^3, \quad (6.9)$$

$$\bar{n}_h \approx 144,700 / [300 \text{ h}^{-1} \text{Mpc}]^3. \quad (6.10)$$

Figures 9-11 also display theoretical predictions. The density prediction is given by equations 5.20 and 5.21. Combining these with 6.8 (which we assume holds for halos as well) yields

$$\frac{\rho_{2|1}(< r)}{\bar{\rho}_2} = 1 + \xi_{12}(< r) = 1 + \int d \ln k \Delta_{12}^2(k) \frac{3j_1(kr)}{kr}. \quad (6.11)$$

For $\rho_{p|p}(< r)$ we used the nonlinear power spectrum $P_{pp}(k)$ given by the revised halofit (see the previous section). For the other pairs we used this nonlinear power spectrum in addition to a linear bias (equation 2.86 and the equations 6.1 and 6.2 averaged between $10 - 50 \text{ h}^{-1} \text{Mpc}$)

$$P_{hp}(k) \approx b_{hp} P_{pp}(k), \quad (6.12)$$

$$P_{hh}(k) \approx b_{hh}^2 P_{pp}(k). \quad (6.13)$$

We also offer more approximate models by using the linear power spectrum from our CLASS output. We expect all of our density predictions and data to agree on large scales. And clearly the nonlinear and linear models will disagree on small scales.

For our biased density $\rho_{p|p}$ we found agreement between the nonlinear theory and data on all scales. Our other densities $\rho_{p|h}$ and $\rho_{h|h}$ only agree on large scales, which is not surprising because we use a linear bias model. A more accurate model would either directly compute the power spectra P_{hp} and P_{hh} or use a higher-order bias model that is accurate on smaller scales.

The agreement on large scales can be expressed as

$$\frac{\rho_{2|1}(< r)}{\bar{\rho}_2} \approx 1. \quad (6.14)$$

We also plot the dimensionless form of the simulation-averaged conditional variance 6.7 by itself in figures 12 - 14 along with theoretical predictions. Recall equation 4.37 which gives us the conversion from our previous variance to a dimensionless one

$$\sigma_{2|1}^2(< r) \rightarrow \frac{\sigma_{2|1}^2(< r)}{\rho_{2|1}^2(< r)}. \quad (6.15)$$

The error on this dimensionless simulation-averaged variance is the standard error 4.42

$$\text{Error for dimensionless variance: } \frac{1}{N_s - 1} \sum_s^{N_s} \left[[\sigma_{2|1}^s(< r)]^2 - \sigma_{2|1}^2(< r) \right]^2. \quad (6.16)$$

The theoretical prediction for $\sigma_{2|1}^2(< r)$ is given by equation 5.38 which is fully determined by the power spectrum $P_{12}(k)$, the 3-point correlation coefficient $Q_{1AB}(r_A, r_B, r_{AB})$, and the joint probability distribution $f(r_A, r_B, r_{AB})$. We determine the power spectrum the same way as before with the nonlinear revised halofit model and linear bias (we do not use the linear CLASS power spectrum this time). The linear bias equations are given by 6.12 and 6.13. We use both the approximation $Q_{1AB} \approx 1.3$ and an approximation to the simulation results of [52]. This approximation uses their figure 6 as a piecewise function. We take our r_A, r_B, r_{AB} values and round to the nearest value appearing on their graph. We then use the given functional form for Q_{1AB} . And for r_A, r_B, r_{AB} values outside of this range we take $Q_{1AB} = 0$. This second model is only meant as a proof-of-concept to gauge the ballpark validity of our theoretical variance equation. Lastly, we use a Monte Carlo method to approximate $f(r_A, r_B, r_{AB})$.

We find very approximate agreement between $\sigma_{p|p}^2(< r)$ and our theoretical model for a variable Q_{1AB} . The constant Q model behaves worse. It is evident that the prediction for the variance is sensitive to our model for the volume-averaged 3-point correlation function. For the other pairs $\sigma_{h|p}^2(< r)$ and $\sigma_{h|h}^2(< r)$ we only expect approximate agreement on large scales because we only use a linear bias model. We discuss the variance more in the following section.

6.5 The conditional density's mean and variance (variable solid angle)

We compute the conditional density and its variance again but with spherical subsections instead of only full spheres. Let us consider taking a sphere and cutting it according to section 4.5 such that our resulting solid angles are $4\pi/2^n$ steradians. The volume of our bins are now

$$V(r) = \int_0^r dr' r'^2 \frac{4\pi}{2^n} = \frac{4}{3}\pi r^3 \frac{1}{2^n}. \quad (6.17)$$

Our equations from the previous section retain the same functional form but with a few key differences. The density within a sphere 6.3 becomes the density within the i^{th} angular bin

$$\bar{m}_2 \cdot \frac{\# \text{ of type 2 particles within distance } r \text{ of } \mathbf{x}_1 \text{ and inside angular bin } i}{4\pi r^3 / (3 \cdot 2^n)}, \quad (6.18)$$

where there are $4\pi/2^n$ angular bins per sphere. To compute the average over spheres and its variance (6.4 and 6.5) we not only average the density over \mathbf{x}_1 but the angular bins i as well. And then the equations 6.6 and 6.7 remain the same.

We expect the simulation-averaged density 6.6 to be independent of n . Thus the density in figures 9 - 11 should be the same for any n (assuming each spherical subsection has enough particles for the mean to converge). And in fact, we have confirmed this for $n = 0$ to $n = 11$.

The simulation-averaged variance 6.7 however does depend on n . We plot this in figures 15, 16, and 17 for DM conditioned on DM, DM conditioned on MW-like halos, and MW-like halos conditioned on MW-like halos, respectively. We plot the variance for volumes using $n = 0$ to $n = 11$. As expected, we find that the variance increases as the volume decreases. Note that the curves for $n = 0$ are the same as that of figures 12 - 14.

We find that σ_{ph}^2 decreases on small scales. We believe this is due to us using only MW-like halos of masses 3.5. This restricts the density to a small range when we consider small scales.

The direct computation of the variance underestimates the variance when the spherical volume is close to the box-size. This is because our periodic boundary conditions end up repeating particles and thus artificially deflate the variance.

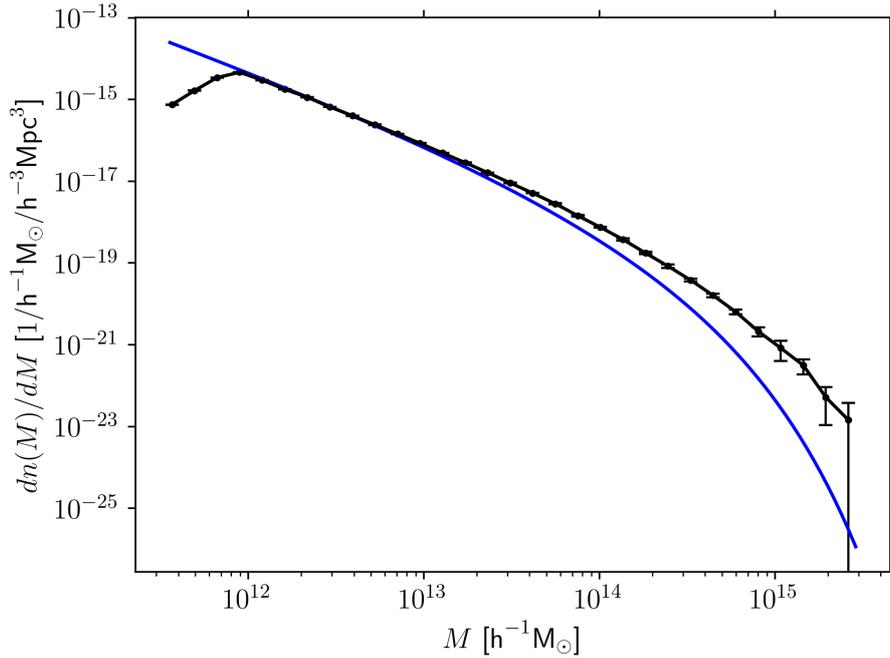


Figure 5: The black curve is our halo mass function for halos above $5.0 \times 10^{11} h^{-1} M_{\odot}$ (we plot the simulation-averaged mean and its standard error). We used Rockstar to identify halos. The blue curve is the Press-Schechter model which offers a rough approximation.

