Amended Model of Large Scale Dark Matter Structure

by

Alice Chen

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

Waterloo, Ontario, Canada, 2020

Author's Declaration

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

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

Statement of Contributions

Chapter 1: Introduction: Alice Chen was the sole author of Chapter 1, which was written under the supervision of Dr. Niayesh Afshordi.

Chapter 2: Amended Halo Model: The research presented in this paper was done under the supervision of Dr. Niayesh Afshordi, who is the sole co-author on the paper that contains the research presented in chapters 2-4. Dr. Niayesh Afshordi suggested the idea for this amended halo model, while Alice Chen matched the theoretical predictions to data and modified them accordingly. Alice Chen wrote the paper draft, which Dr. Niayesh Afshordi edited and contributed intellectual input on, which is published as:

Alice Chen and Niayesh Afshordi, "Amending the halo model to satisfy cosmological conservation laws", Phys. Rev. D 101, 103522 (2020).

This chapter also contains simulation data from Ryuichi Takahashi, Masanori Sato, Takahiro Nishimichi, Atsushi Taruya, and Masamune Oguri's 2011 paper [35].

Chapter 3: Beyond Mass Density Power Spectrum: This chapter compares the predictions for halo-density cross-power spectrum from Amended Halo Model with the Illustris-TNG simulations [1], and is not yet submitted for publication.

Chapter 4: Conclusions and Future Prospects: Extended conclusions, originally presented in Alice Chen and Niayesh Afshordi, "Amending the halo model to satisfy cosmological conservation laws", Phys. Rev. D 101, 103522 (2020).

Abstract

Dark matter makes up around a quarter of the total energy density in the universe, but its identity remains elusive. Current ways of studying dark matter have centered around its macroscopic properties, such as density distribution and large scale structure formation. The halo model of large scale structure is an important tool that cosmologists use to study the phenomenological behaviour and nonlinear evolution of structure in the universe. However, it is well known that there is no simple way to impose conservation laws in the halo model. This can severely impair the predictions on large scales for observables such as weak lensing or the kinematic Sunyaev-Zel'dovich effect, which should satisfy mass and momentum conservation, respectively. For example, the standard halo model overpredicts weak lensing power spectrum by $> 8\%$ on scales > 20 degrees. To address this problem, we present an *Amended Halo Model*, explicitly separating the linear perturbations from *compensated* halo profiles. This is guaranteed to respect conservation laws, as well as linear theory predictions on large scales. We also provide a simple fitting function for the compensated halo profiles, and discuss the modified predictions for 1-halo and 2-halo terms, as well as other cosmological observations such as weak lensing power spectrum.

Similar to previous and recent works centered around the halo model, this work is physically motivated and matches simulation data to a greater degree of accuracy than the standard halo model currently does. We compare our results to previous work, and argue that the amended halo model provides a more efficient and accurate framework to capture physical effects that happen in the process of large scale cosmological structure formation.

Acknowledgements

This research would have not been possible without funding from NSERC, University of Waterloo and Perimeter Institute, so a big thank you to them.

First and foremost, I would like to thank my supervisor Niayesh Afshordi for his guidance, contributions, and patience during the duration of this project. His intuition for physics and quick ability to fix issues as well as come up with astrophysical equations on the spot never cease to amaze me. Thank you so much for providing me with the direction for this project, and for putting up with and being able to help me fix my many, many mistakes. You are the astrophysicist I aspire to become like one day.

I would also like to thank my committee members Michael Balogh and Rafael Sorkin for giving me counsel during this project and keeping me on the right track. Thank you to Michael Balogh for always replying to my emails quickly and making scheduling meetings so much easier. A special thanks is given to Rafael Sorkin as well for staying to talk to me after group meetings about Causal Set theory and quantum gravity conjectures in general - I have learned so much from you about physics and it was often those conversations that continuously renewed my interest in theoretical physics, even though my current project was almost a separate field.

I would like to further thank Perimeter Institute members Latham Boyle, Neal Dalal, and Simon Foreman for helpful discussions and comments. A special thanks to Neal Dalal for being on my thesis defence committee as well.

Of course, no acknowledgements section is complete without naming nearly everyone that one has spoken to during the duration of one's degree, so here it goes. Thank you to Tianyi Yang, Chengyu Xi, Muzi Li, Saoussen Mbarek, and Matthew Robbins for your senior students' advice and help, either pertaining to academics or otherwise; this is not to call you guys old (although by the time this thesis is finished, you probably will be). An extra thank you to Matthew for editing this thesis despite being very busy, and for being an overall awesome mentor figure. Also, a thanks to my friend and equally neurotic "academic fraternal twin" Samantha "Sam" Hergott for putting up with me from undergrad to graduate school, and being at times my sole social contact in Waterloo.

Last and most, a special thanks is given to my family for their support.

Table of Contents

List of Figures

List of Tables

1 Introduction

1.1 Dark Matter and Cosmology

Cosmology is a topic that has been extensively studied in the past few decades. With new developments in technology and a resulting increase in cosmological data, we gain more and more information that will help us understand the evolution and composition of our universe. One of the key unanswered questions remaining is the mysterious nature of dark matter.

Dark matter makes up around 25% of the total energy of our universe [6], but we currently have no way of observing it directly. We know about the existence of dark matter through its gravitational effects and that it must have had a significant role to play in cosmic evolution; it has been commonly theorized that without dark matter, large galaxies like the Milky Way could not have formed [12, 28]. However, what dark matter is made of, what kinds of structure it has formed throughout the universe, and how exactly it has affected galactic evolution remain active areas of research.