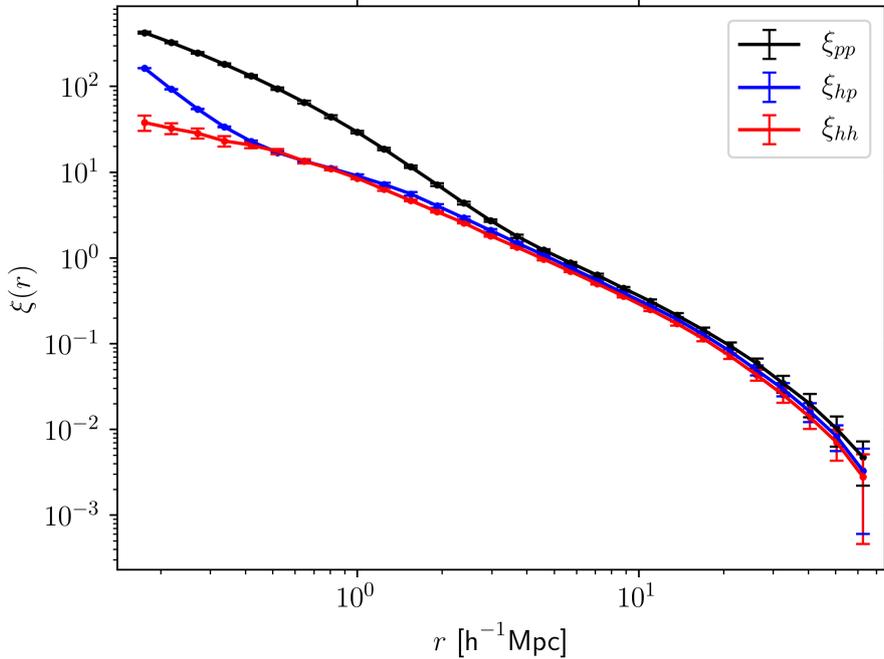


Figure 6: The 2-point cross-correlation (hp) and auto-correlation functions (pp and hh) using our simulation data. p is DM and h is a MW-like halo center. (We plot the simulation-averaged correlations and their standard errors). We computed these correlations using the pair counting estimator 5.12.

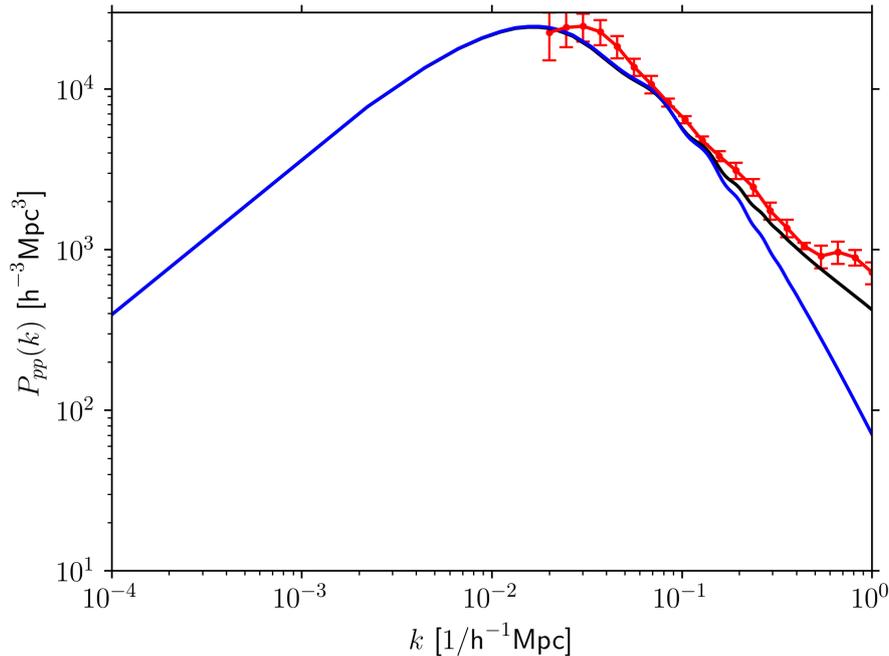


Figure 7: The DM power spectra at $z = 0$. The black curve gives the nonlinear spectra from the revised halofit model [47] and the blue curve is the linear spectra from our CLASS output. The red curve is computed by plugging the DM-DM correlation function of figure 6 (this uses our GADGET simulation output) into 2.43. (We plot the simulation-averaged power spectrum and its standard error.)

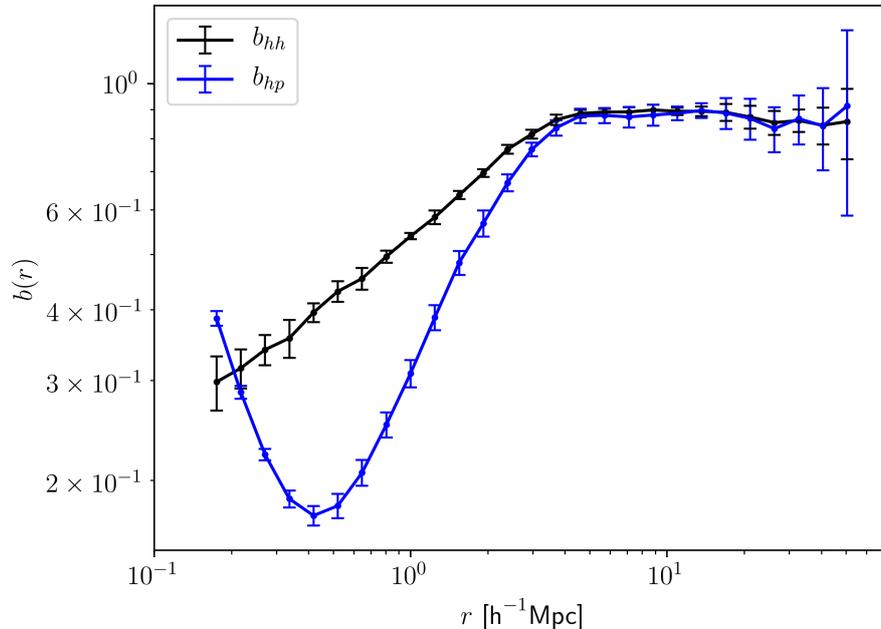


Figure 8: The hh and hp linear biases where p is DM and h is a MW-like halo center. The bias is computed using the correlation functions according to 6.1 and 6.2. (These curves are the simulation-averaged biases and their standard errors).

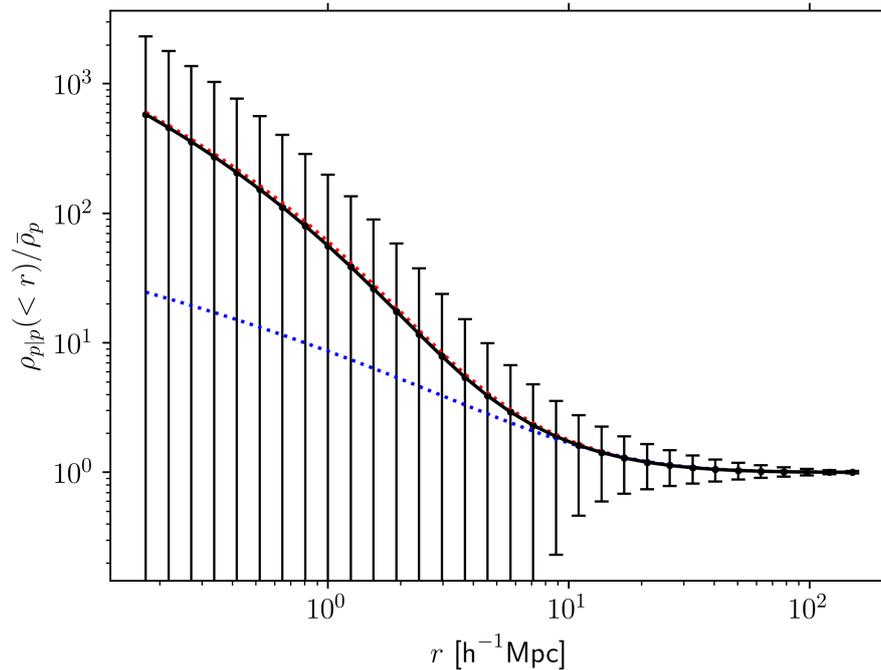


Figure 9: The black curve is our DM (p) density biased on other DM particles in spherical windows of radius r . (We plot the simulation-averaged density 6.6 and simulation-averaged variance 6.7). We offer theoretical fits 6.11 using the nonlinear halofit power spectrum [47] (red curve) and the linear power spectrum from CLASS (blue curve). The normalization is given by 6.9.

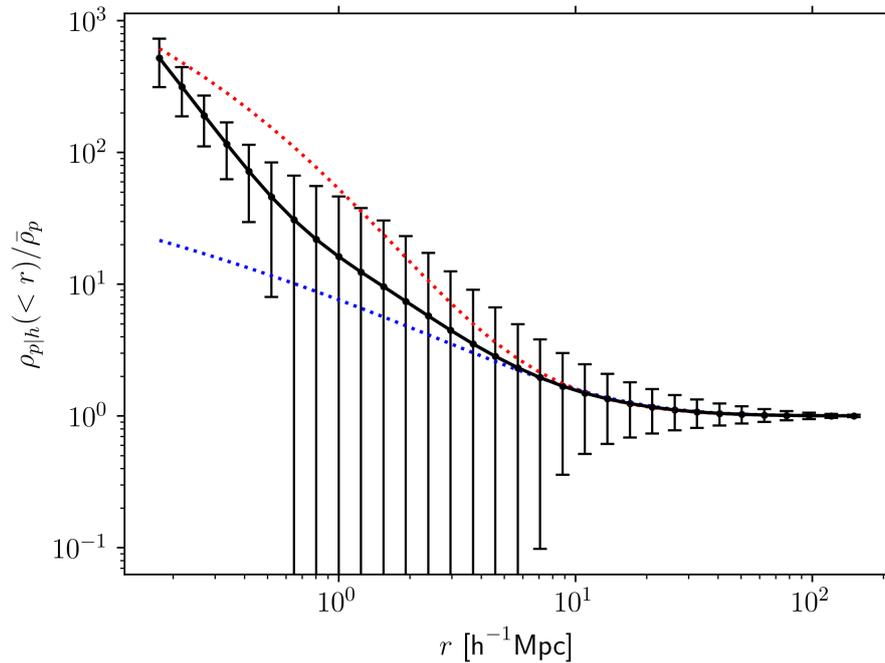


Figure 10: The black curve is our DM (p) density biased on MW-like halo centers (h) in spherical windows of radius r . (We plot the simulation-averaged density 6.6 and simulation-averaged variance 6.7). We offer theoretical fits 6.11 using the nonlinear halofit power spectrum [47] (red curve) and the linear power spectrum from CLASS (blue curve). The normalization is given by 6.9 and we use a linear bias model 6.12.

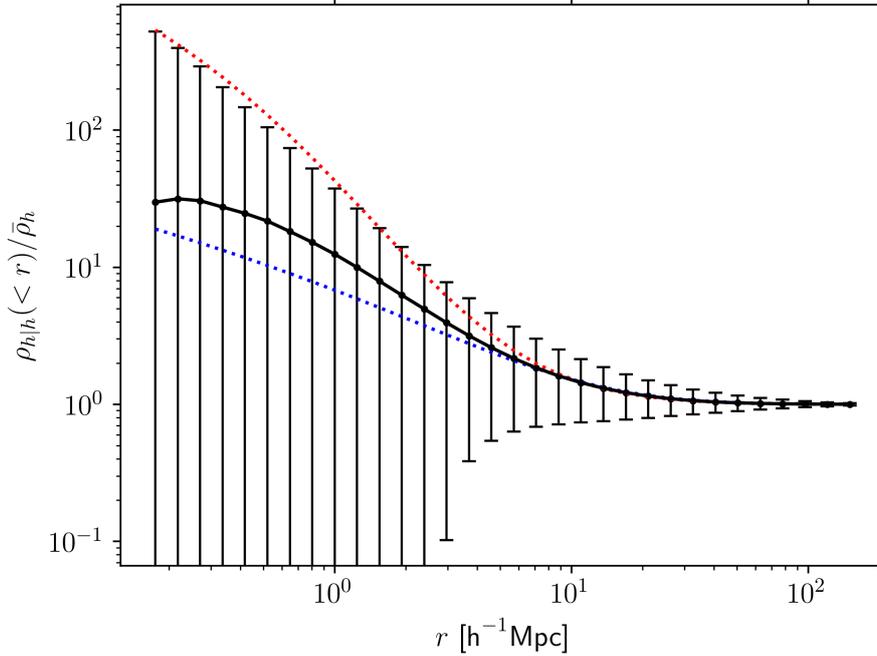


Figure 11: The black curve is our MW-like halo (h) density biased on MW-like halo centers averaged in spherical windows of radius r . (We plot the simulation-averaged density 6.6 and simulation-averaged variance 6.7). We offer theoretical fits 6.11 using the nonlinear halofit power spectrum [47] (red curve) and the linear power spectrum from CLASS (blue curve). The normalization is given by 6.9 and we use a linear bias model 6.13.

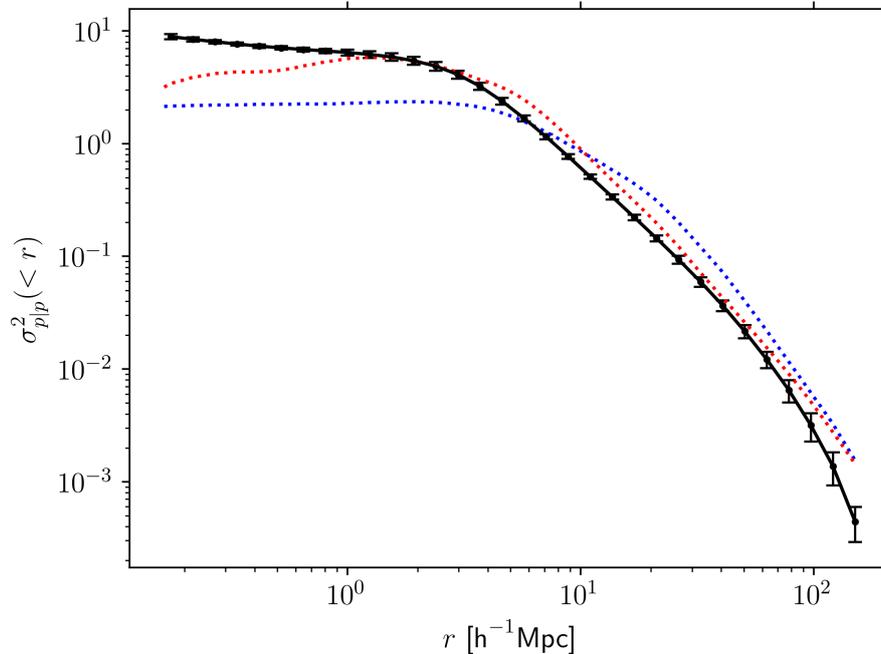


Figure 12: The black curve is our dimensionless DM (p) variance biased on DM particles (the RHS of equation 6.15). (We plot the simulation-averaged variance and its error 6.16). We offer a theoretical prediction 5.38 with a variable Q_{1AB} (red curve) and a more approximate constant $Q_{1AB} = 1.3$ (blue curve).