Since we are unable to observe dark matter in other ways than through its gravity, inferences about dark matter have so far been solely based on its gravitational effects [28]. Dark matter's gravity bends the path of light as it travels past, and attracts baryonic matter to clump together inside dense dark matter regions. By observing the path of distant sources' light bending around invisible structures, and the evolution of large scale baryonic structures probed by cosmic microwave background (CMB), galaxies, and stars, we now have evidence of not only dark matter's existence, but also clues to infer some of its properties [28].

The earliest known evidence for dark matter dates back to the 1930's, with Fritz Zwicky's observation of galactic velocities in the Coma cluster of galaxies [28]. He noticed that the dispersion among the line of sight velocities of galaxies in the cluster was much larger than what the luminous matter - such as stars and gas - in the cluster could gravitationally hold together, and concluded that an additional unseen source of matter must be present. Without this unseen matter and its subsequent gravity, galaxies would fly off from the cluster [28]. Further evidence for this unseen matter was discovered in the 1960's by Vera Rubin. She measured the rotation curves of spiral galaxies, also concluding that the velocities were too high to be gravitationally

held together by the luminous matter that was observed in the galaxies. The unseen matter that Zwicky observed in the 1930's then had further evidence to substantiate its existence, becoming what we now call dark matter. Today, the most compelling evidence for its existence comes from recent measurements of the CMB (e.g., [2]). These measurements showed that the ratio of odd to even acoustic peaks in the CMB anisotropy power spectrum require the density to be dominated by a cold dark matter at the time of recombination, 380,000 years after the big bang. This matches the unseen matter characteristics that Zwicky and Rubin have observed around clusters and galaxies, almost completely confirming its existence [2, 28].

Currently, the best way we have of detecting dark matter is through weak lensing. By observing how light from distant sources get distorted by dark matter's gravity on the path to crossing Earth, we can make inferences about the abundance of dark matter and approximately where in the universe it resides. However, this method can only probe dark matter's gravitational effect. Consequently, an alternate observation method that has been to search for the byproducts of dark matter annihilation/decay astrophysical observations (should it have strong enough coupling to the standard model of particle physics) [8]. For example, in [9] it is hypothesized that ultra high energy extragalactic cosmic rays are the result of dark matter particle annihilation. However, high energy astrophysical processes such as blazars or magnetars/pulsars are also capable of producing these cosmic rays [40] and are thus a more conventional explanation for them. As this example demonstrates, the best way of studying dark matter in the universe currently is still through its gravitational effect.

From recent and ongoing weak lensing observations, we can infer dark matter to be cold and collisionless - its particles have negligible self-interactions and interactions with any Standard Model particle that we know so far [5, 6]. Their "darkness" - or lack of observational evidence for interactions with photons - is what makes the second point likely, since interactions with standard model particles often result in some form of photon production. These dark matter properties have led to the proposed dark matter candidates that are discussed in Section 1.2 [5, 6].

In addition to constraining dark matter particle candidates, we can make inferences about the type of structures dark matter particles would form based on their cold, collisionless properties, since any structure formation would be solely due to gravitational effects. Dark matter particles are theorized to

clump together gravitationally to form virialized structures known as halos [12]. Halos typically have virial radii of scale around $10^{21} - 10^{22}$m, or 0.1-1 Mpc, and masses of $10^{10} - 10^{15}$ solar masses. Their densities are thought to follow a smooth distribution - popular models so far include the Navarro-Frenk-White (NFW) profile [22] and Einasto profile [15], which parameterize halo density as a function of halo radius. They will also both be discussed in more detail in later sections. These structures' subsequent physical effects - or lack thereof - on observable baryonic structures also support the current consensus about dark matter's weakness of interaction with ordinary matter.

1.2 Current Hypothesized Candidates of Dark Matter

The true composition of dark matter remains a question, but we can set some constraints for potential candidates based on observational evidence. Any dark matter candidate proposals must at least satisfy the following criteria [5]:

1. The abundance of this particle must match the observed abundance of dark matter today, based on its gravitational effects and observed CMB temperature fluctuations;

2. This particle candidate must have been non-relativistic during the epoch of matter-radiation equality to form the currently observed galactic structures - particles with masses below a few keV are likely ruled out [5] due to their large free streaming length which washes out structure on small scales;

3. The dark matter candidate should have very weak electromagnetic interactions [5], as they are highly constrained by astrophysical observations;

4. Dark matter particles should be cosmologically stable, i.e. have a lifetime longer than the age of the universe [5] in order to not have decayed.

5. Any self-interactions dark matter particles are constrained by observations such as cluster collisions, e.g. the Bullet cluster [5].

Historically, the Weakly Interacting Massive Particle - or WIMP - has been the most popular candidate for dark matter [5]. WIMPs are by design non-relativistic, non-interacting, and electromagnetically neutral, automatically satisfying requirements 2, 3, and 5. Their cross-sections are assumed to be within the scale of weak interactions, making WIMPs dominated by the weak force. The time of "freeze-out" in the early universe - when WIMPs decoupled from leptons and baryons to scatter across the universe - is when weak interaction processes would have set the dark matter mean (comoving) density to match its current observed value. This also occurs at the scale where many Beyond Standard Model theories of particle physics predict new particles, in order to address the Higgs hierarchy problem. This coincidence of scales, often referred to as the *WIMP miracle*, made WIMPS very popular [5]. This additionally satisfies candidate requirements 1 and 4, since the abundance matches the observed cosmological dark matter density and particles governed by the weak force are generally stable and have little interaction following their freeze-out (neutrinos are an example). However, decades of laboratory searches to find WIMPs have not yet been successful in finding them, leading to a current fade in their popularity [5].