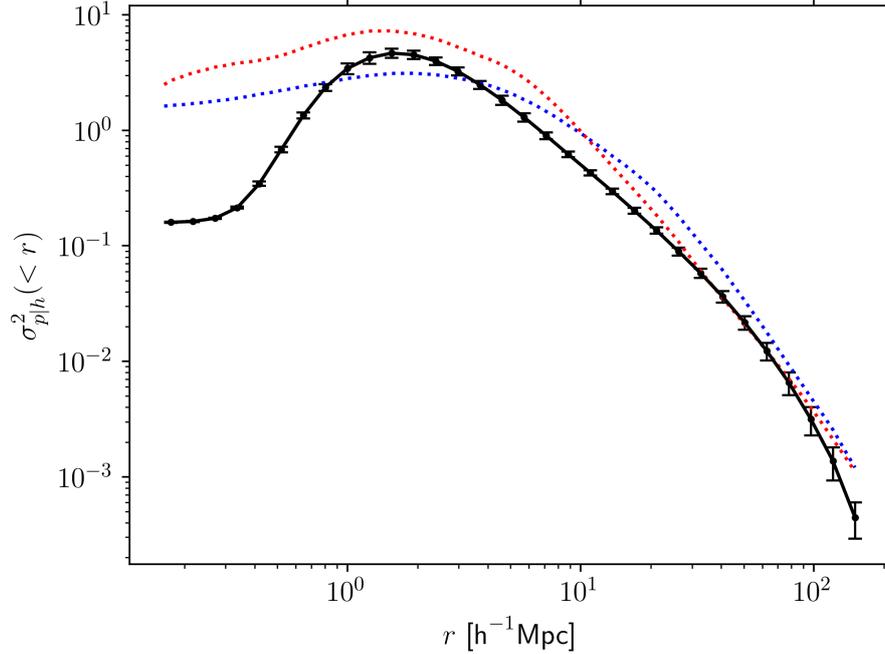


Figure 13: The black curve is our dimensionless DM (p) variance biased on MW-like halo centers (h) (the RHS of equation 6.15). (We plot the simulation-averaged variance and its error 6.16). We offer a theoretical prediction 5.38 with a variable Q_{1AB} (red curve) and a more approximate constant $Q_{1AB} = 1.3$ (blue curve). We use a linear bias model 6.12.

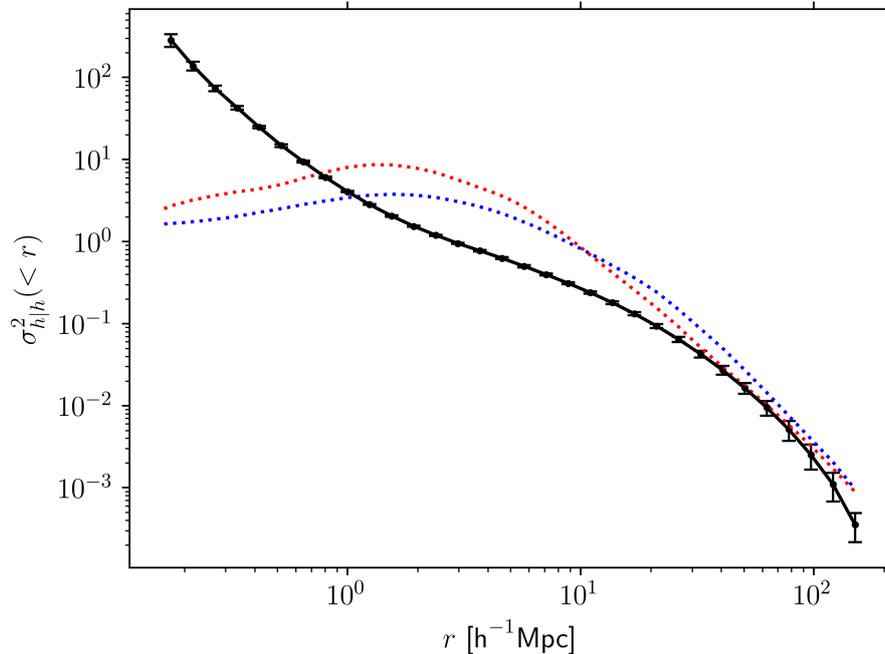


Figure 14: The black curve is our dimensionless MW-like halo (h) variance biased on MW-like halo centers (the RHS of equation 6.15). (We plot the simulation-averaged variance and its error 6.16). We offer a theoretical prediction 5.38 with a variable Q_{1AB} (red curve) and a more approximate constant $Q_{1AB} = 1.3$ (blue curve). We use a linear bias model 6.13.

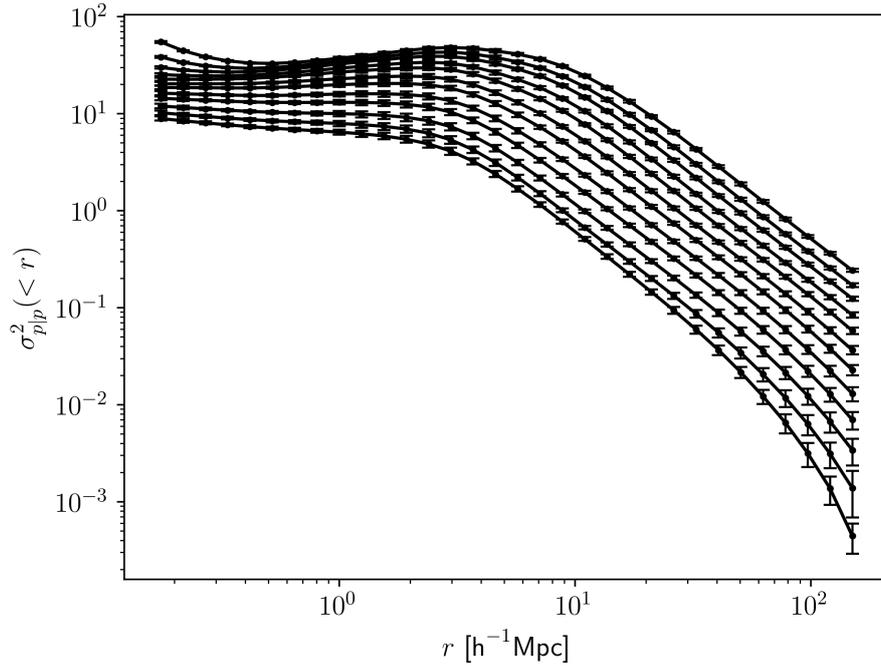


Figure 15: The dimensionless DM (p) variance biased on DM particles (the RHS of equation 6.15) for variable solid angles. (We plot the simulation-averaged variance and its error 6.16). The solid angle is $4\pi/2^n$ where n varies from 0 to 11 (bottom to top).

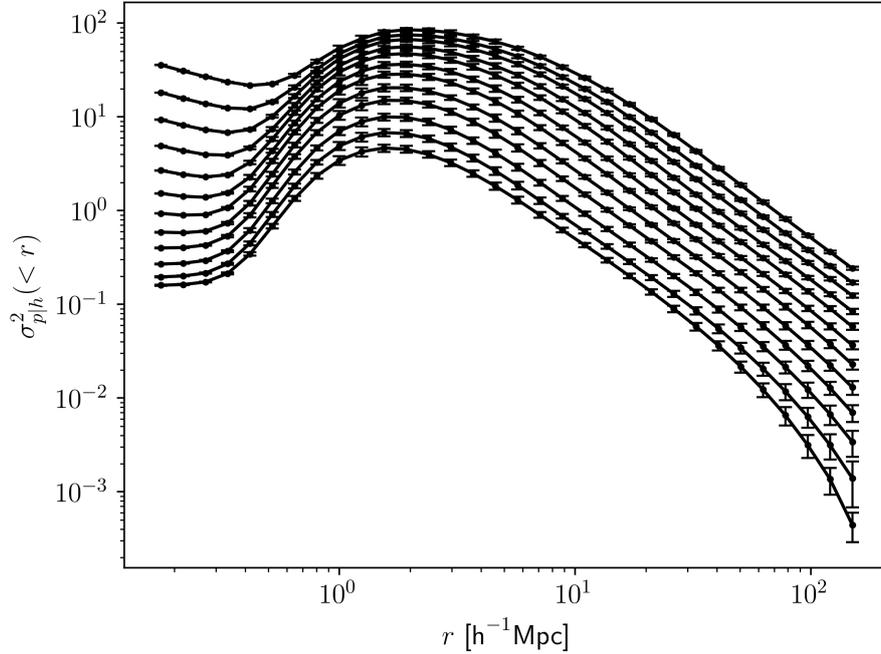


Figure 16: The dimensionless DM (p) variance biased on MW-like halo centers (h) (the RHS of equation 6.15) for variable solid angles. (We plot the simulation-averaged variance and its error 6.16). The solid angle is $4\pi/2^n$ where n varies from 0 to 11 (bottom to top).

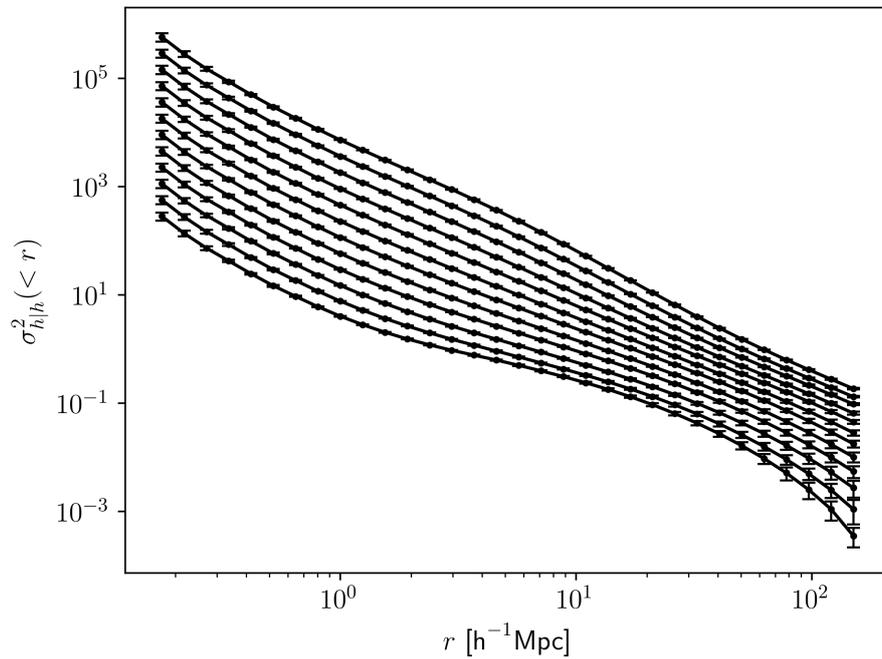


Figure 17: The dimensionless MW-like halo (h) variance biased on MW-like halo centers (the RHS of equation 6.15) for variable solid angles. (We plot the simulation-averaged variance and its error 6.16). The solid angle is $4\pi/2^n$ where n varies from 0 to 11 (bottom to top).

7 Future direction and conclusion

The main results of this work are the direct computation of the conditional mean and variance (also known as a cosmic variance) as presented in figures 9-17. We also present a theoretical model for the conditional variance 5.38.

There is some pre-existing literature describing the cosmic variance using a number of different methods but with different limitations. For example, [55] also offers a theoretical pair-counting model for the variance but they neglect the 3-point correlation function which we have shown highly affects the variance on small scales. Additionally they only use a linear bias and consider a small number of geometries (deep-pencil surveys). However, they offer a fitting formula to extend the variance model to account for redshift that could also be applied to our work. A future goal of ours is to compare our results with theirs and try to build a more general model. Some of this work has been done already in [8] and the references therein.

Often the cosmic variance is not mentioned in published works. For example, in searching for the missing baryons, the authors of [7] do not explicitly compute this value. We will very roughly approximate the cosmic variance here. Their FOV along one side is 5' which is extremely narrow. This motivates us to believe their variance will be quite large. This correspond to a solid angle of $4\pi/2^{22.5}$ steradians! For reference, we only computed the variance for steradians down to $4\pi/2^{11}$ but perhaps we can compare overall volumes. One of their observations is at $z \approx 0.44$ which corresponds to $\chi \approx 1160 \text{ h}^{-1}\text{Mpc}$ (equation 2.21). So the volume of their observation is approximately $1,100\text{h}^{-3}\text{Mpc}^3$. This has the same volume as $4\pi/2^{11}$ steradians over $\chi \approx 81 \text{ h}^{-1}\text{Mpc}$. Their observations is of baryonic matter from the conditioned position of our MW halo. If we assume the linear bias on this scale between DM and baryons is b than we can use our $\sigma_{p|h}^2(< r)$ measurements from figure 16. This corresponds to $\sigma_{p|h}(< r) \approx 0.88\sqrt{b}$. This is not meant to be a rigorous calculation (e.g. we neglect the effect of varying redshift and the effect of different geometries on the variance) but as motivation for further study. Also see [56] for another approximate calculation of the cosmic variance relating to the missing baryon problem.

Quite clearly, our theoretical models are most limited by our linear bias models. Better bias models have long since been studied elsewhere though. For example, see the EFT inspired “peak-background-split” (PBS) approach [24] and [57] for a good description. Our theoretical computation is also limited by our model of the volume-averaged 3-point correlation. We could alternatively have used tried to use the bispectrum (e.g. through the halo model). Additionally, we only approximate the joint distribution $f(r_A, r_B, r_{AB})$ using a Monte Carlo method. We believe it is possible to analytically solve for this distribution (e.g. see [50]) but this was outside the scope of this work.

Our simulation data is limited by its box-size and simultaneity. And for variable solid angles, we would need to perform simulations with higher particle counts to consider

much smaller spherical subsections than $4\pi/2^{11}$. It would be difficult to directly compute the conditional mean and variance over a variable redshift using the methods presented here. Our observer's FOV would have to extend from a biased position outward and span different redshift.

We are also interested in applying and extending conditional measurements using the mathematical framework presented in [58]. There they have found analytic expressions for a mean and correlation conditioned on a local density field. We are interested in finding a quantitative relationship between their and our expressions. It is possible that an analytic relationship between conditioning on halos versus DM (e.g. $\sigma_{p|h}$ versus $\sigma_{p|p}$) could yield a more useful model than our linear bias.