As a result of the WIMP popularity decline, alternate models of hypothetical non-baryonic particles have been proposed as candidates for dark matter. The one that is gaining the most popularity at the present is the axion [16]. Axions have recently been proposed as alternate candidates, due to many of the underlying assumptions required for WIMPs to be unconfirmed (such as new Beyond Standard Model physics) and their subsequent observation predictions unobserved. Axions are also by definition non-relativistic, non-interacting, and electromagnetically neutral, automatically satisfying criteria 2, 3, 4, and 5 [5, 6, 16]. They can also be integrated into the Standard Model of particle physics (not to be confused with the Standard *Halo* Model, which will be elaborated on later) as easily as WIMPs, due to them explaining a lack of observed strong CP violation [16, 6]. In the Standard Model, strong interaction forces break the combined symmetries of charge and parity (CP symmetry), resulting in the hypothetical existence of a light neutral particle that is the axion [6]. These axions can be created in the early universe and be abundant enough to match current observations [16], satisfying requirement 1. There are also less assumptions to be made about the Standard Model to explain their existence over WIMPs and their size can explain the lack of direct detection so far - these are reasons why researchers are leaning more

and more towards dark matter being axions over WIMPs.

Another candidate gaining popularity at the moment is the massive right-handed neutrino [10], proposed to exist as an extension to the standard model. Symmetrically, it would be the mirror image of the left-handed light neutrino we know to exist, and would have the abundance, electromagnetic neutrality, and stability for dark matter [10]. These right handed neutrinos have no coupling to matter or each other and are non-relativistic, satisfying all the requirements above and making this another viable dark matter candidate.

Other candidates for dark matter include primordial black holes, which are black holes created in the beginning of the universe shortly after the Big Bang [11]. These black holes would be electromagnetically neutral and theoretically have masses ranging from 10^{16} - 10^{33}kg, although recent observational efforts have constrained this mass range to be from 10^{19} - 10^{24}kg [23]. This constraint ensures that these black holes not have yet completely evaporated via Hawking radiation [11] - satisfying requirement 1 and 3, - but are also not large enough to violate current observational limits. Since primordial black holes are predicted to exist independently of dark matter predictions and have little interaction with baryonic matter (satisfying requirements 2, 4, and 5), it makes sense for them to be a dark matter candidate. Alternate models of gravity, which postulate that general relativity works differently on large scales rather than an existence of additional matter, are also an exotic proposed explanation for the gravitational effect thought to be produced by "dark matter", albeit not a very popular model.

For the purpose of this research project, we assume dark matter to be made of non-relativistic particles that approximately evolve based on Newtonian gravity. It is not of high importance to be stringent on which candidate we choose, as long as it is cold (non-thermal), collisionless, has negligible self-interactions or interactions with baryonic matter, and starts with standard adiabatic ΛCDM (see below) initial conditions. Their differences in mass and standard model particle behaviours can be ignored on the large scales that we are using to characterize dark matter density and structure.

1.3 Cosmological Density Distribution and Large Scale Structure

Currently, the most widely used model in cosmology is the Λ Cold Dark Matter (ΛCDM) model. As the name suggests, dark matter here has the cold and collisionless properties mentioned above, making it ideal background model for this project. In ΛCDM cosmology, all matter was initially smoothly distributed in the universe, with some small fluctuations very similar to a Gaussian random field [12]. These fluctuations would perturb the matter field slightly and eventually lead to matter becoming slightly more dense in some areas than in others. The denser regions would then gravitationally attract more and more matter to the area. This growth would in turn result in larger fluctuations within the matter field, causing them to no longer remain Gaussian. Consequently, matter became non-linearly distributed and denser regions became more dense, forming halo structures, while less dense regions lost more and more matter, becoming what we know today as voids. Early attempts have been made to use perturbation theory to model how dark matter and galaxies began clustering, but it was found that these theories broke down in the highly nonlinear regime [12] when matter density fluctuations became very large and were no longer Gaussian (although see [18] and [31] for a more recent approach using Zel'dovich approximations, as well as [27] which incorporates the effective field theory into the halo model using perturbative effects). Therefore, a new model is needed to predict matter distribution at late times, in the nonlinear regime.

Once in the nonlinear regime, the matter regions can be modelled using their overdensities, which trace how the density in a local area compares to the average linear density. The overdensity is defined mathematically by the expression,

$$\delta(\vec{r}) = \frac{\rho(\vec{r})}{\bar{\rho}} - 1, \tag{1}$$

where $\bar{\rho}$ is the average matter density of the universe and $\rho(\vec{r})$ is the matter density at position $\vec{r}$. The regions with a high positive overdensity would then collapse to form virialized spherical structures known as halos. The point of collapse depends on the mass of the halo and the overdensity value at the point of collapse, which is often denoted Δ in papers and generally varies from 200-1600 times the average background density $\bar{\rho}$, depending on the model of collapse used.

An interesting route to quantify cosmic matter distribution is to model the statistical properties of objects formed from matter (e.g., galaxies or CDM haloes) in the universe. This is often done by using correlation functions, such as the 1-point correction function, which describes the number density of objects, and the 2-point correlation function, which describes number density as well as objects' correlations with each other. Both these functions are used in the formulation of the standard halo model, which is summarized below.

1.3.1 Standard Halo Model

Most of the dark matter inferred to exist throughout the universe today is thought to be within virialized halo structures. There exists commonly agreed upon models for how exactly dark matter is distributed within halos (see 1.3.2 as well as [12], [15] and [22]), while a large scale description is typically described by the standard halo model (SHM).