Appendices

A Power spectrum manipulation

We can prove equation 2.38 by considering the following FT

$$\langle \delta(\mathbf{k})\delta(\mathbf{k}') \rangle = \left\langle \int d^3x e^{-i\mathbf{k}\cdot\mathbf{x}} \delta(\mathbf{x}) \int d^3x' e^{-i\mathbf{k}'\cdot\mathbf{x}'} \delta(\mathbf{x}') \right\rangle, \quad (7.1)$$

where the ensemble average can be brought inside the integral and only acts on the perturbations (it only acts on these stochastic variables). We perform a change of variables $\mathbf{x} = \mathbf{x}' + \mathbf{r}$

$$\langle \delta(\mathbf{k})\delta(\mathbf{k}') \rangle = \int d^3r e^{-i\mathbf{k}\cdot\mathbf{r}} \xi(\mathbf{r}) \int d^3x e^{-i(\mathbf{k}+\mathbf{k}')\cdot\mathbf{x}}. \quad (7.2)$$

The first integral is our power spectrum by definition. Lastly, by taking the inverse FT of the Dirac delta function it can be shown that $\delta(k+k') = (2\pi)^{-1} \int dx \exp[-i(k+k')x]$. Hence

$$\int d^3x e^{-i(\mathbf{k}+\mathbf{k}')\cdot\mathbf{x}} = (2\pi)^3 \delta^3(\mathbf{k} + \mathbf{k}'), \quad (7.3)$$

which yields our result. We can continue by plugging in $\mathbf{k} = -\mathbf{k}'$

$$\langle \delta(\mathbf{k})\delta(-\mathbf{k}) \rangle = (2\pi)^3 \delta^3(0) P(\mathbf{k}). \quad (7.4)$$

It can be shown from the properties of FTs that $\delta(\mathbf{x})$ is real iff $\delta(\mathbf{k})$ is Hermitian i.e. we can write the complex conjugate $\delta^*(\mathbf{k}) = \delta(-\mathbf{k})$. Substituting this into the above results in 2.39. And lastly, by definition of the variance

$$\sigma^2 = \xi(0) = \int \frac{d^3k}{(2\pi)^3} P(\mathbf{k}) = \iiint d\theta d\phi dk \frac{k^2 \sin \theta}{(2\pi)^3} P(\mathbf{k}). \quad (7.5)$$

Then, by the homogeneity and isotropy of our matter perturbations, $P(\mathbf{k})$ only depends on the wavevector's magnitude

$$\sigma^2 = \int dk \frac{k^2}{2\pi^2} P(k), \quad (7.6)$$

which yields 2.40.

B Power spectrum and the transfer function

To derive equation 2.65 we start by noting that, by definition of our FT, it can be easily shown that the FT of $\nabla^2 \delta(\mathbf{x})$ is $-k^2 \delta(\mathbf{k})$. With this we know the FT of the nonlinear gravitational perturbation equation 2.33

$$-k^2 \delta\Phi(\mathbf{k}) = 4\pi G \bar{\rho} a^2 \delta(\mathbf{k}). \quad (7.7)$$

Next we note that for an Einstein-de Sitter Universe $D(t) \propto a(t)$ such that $D(t_i) = a(t_i)$. We can combine this, the above equation, and the potential perturbation 2.33 as

$$\delta(k) = -\frac{k^2}{4\pi G \bar{\rho} a(t)^2} N(k) T(k) \frac{D(t)}{a(t)}. \quad (7.8)$$

The matter density scaling $\bar{\rho} \propto a^{-3}$ yields

$$\delta(k) = -\frac{2}{3} \frac{k^2}{H_0^2 \Omega_{m,0}} N(k) T(k) D(t), \quad (7.9)$$

$$\langle |\delta(k)|^2 \rangle = \frac{4}{9} \frac{k^4}{H_0^4 \Omega_{m,0}^2} \langle |N(k)|^2 \rangle T^2(k) D^2(t). \quad (7.10)$$

We can define the initial power spectrum using 2.39

$$P_p \equiv \frac{4}{9} \frac{k^4}{H_0^4 \Omega_{m,0}^2} \langle |N(k)|^2 \rangle \frac{1}{(2\pi)^3}, \quad (7.11)$$

which yields the desired result.

C Indexing $2^n \cdot \mathbf{R}$ spherical sub-volumes

When computing distances between a large number of particles it becomes necessary to bin particles, otherwise we would be severely limited by memory. A particularly useful method for binning (when the bins are uniformly spaced) is integer division (which we'll denote with “\”). This single operation takes a distance measurement and converts it into a bin index.

Since we are also interested in dividing the solid angle into smaller portions, integer division binning will also be useful for our angular coordinates. So, we will now come up with a scheme to evenly divide our angular coordinates. Let (θ, ϕ, r) be our spherical coordinates, where θ is the polar angle, ϕ the azimuthal angle, and r the comoving radial component. We choose the parametrization

$$\mu = -\cos \theta \quad (7.12)$$

then $\mu \in [-1, 1]$ and the differential solid angle is $d\Omega = d\mu d\phi$. In particular, this is useful because cutting the range of μ in equally sized pieces or the range of ϕ (as many times as needed) results in equal solid angles. We can describe this cutting process in terms of algorithmic steps. At step 0 we make no divisions so that ϕ and μ vary over their whole range. At step 1 we take the previous division and cut each ϕ interval in half. At step 2 we take the previous division and cut each μ interval in half. For later steps we just alternate between step 1 and 2. See Table 2 for a generalization of these steps.

The solid angle at step n can be clearly shown to be constant regardless of the number of ϕ and μ interval-halvings chosen

$$\begin{aligned} \int_{-1}^{-1+\Delta_{\mu,n}} \int_0^{\Delta_{\phi,n}} d\mu d\phi &= \int_{-1}^{-1+\Delta_{\mu,n}} \int_{\Delta_{\phi,n}}^{2\Delta_{\phi,n}} d\mu d\phi = \dots = \int_{1-\Delta_{\mu,n}}^1 \int_{2\pi-\Delta_{\phi,n}}^{2\pi} d\mu d\phi \quad (7.13) \\ &= \Delta_{\mu,n} \Delta_{\phi,n} = \frac{4\pi}{2^n}. \end{aligned}$$

Step	μ intervals	ϕ intervals
0	$[-1, 1]$	$[0, 2\pi]$
1	$[-1, 1]$	$[0, \pi), [\pi, 2\pi]$
2	$[-1, 0), [0, 1]$	$[0, \pi), [\pi, 2\pi]$
3	$[-1, 0), [0, 1]$	$[0, \pi/2), [\pi/2, \pi), [\pi, 3\pi/2), [3\pi/2, 2\pi]$
.	.	.
.	.	.
.	.	.
n	$[-1, -1 + \Delta_{\mu,n}), \dots, [1 - \Delta_{\mu,n}, 1]$	$[0, \Delta_{\phi,n}), \dots, [2\pi - \Delta_{\phi,n}, 2\pi]$

Table 2: The angular binning algorithm we adopt for any number of interval-halvings. We define $\Delta_{\mu,n} \equiv 2/[\# \text{ of } \mu \text{ bins at step } n]$ and $\Delta_{\phi,n} \equiv 2\pi/[\# \text{ of } \phi \text{ bins at step } n]$. The number of μ bins at step n is $2^{n\setminus 2}$ and ϕ bins is $2^{(n+1)\setminus 2}$. The number of solid angles that can be formed at step n is 2^n .

This implies that the smallest algorithmic step resulting in $\lesssim 1$ square degree solid angles is $n = 16$.

Let us now consider varying all spherical parameters. We have 2^n angular bins and R equally spaced radial bins. Therefore, after we perform integer division to convert distances into bins (μ_i, ϕ_j, r_k) for all our particles, we have $2^n \cdot R$ possible bins. Lastly, we need to convert from differential bin counts to total counts (to, for example, compute the mean and variance). As we have seen, a deep pencil survey results in $2^{16} \approx 65,000$ unique angular bins (not including possible radial bins)! So, instead of using a triplet (μ_i, ϕ_j, r_k) to index particles into sub-volumes, we will now consider a simpler mixed base system.

For the mapping that follows, our bins should have non-negative values. So, we have to add 1 to μ before binning; the bins μ_i now ranges from 0 to $2^{n\setminus 2} - 1$. And, if we are considering $\log r$, we have to transform the bins to make them non-negative as well. As an aside, it can be useful to use a parameterization for r

$$v = \frac{r^3}{3}. \quad (7.14)$$

This results in the equally sized volume elements $dV = dvd\mu d\phi$. This is particularly useful for uniformly sampling spheres (e.g. see equation 5.37 whose probability distribution is numerically approximated using Monte Carlo sampling). However, it is more convenient for us to use $\log r$ bins which vary better with the rate of change of the number counts. Next we define $b_\mu \equiv 2^{n\setminus 2}$ and refer to it as the base of μ . Similarly, for the other coordinates $b_\phi \equiv 2^{(n+1)\setminus 2}$, $b_r \equiv R$. This allows us to think of (μ_i, ϕ_j, r_k) in a mixed base system where r_i has the lowest position and μ_i has the highest. Then there is a unique

mapping from this triplet into the decimal system with

$$f(\mu_i, \phi_j, r_k) = b_r b_\phi \mu_i + b_r \phi_j + r_k, \quad (7.15)$$

where $f : \mathbb{R}^3 \rightarrow \mathbb{R}$. For example, for $b_r = 2$, $b_\phi = 1$, $b_\mu = 2$, there are $b_\mu b_\phi b_r = 4$ possibilities $f(0, 0, 0) = 0$, $f(0, 0, 1) = 1$, $f(1, 0, 0) = 2$, $f(1, 0, 1) = 3$. This mapping allows easy conversion from our differential bins to cumulative bins because the radial bins that need to be summed together are always sequential. Using our previous example, we have to add the differential number count in bin $f = 0$ to bin $f = 1$ and then, separately, the count in $f = 2$ with $f = 3$. This reduces the number of indices we have to keep track of by two-thirds and allows an easy recursive transformation between steps n and $n - 1$.

D A probability density function transformation

Equation 5.27 is a 6-dimensional integral that only depends on one variable. We found that we can solve for it if we know the probability distribution of randomly distributed distances r_{AB} in a sphere of radius r . The probability density function f is given by (see e.g. [50])

$$f_S(s) = 12s^2(1 - s)^2(2 + s), \quad (7.16)$$

where S is a continuous random variable with $s = r_{AB}/2r$. Let R_{AB} be the continuous random variable for r_{AB} values. Now we will solve for $f_{R_{AB}}(r_{AB})$

$$F_{R_{AB}}(r_{AB}) = F_{2rS}(r_{AB}) = \mathbb{P}(2rS \leq r_{AB}) = F_S(r_{AB}/2r), \quad (7.17)$$

$$F_S(s) = \int_0^s 12s'^2(1 - s')^2(2 + s')ds', \quad (7.18)$$

$$f_{R_{AB}}(r_{AB}) = \frac{d}{dr_{AB}} F_{R_{AB}} = \frac{1}{2r} f_S(r_{AB}/2r), \quad (7.19)$$

where F is a cumulative density function and \mathbb{P} a probability. We can now solve our previous integral

$$\begin{aligned} \frac{1}{V_2^2(r)} \iint dV_A dV_B j_0(kr_{AB}) &= \int_0^{2r} j_0(kr_{AB}) f_{R_{AB}}(r_{AB}) dr_{AB} \\ &= \int_0^1 j_0(2krs) f_S(s) ds = \left[\frac{3j_1(kr)}{kr} \right]^2. \end{aligned} \quad (7.20)$$

References

- [1] A.G. Riess, L. Macri, S. Casertano, et al. *A 3% Solution: Determination of the Hubble Constant with the Hubble Space Telescope and Wide Field Camera 3*. *ApJ* **730** (2011) 119, [arXiv:1103.2976 \[astro-ph.CO\]](#).
- [2] Planck Collaboration. *Planck 2013 Results. XVI. Cosmological Parameters*. . *A&A* **571** (2014) A16, [arXiv:1303.5076 \[astro-ph.CO\]](#).
- [3] V. Marra, L. Amendola, I. Sawicki, and W. Valkenburg. *Cosmic Variance and the Measurement of the Local Hubble Parameter*. *Phys. Rev. Lett.* **110** (2013) 241305, [arXiv:1303.3121 \[astro-ph.CO\]](#).
- [4] P. Coles, A.L. Melott, and S.F. Shandarin. *Testing Approximations for Non-Linear Gravitational Clustering*. *MNRAS* **260** (1993) 765.
- [5] H.-Y. Wu and D. Huterer. *Sample Variance in the Local Measurements of the Hubble Constant*. *MNRAS* **471** (2017) 4946, [arXiv:1706.09723 \[astro-ph.CO\]](#).
- [6] J.M. Shull, B.D. Smith, and C.W. Danforth. *The Baryon Census in a Multiphase Intergalactic Medium: 30% of the Baryons May Still be Missing*. *ApJ* **759** (2012) 15, [arXiv:1112.2706 \[astro-ph.CO\]](#).
- [7] F. Nicastro, J. Kaastra, Y. Krongold, et al. *Observations of the Missing Baryons in the Warm-Hot Intergalactic Medium*. *Nature* **558** (2018) 406, [arXiv:1806.08395 \[astro-ph.GA\]](#).
- [8] S.P. Driver and A.S.G. Robotham. *Quantifying Cosmic Variance*. [arXiv:1005.2538 \[astro-ph.CO\]](#).
- [9] S. Dodelson. *Modern Cosmology*. Academic Press, Cambridge, US (2003), [[Abstract](#)].
- [10] H. Mo, F.C. Van Den Bosch, and S. White. *Galaxy Formation and Evolution*. Cambridge University Press, Cambridge, UK (2010), [[Abstract](#)].
- [11] F. Van Den Bosch. *ASTR 610, Theory of Galaxy Formation, Lecture 12: Large Scale Structure*. . Yale University (2018), [[Lecture Notes](#)].
- [12] F. Bernardeau, S. Colombi, E. Gaztanaga, and R. Scoccimarro. *Large-Scale Structure of the Universe and Cosmological Perturbation Theory*. *Physics Reports* **367** (2002) 1, [arXiv:astro-ph/0112551](#).
- [13] Y.B. Zel'dovich. *Gravitational Instability: An Approximate Theory for Large Density Perturbations*. *A&A* **5** (1970) 13.
- [14] M. White. *The Zel'dovich Approximation*. *MNRAS* **439** (2014) 3630, [[astro-ph.CO](#)].
- [15] C.-P. Ma, and E. Bertschinger. *Cosmological Perturbation Theory in the Synchronous and Conformal Newtonian Gauges*. *ApJ* **455** (1995) 7, [[arXiv:astro-ph/9506072](#)].