The SHM is the most commonly used way to describe dark matter density throughout the universe. It is modelled analytically by the power spectrum, which is the Fourier transform of the 2-point correlation function in real space - Fourier space helps simplify some calculations that would otherwise be very complicated (turns a convolution integral to a multiplication between two functions). There are two key components to SHM - the density of dark matter inside halos - described by the 1-halo term, and the distribution of dark matter halos themselves, described by the 2-halo term. Far outside halos, dark matter is very sparsely distributed, with a density almost close to 0 (vacuum). The standard halo model encompasses both the densities inside and outside halos, by describing density using the power spectrum, or the density in Fourier space. This power spectrum has a 1-halo term in the following form,

$$P_{1-\text{halo}}(k) = \frac{1}{\bar{\rho}^2} \int dM n(M) M^2 |u(k|M)|^2, \tag{2}$$

which describes how matter correlates with each other inside halos, and a 2-halo term,

$$P_{2-\text{halo}}(k) = \left[\frac{1}{\bar{\rho}} \int dM M n(M) b(M) u(k|M) \right]^2, \tag{3}$$

which describes how halos themselves distributed throughout space correlate with each other. The SHM's full analytical form is,

$$P(k)_{\text{standard}} = P_{1-\text{halo}} + P_{2-\text{halo}} P_L(k), \tag{4}$$

where $P_L(k)$ is the linear power spectrum, which describes how matter was initially distributed in the early universe, when density was approximately uniform. Furthermore, in Eqn. (2)-(3), $n(M)$ is the halo mass function [39], $b(M)$ is the bias function [38, 39], and $P_L(k)$ is the linear matter power spectrum (for more details refer to Appendix A).

On large scales, the standard halo model provides a good estimate of dark matter distribution. However, in certain regions the standard model is not a good fit, and it has some physical issues as well, both of which will be elaborated in the Section 1.4.

1.3.2 Density Profiles Inside Halos

One of the most commonly used models for describing dark matter density inside halos in real space is the Navarro-Frenk-White (NFW) profile. This was an empirical density distribution of halos found in the 1990's by Julio Navarro, Carlos Frenk, and Simon White (thus aptly named NFW) modelling large scale simulation data [22]. The spherically-averaged profile for dark matter density as a function of radius inside halos is as follows,

$$\rho(r) = \frac{\rho_{crit}\delta_c}{(r/r_s)(r/r_s+1)^2} \tag{5}$$

and

$$\delta_c = \frac{\Delta}{3}\frac{c^3}{\ln(1+c) - c/(c+1)} \tag{6}$$

where the concentration parameter $c{=}r_{virial}/r_s$ and Δ is the overdensity value at the time of collapse. Putting Eqn. (5) and (6) together, we can obtain the halo density in a more general form,

$$\delta_{\rm NFW}(r) \equiv \frac{\rho(r)}{\bar{\rho}} = \frac{\Omega_m\delta_c}{(r/r_s)(1+r/r_s)^2}, \tag{7}$$

where $\rho(r)$ is the density of the halo region, $\bar{\rho} = \frac{3\Omega_m H^2}{8\pi G}$ is the mean density of the universe, r is the radius from the halo centre, δ_c is defined in Eqn. (6) above, and r_s is a fitting parameter found from simulations known as the scale radius [22].

This density profile was found by Navarro, Frenk, and White performing and analyzing ΛCDM simulations. Testing dark matter density inside halos

to see if they fit the NFW profile is a good check to see if our model and simulation results are reliable, since NFW is widely used - the main reason being that NFW is largely independent of cosmology and scale.

Another profile that is commonly looked at is the Einasto profile. This profile (sometimes called "Einasto's $r^{1/n}$ model" [15]) is a special case of de Vaucouleurs' model of spherical density, and states that

$$\rho(r) \propto e^{-Ar^{\alpha}}, \tag{8}$$

where α is a parameter describing the degree of curvature of the profile (this depends on the model used), A is a parameter given as,

$$A = -\frac{2}{\alpha r_{-2}^{\alpha}}, \tag{9}$$

and r_{-2} is the radius at which the logarithmic slope of the density is -2 [15]. Einasto provides a good fit closer to the centre of halos, as r approaches 0. Elsewhere, NFW is generally provides good density predictions.

These profiles do a good job of describing dark matter density inside halos, but not outside halos, or even at the transition region/boundary between halos and open space. While on large scales dark matter is sparsely distributed throughout space, coming up with a model that both describes halo density as well as the dark matter density inside and outside halos is actually very complicated, so we once again turn to the standard halo model in Fourier space.

1.4 Issues with Current Models

While the standard halo model gives a description of universal dark matter density distribution, it does not come without problems. One issue is that the standard halo model becomes inaccurate in the transition regions around the halo virial radius out to the boundary between the halo region and void regions. This is due to a lack of a connection mechanism in the standard halo model, as the analytic transition from 1-halo term dominance to 2-halo term dominance is characterized by a sharp cutoff at the virial radius (see Figure 1 for a visual summary).

Furthermore, conservation laws, such as mass and momentum, are not taken into consideration in this model, resulting in unphysical behaviour. A