- [16] A. Refregier, L. Gamper, A. Amara, and L. Heisenberg. *PyCosmo: An Integrated Cosmological Boltzmann Solver*. [arXiv:1708.05177 \[astro-ph.CO\]](#).
- [17] J.F. Navarro, C.S. Frenk, and S.D.M. White. *A Universal Density Profile from Hierarchical Clustering*. *ApJ* **490** (1997) 493, [arXiv:astro-ph/9611107](#).
- [18] A. Klypin, G. Yepes, S. Gottlöber, et al. *MultiDark simulations: The Story of Dark Matter Halo Concentrations and Density Profiles*. *MNRAS* **457** (2016) 4340, [arXiv:1411.4001 \[astro-ph.CO\]](#).
- [19] K. Tomita. *Formation of Gravitationally Bound Primordial Gas Clouds*. *Progress of Theoretical Physics* **42** (1969) 9.
- [20] J.E. Gunn and J.R. Gott III. *On the Infall of Matter Into Clusters of Galaxies and Some Effects on Their Evolution*. *ApJ* **176** (1972) 1.
- [21] J. Yoo. *AST802, Advanced Topics of Theoretical Cosmology*. University of Zurich (2019), [\[Lecture Notes\]](#).
- [22] W.H. Press and P. Schechter. *Formation of Galaxies and Clusters of Galaxies by Self-Similar Gravitational Condensation*. *ApJ* **187** (1974) 425.
- [23] W. Hu. *Lecture 3: The Halo Model, Structure Formation and the Dark Sector*. Trieste, Italy (2002), [\[Lecture Notes\]](#).
- [24] R.K. Sheth and G. Tormen. *Large-Scale Bias and the Peak Background Split*. *MNRAS* **308** (1999) 119 [arXiv:astro-ph/9901122](#).
- [25] A. Jenkins, C.S. Frenk, S.D.M. White, et al. *Mass Function of Dark Matter Halos*. *MNRAS* **321** (2001) 372, [arXiv:astro-ph/0005260](#).
- [26] J.L. Tinker, A.V. Kravtsov, A. Klypin, et al. *Toward a Halo Mass Function for Precision Cosmology: The Limits of Universality*. *ApJ* **688** (2008) 709, [arXiv:0803.2706 \[astro-ph\]](#).
- [27] J. DeRose, R.H. Wechsler, J.L. Tinker, et al. *The Aemulus Project I: Numerical Simulations for Precision Cosmology*. *ApJ* **875** (2019) 69, [arXiv:1804.05865 \[astro-ph.CO\]](#).
- [28] Planck Collaboration. *Planck 2018 Results. I. Overview and the Cosmological Legacy of Planck*. . (2018), [arXiv:1807.06205 \[astro-ph.CO\]](#).
- [29] R. Teyssier. *Cosmological Hydrodynamics with Adaptive Mesh Refinement: A New High Resolution Code Called RAMSES*. *A&A* **385** (2002) 337, [arXiv:astro-ph/0111367](#).
- [30] R.A. Gingold and J.J. Monaghan. *Smoothed Particle Hydrodynamics: Theory and Application to Non-Spherical Stars*. *MNRAS* **181** (1977) 375.
- [31] L.B. Lucy. *A Numerical Approach to the Testing of the Fission Hypothesis*. *AJ* **82** (1977) 1013.
- [32] W. Dehnen. *Towards Optimal Softening in 3D N-Body Codes: I. Minimizing the Force*

- Error.* [MNRAS **324** \(2001\) 273](#), [arXiv:astro-ph/0011568](#).
- [33] D. Blas, J. Lesgourgues, and T. Tram. *The Cosmic Linear Anisotropy Solving System (CLASS) II: Approximation schemes.* [JCAP textbf07 \(2011\) 034](#), [arXiv:1104.2933 \[astro-ph.CO\]](#).
- [34] U. Seljak, N. Sugiyama, M. White, and M. Zaldarriaga. *A Comparison of Cosmological Boltzmann Codes: Are We Ready for High Precision Cosmology?* [Phys. Rev. D **68** \(2003\) 083507](#), [arXiv:astro-ph/0306052](#).
- [35] M. Crocce, S. Pueblas, and R. Scoccimarro. *Transients from Initial Conditions in Cosmological Simulations.* [MNRAS **373** \(2006\) 369](#), [arXiv:astro-ph/0606505](#).
- [36] V. Springel. *The Cosmological Simulation Code GADGET-2.* [MNRAS **364** \(2005\) 1105](#), [arXiv:astro-ph/0505010](#).
- [37] M. Boylan-Kolchin, V. Springel, S.D.M White, and A. Jenkins. *There's No Place Like Home? Statistics of Milky Way-Mass Dark Matter Halos.* [MNRAS **406** \(2010\) 896](#), [arXiv:0911.4484 \[astro-ph.CO\]](#).
- [38] M. Davis, G. Efstathiou, C.S. Frenk, and S.D.M White. *The Evolution of Large-Scale Structure in a Universe Dominated by Cold Dark Matter.* [ApJ **292** \(1985\) 371](#).
- [39] P.S. Behroozi, R.H. Wechsler, and H.-Y. Wu. *The Rockstar Phase-Space Temporal Halo Finder and the Velocity Offsets of Cluster Cores.* [ApJ **762** \(2013\) 109](#), [arXiv:1110.4372 \[astro-ph.CO\]](#).
- [40] P.S. Behroozi, A. Loeb, and R.S. Wechsler. *Unbound Particles in Dark Matter Halos.* [JCAP **06** \(2013\) 019](#), [arXiv:1208.0334 \[astro-ph.CO\]](#).
- [41] A. Knebe, S.R. Knollmann, S. I. Muldrew, et al. *Haloes Gone MAD: The Halo-Finder Comparison Project.* [MNRAS **415** \(2011\) 2293](#), [arXiv:1104.0949 \[astro-ph.CO\]](#).
- [42] J. Neyman and E.L. Scott. *A Theory of the Spatial Distribution of Galaxies.* [ApJ **116** \(1952\) 144](#).
- [43] R.J. Scherrer and E. Bertschinger. *A Theory of the Spatial Distribution of Galaxies.* [ApJ **381** \(1991\) 349](#).
- [44] F. Van Den Bosch. *ASTR 610, Theory of Galaxy Formation, Lecture 13: The Halo Model & Halo Occupation Statistics.* . Yale University (2018), [\[Lecture Notes\]](#).
- [45] A. Cooray, and R. Sheth. *Halo Models of Large Scale Structure.* [Physics Reports **372** \(2002\) 1](#) , [arXiv:astro-ph/0206508](#).
- [46] R.E. Smith, J.A. Peacock, A. Jenkins, et al. *Stable Clustering, the Halo Model and Nonlinear Cosmological Power Spectra.* [MNRAS **341** \(2003\) 1311](#), [arXiv:astro-ph/0207664](#).
- [47] R. Takahashi, M. Sato, N. Takahiro, et al. *Revising the Halofit Model for the Nonlinear*

- Matter Power Spectrum*. *ApJ* **761** (2012) 152, [arXiv:1208.2701 \[astro-ph.CO\]](#).
- [48] Y.P. Jing. *Correcting for the Alias Effect When Measuring the Power Spectrum Using a Fast Fourier Transform*. *ApJ* **620** (2005) 559, [arXiv:astro-ph/0409240](#).
- [49] P.J.E. Peebles. *The Large-Scale Structure of the Universe*. Princeton University Press, Princeton, US (1980), [\[abstract\]](#).
- [50] S.-J. Tu, and E. Fischbach. *Random Distance Distribution for Spherical Objects: General Theory and Applications to n-Dimensional Physics*. *J. Phys. A: Math. Gen.* **35** (2002) 6557, [arXiv:math-ph/0201046](#).
- [51] P.J.E. Peebles and E. J. Groth. *Statistical Analysis of Catalogs of Extragalactic Objects. V - Three-Point Correlation Function for the Galaxy Distribution in the Zwicky Catalog*. *ApJ* **196** (1975) 1.
- [52] M. Takada and B. Jain. *Three-Point Correlations in Weak Lensing Surveys: Model Predictions and Applications. Version 3*, [arXiv:astro-ph/0304034](#).
- [53] N. Khandai, T. Di Matteo, R. Croft, et al. *The MassiveBlack-II Simulation: The Evolution of Haloes and Galaxies to $z \sim 0$* . *MNRAS* **450** (2015) 1349, [arXiv:1402.0888 \[astro-ph.CO\]](#).
- [54] U. Seljak, and M.S. Warren. *Large Scale Bias and Stochasticity of Halos and Dark Matter*. *MNRAS* **355** (2004) 129, [arXiv:astro-ph/0403698](#).
- [55] B.P. Moster, R.S. Somerville, J.A. Newman, and H.-W. Rix. *A Cosmic Variance Cookbook*. *ApJ* **731** (2011) 113, [arXiv:1001.1737 \[astro-ph.CO\]](#).
- [56] J. Yoo, E. Mitsou, N. Grim, et al. *Cosmological Information Contents on the Light-Cone*. (2019) [arXiv:1905.08262 \[astro-ph.CO\]](#).
- [57] F. Schmidt. *MPA Lectures on Galaxy Clustering*. Max Planck Institute for Astrophysics (2013), [\[Lecture Notes\]](#).
- [58] A. Hall. *The Impact of our Local Environment on Cosmological Statistics*. [arXiv:1911.07855 \[astro-ph.CO\]](#).