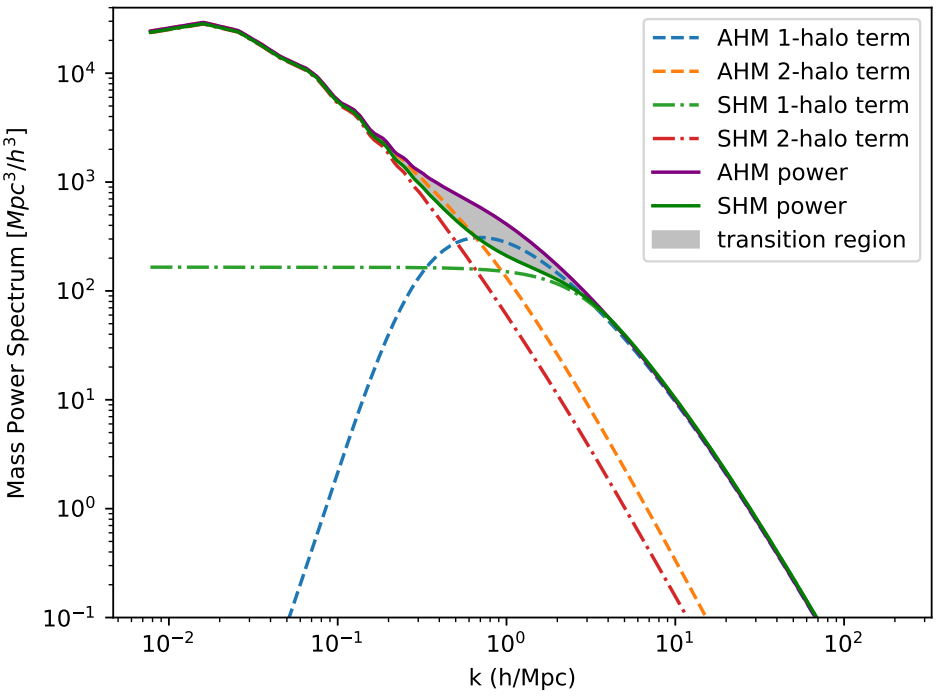

Figure 1: The standard halo model and amended halo model power spectra, as well as their 1-halo and 2-halo terms. The grey shaded region near the middle is where the transition region - where the standard halo model doesn't match simulation power - occurs at.

result of this lack of conservation leads to matter overdensity levelling off to a constant value even as the wavenumber k approaches 0 (as volume approaches infinity, essentially going to all of space), which does not match the linear power spectrum. This is shown in Figure 2, where the standard halo model's power spectrum starts approaching a constant value at very small k's, while it should be approaching the linear power spectrum (which is $\propto k$), as a result of mass conservation.

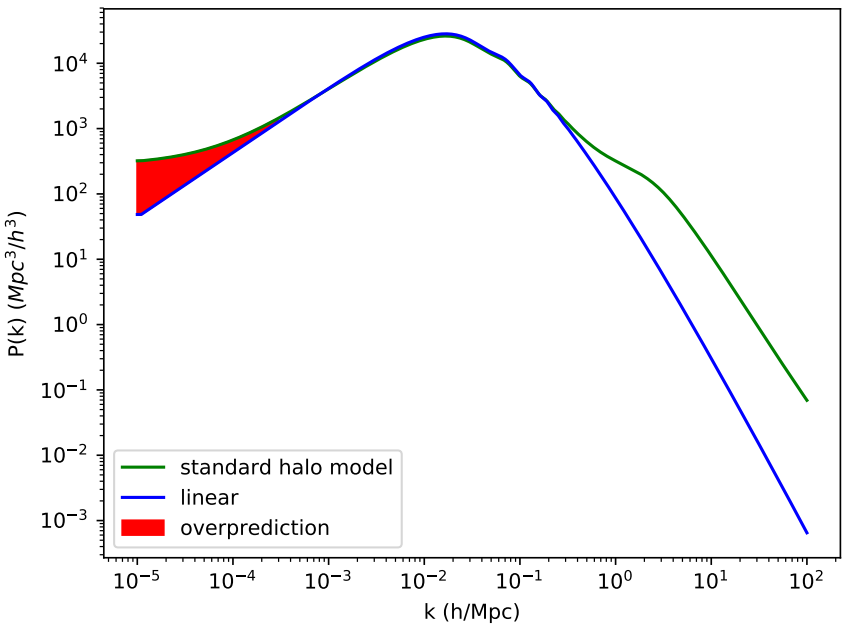

Figure 2: The standard halo model vs linear power spectrum. It can be seen in the red shaded region where the standard halo model overpredicts power at small k's. While the power should approach the linear power spectrum, it is starting to level off to a constant. This is due to the unphysical behaviour resulting from a lack of mass conservation in the model.

This unphysical picture of the standard halo model at small k's is the main motivation for this project. To provide a more accurate model for future lensing projects, we aim to modify the current halo model in a way that replicates its success on large scales, but also describes the dark matter distribution well in transition regions as well as enforce mass conservation and consequently,

momentum conservation.

1.5 Previous Attempts at Fixing the Halo Model

The standard halo model's problems with accuracy at small scales, and lack of enforcing physical conservation laws have been recognized in previous literature. Seljak and Vlah [31] noticed that while the standard halo model did not work well with standard perturbation theory, it was still necessary to try to incorporate perturbation theory into the model because without it, there would be no covariance matrix convergence unless simulations go up to scales of around 1000Gpc^3 or larger [31]. Since this would take too long and is too expensive to run, a reliable perturbation theory is needed to extrapolate the standard halo model to larger scales [31]. To fix this, Seljak and Vlah proposed to apply Zel'dovich approximations with 1 loop perturbation to the halo model on larger scales, where the 2-halo term dominates, and provide a compensation with a Padé series ansatz on smaller scales where the 1-halo term dominates [31]. As a result, the halo power spectrum becomes

$$P(k) = P_{Zel} + P_{BB}, \tag{10}$$

where P_{Zel} is the power term from the Zel'dovich approximations and P_{BB} is the broadband power term using the Padé series ansatz. The broadband term can be written as

$$P_{BB} = A_0 F(K) \left(\frac{1 + \sum_{m=1}^{n_{max}-1} (kR_m)^2}{1 + \sum_{m=1}^{n_{max}-1} (kR_{nh})^2} \right), \tag{11}$$

where F(k) is a fitting parameter that vanishes as k goes to 0 and depends on a fitting parameter R, and A_0 is of the form

$$A_0 = \frac{1}{\rho^2} \int M^2 dn(M). \tag{12}$$

Their model satisfied mass conservation laws as well as applied perturbation theory more successfully to the halo model. However, it is unclear what the cosmology dependence for this model is, as it is unspecified what cosmological parameters were used in this model. The k range fitted was also less than 1Mpc/h, so it is unclear how this model would perform for simulations with

larger box sizes and the resulting larger k ranges. The model is also rather mathematically complicated and has parameters with little physical meaning, indicating further work could still be done with correcting the standard halo model's issues.

Another model that set out to fix the mass conservation problem was published by Schmidt in 2015 [29], who suggested using halo stochastic covariance - referred to as shot noise - as compensation for the halo model. Schmidt's model has also been extrapolated to higher order statistics beyond the power spectrum, such as bispectrums, trispectrum, and matter velocity fields. However, his model is still inaccurate in the transition regions near the halo virial radius, which is mentioned in the paper as well [29].

Most recently, the work by Philcox et al. [27] uses effective field theory (EFT) to make corrections to the standard halo model. Their model rests on 2 key assumptions: the position of dark matter halos are a function of the underlying nonlinear density field smoothed on an unknown scale R and the long wavelength density field can be described by EFT at the 1-loop order (first integral over momentum in perturbation theory expansion) [27]. Since EFT arises from smoothed fluid equations, it works well in describing a smoothed nonlinear density field, and is a good empirical fit to the halo number density [27], a key ingredient in large scale halo models. EFT can incorporate physical effects, such as density fluctuations at the quasi-nonlinear and nonlinear regimes, into the halo model as perturbative terms. While this work also uses a Padé ansatz with perturbation theory to achieve a more accurate result for the halo model on all scales, it differs from Seljak and Vlah (2015) [31] in that the power spectrum doesn't require a complex compensation function with many parameters. This model only has two parameters that are fitted empirically - the scale R in its first assumption and some effective sound speed to incorporate small scale physics [27]. Instead of assuming the 2-halo term in the halo model as being proportional to the linear power spectrum - as the standard halo model and our amended model below do - Philcox et al. (2020) modifies the 2-halo term to be proportional to a power term arising from EFT [27]. It is also different from the model presented in this thesis because it uses perturbation theory to predict halo statistics that can be extrapolated to higher orders, then compares the result to simulations, whereas we have calibrated the compensation term in the halo model to match simulations. However, we have been able to fit the power spectrum on a wider range of scales, as Philcox et al. (2020) only goes up to k $\leq$1.0 h/Mpc, while we fit to

a range of 0.001h/Mpc $\leq$ k $\leq$30.0h/Mpc. At the time of this paper's release, the EFT and semi-analytic model does not address the un-physical behaviour due to a lack of conservation mechanism in the halo model either.

Currently, a commonly used numerical package that predicts the nonlinear power spectrum on a wide range of scales and redshifts is HALOFIT ([26], [35]). HALOFIT does a good job of numerically estimating the simulated nonlinear power spectrum for many cosmologies, but it largely works for the mass power spectrum and cannot be extended easily to other models (such as halo-matter cross spectrum, halo-halo spectrum), nor is the model based on an obvious physical framework.

In our work below, we attempt to address the mass conservation problem by fitting a compensated halo profile to a larger range of wavenumber k's, as well as model the transition region between halo and vacuum. We were able to find a semi-analytic empirical model for the power spectrum that fits simulation data well on scales $0.001 \leq k \leq 30.0$ that also takes into account mass conservation. We have not yet attempted to extrapolate this model to higher order statistics such as the bispectrum and trispectrum.

2 Amended Halo Model

2.1 Main Motivation

As mentioned in the introduction above, the nature and composition of dark matter has been a long-standing problem in cosmology. All observational evidence for the existence of dark matter so far has been purely gravitational, and it is based upon these observations that we currently infer dark matter's nature and properties. Due to these observations, it is currently hypothesized that dark matter particles do not have any other detectable signatures aside from gravity. As a result, they also do not interact with standard model particles or photons, other than through their gravitational pull, which is the key underlying assumption we use here in this model (i.e. all thermal effects can be ignored).

The most commonly used analytic framework for the formation of dark matter structure in cosmology has been the SHM, where dark matter particles clump together to form (nearly-) spherical virialized structures known as halos. Dark matter particle properties, along with the cosmological initial conditions, determine the properties of SHM, which describes how halos are formed and what their internal structures are like (e.g., [30, 12]). In spite of its success in describing the statistics of nonlinear structures on small scales (e.g., [33]), the SHM is not dynamical, and thus has no way to guarantee conservation laws, such as for mass or momentum. This leads to unphysical behaviour, such as significant deviations from linear theory predictions at small wavenumbers, $k \to 0$, due to a significant contribution from the 1-halo term that shouldn't be present on these scales, e.g., [12, 14]. This 1-halo term mainly describes how dark matter densities inside halo structures correlate with each other, and hence why they are dominant on small scales while approaching 0 at large scales.

Even though at first this may sound like an academic question, current and upcoming wide-field surveys of weak lensing, the kinematic Sunyaev-Zel'dovich effect, and pre-reionization 21-cm intensity will be probing total mass, momentum, and hydrogen mass on large scales [34]. Thus, they will be sensitive to theoretical deficiencies such as violation of conservation laws that the SHM entails. For example, the current SHM does not match the simulated power well in the region around the halo radius (in the transition regions, see Figure 4 where the green curve deviates from the simulation data). As a result,

we need to make amendments to the current halo model in order to obtain a more accurate picture of cosmic structures on all scales. This project aims to provide a simple and user-friendly prescription to implement this amendment, what we will call the *Amended Halo Model*. Our result can be visually summarized in Figure 3 in Section 2.3, which compares our amended halo model predictions for the matter power spectrum to the standard halo model which overpredicts power on large scales.

2.2 Review and Connection to Standard Halo Model

The most important ingredient of the SHM is that all cosmological halos approximately follow a parametrized universal density profile. The earliest proposal for this density profile was the NFW profile [22], although more precise extensions have been considered more recently [15, 36, 17]. The NFW profile is one of the most widely used profiles to date, and was developed through N-body simulations of dark matter particles [22]. By using the data from these large scale simulations, they came up with a formula that describes the spherically averaged density of dark matter within halo structures. This density was fitted by Eqn. (7) above.

The NFW profile provides a good fit for dark matter density inside the virialized halo region; however, significantly outside this region (i.e. on large scales), where $r \gtrsim r_{\rm vir}$, there does not exist a clear consensus on a universal dark matter density profile (but see [36] for one proposal). It should also be mentioned here that we need dark matter models, such as SHM and AHM, beyond the NFW profile, because the NFW profile is limited in what it can describe.

In the SHM, the matter overdensity is written as a sum over halos:

$$\delta(\mathbf{x}) = \sum_j \delta^j(\mathbf{x} - \mathbf{x}_j), \tag{13}$$

or in Fourier space as:

$$\delta_{\mathbf{k}} = \sum_j \delta^j_{\mathbf{k}} \exp(i\mathbf{k} \cdot \mathbf{x}_j), \tag{14}$$

where $\mathbf{x}$ and $\mathbf{x_j}$ are position coordinates in real space and $\mathbf{k}$ is the wavenumber vector in Fourier space.

For individual halo profiles in SHM, we often use the Fourier transform of a mean profile (assuming its universality):

$$\delta^j_{\mathbf{k}} \simeq \frac{M^j}{\bar{\rho}} u(k|M^j) \equiv \int d^3\mathbf{x} \exp(i\mathbf{k}\cdot\mathbf{x}) \delta^j_{\rm mean}(|\mathbf{x}||M^j). \tag{15}$$

Here, M^j is not the total mass of the j-th halo, which is not well defined to begin with, but rather the mass on a fixed scale. We opt to use M_{200c} as the mass within the radius where the mean halo density is $200\times\bar{\rho}/\Omega_m$, where Ω_m is the cosmological constant for matter.

Note that Eqn. (15) ignores the (possibly correlated) variations in profiles of halos with the same M^j, which is a fundamental limitation of the SHM, and halo models in general. We shall come back to this issue, and our quick fix for it, below.

For the NFW profile (7), $u(k|M)$ has the analytical form:

$$u(k|M) = \frac{4\pi\rho_s r_s^3}{M}\Big\{ \sin(kr_s)\big({\rm Si}[(1+c)kr_s] - {\rm Si}(kr_s)\big) \\ + \cos(kr_s)\big({\rm Ci}[(1+c)kr_s] - {\rm Ci}(kr_s)\big) - \frac{\sin(ckr_s)}{(1+c)kr_s}\Big\}, \tag{16}$$

where Si and Ci are the sine and cosine integral functions respectively [12]. We find the concentration parameter c using the Eqn. (56)-(57) of Okoli & Afshordi (2015) [24]. We then use this c to then find the scale radius from NFW.

We can now discuss the simplest application of the halo model. Given a choice of halo profile $u(k|M)$, the matter power spectrum in SHM is given by Eqn. (4). As we discussed in Section 2, there is no simple mechanism in SHM to enforce conservation laws on large quasilinear scales. It is arbitrary to split the density into multiple halos for small k's, or on large distances that involve several halos. However, requiring $\delta_{\mathbf{k}} \to \delta_{L,\mathbf{k}}$ for small k's (i.e. approximately linear evolution on large scales) will also require a fine-tuned cancellation between the diagonal and off-diagonal parts of the covariance matrix of $\delta^j_{\mathbf{k}}$, for 1-halo and 2-halo terms. For example, this would not be satisfied by the choice of a universal profile, such as Eqn. (15) in NFW,

because NFW does not model dark matter density well beyond the halo virial radius.

While such a constraint is hard to impose in SHM (but see Section 1.5 for a summary of other attempts), in the next section, we develop an *amended halo model* that automatically satisfies this constraint as $k \to 0$, and yet replicates the success of SHM at large k's.

2.3 Amending the Halo Model

Here, we propose a small improvement to the halo model that automatically satisfies mass conservation. To do this, we change Eqn. (13)-(14) to separate the linear overdensity from (now compensated) halo profiles:

$$\delta(\mathbf{x}) = \delta_L(\mathbf{x}) + \sum_j \delta^j(\mathbf{x} - \mathbf{x}_j) \tag{17}$$

and thus

$$\delta_{\mathbf{k}} = \delta_{L,\mathbf{k}} + \sum_j \delta^j_{\mathbf{k}} \exp(i\mathbf{k} \cdot \mathbf{x}_j). \tag{18}$$

We also modify halo profiles u$(k|m)$ to become

$$u_{\rm AHM}(k|M) \to f(kr_{vir})\tilde{u}_{\rm NFW}(k|M), \tag{19}$$

where $f(x)$ is a dimensionless fitting function we find using simulation data. Now, requiring that $f(x) \to 0$ as $x \to 0$ ensures that individual halo profiles are *compensated*, i.e. have zero integral:

$$\int d^3\mathbf{x}\ \delta^j(\mathbf{x}) = \lim_{\mathbf{k}\to 0} \delta^j_{\mathbf{k}} = \frac{M^j}{\bar{\rho}} u_{\rm AHM}(0|M^j) = 0. \tag{20}$$

Furthermore, $\tilde{u}_{\rm NFW}$ is defined to be the same as $u_{\rm NFW}$ for large k's, but without the sharp cutoff at r_{200c}. In other words, we replace the sharp real-space cutoff at virial radius in AHM, with a gentle Fourier space cutoff $f(x)$, that smoothly interpolates between overdense and underdense regions. As such, we let c $\to \infty$ (and thus $r_{200c} \to \infty$) only within the curly brackets in Eqn. (16) (not changing the prefactor):

$$\tilde{u}_{\rm NFW}(k|M) \simeq \frac{4\pi\rho_s r_s^3}{M}\left\{\sin(kr_s)\big(\frac{\pi}{2} - {\rm Si}(kr_s)\big) - \cos(kr_s){\rm Ci}(kr_s)\right\}, \qquad (21)$$

Now, the power spectrum becomes

$$P_{\rm AHM}(k) = \frac{1}{\bar{\rho}^2}\int dMn(M)M^2|u_{\rm AHM}(k|M)|^2 + \left[1 + \frac{1}{\bar{\rho}}\int dMMn(M)b(M)u_{\rm AHM}(k|M)\right]^2 P_L(k). \qquad (22)$$

This new power spectrum will automatically approach linear power when $k \to 0$, as $\tilde{u}_{\rm AHM}(k|M) \to 0$, but will recover SHM on large k's with small corrections. In the next section, we find that this amended model gives a better fit at small k's than the standard halo model does, based on data from N-body simulations. This amended model also yields fits on the same level of accuracy as the widely used numerical HALOFIT package [26, 35], and it is based on a more solid physical picture of structure formation (has a more physical background framework, since it is modelled with enforced mass conservation).

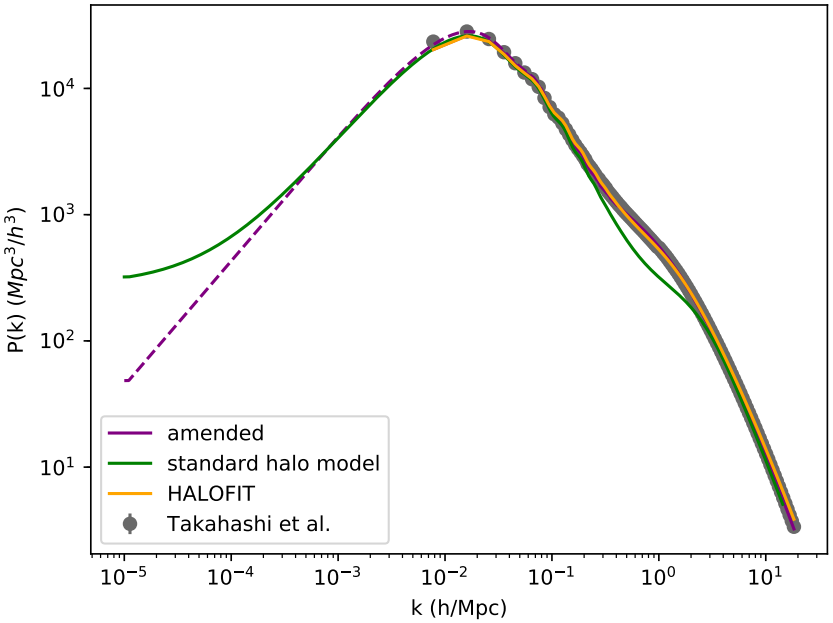

Figure 3: Comparison of predictions for nonlinear matter power spectrum, using Amended Halo Model (introduced here), HALOFIT [35], and the standard halo model [12]. The data points are from Takahashi et al. simulations [35]. It can be seen that the standard halo model power starts to approach a constant value at small k's instead of the linear power spectrum, which is unphysical, as the standard model does not conserve mass. In our amended model, the power starts going to 0 as k→0, similar to linear theory predictions, which is what we should expect.

2.4 Method and Simulations

In order to compare with the HALOFIT model used in the Code for Anisotropies in the Microwave Background (CAMB) package [1][19], the data used to investigate this amended halo model was obtained through N-body simulations of dark matter evolution, using Gaussian ΛCDM linear initial conditions. The simulation data are primarily from Takahashi et al. (2012) [35], using Nishimichi's simulations. We studied the matter power spectra at $z = 0$ for different cosmologies, summarized in Table 1.

[1] `https://camb.info/`

	Ω_b	Ω_m	h	σ_8	n_s
WMAP1	0.044	0.290	0.72	0.90	0.99
WMAP3	0.041	0.238	0.732	0.76	0.958
WMAP5	0.046	0.279	0.701	0.817	0.96
WMAP7	0.046	0.290	0.70	0.81	0.97

Table 1: Cosmology parameters used in Takahashi et al.'s simulations [35].

The simulations had box sizes of 320, 800, and 2000 Mpc/h and a particle number of 1024^3, starting at redshift $z = 99$ and ending at $z = 0$.

2.5 Results and Discussion

We apply Eqn. (17)-(22) to the power spectra obtained by Takahashi et al.'s simulation data [35] (used to calibrate the HALOFIT model) and attempt to parametrize the cutoff function $f(x)$ in Eqn. (19) that can fit the data with an error $\leq 5\%$.

Furthermore, we require $f(x) \propto x^2$ for $x \ll 1$, while it approaches 1 for $x \gg 1$. The former ensures that the (spherically) averaged halo profile is analytic in $\mathbf{k}$ and compensated, i.e. the leading term in $\delta_{\mathbf{k}}$ should be $\mathbf{k} \cdot \mathbf{k}$ in the Taylor expansion. The latter ensures that we recover the SHM with NFW profiles on small scales/large k's.

We find that the following parametrization for $f(x)$ satisfies these requirements:

$$f(x) = \frac{ax^2 + bx^3 + dx^4}{1 + cx^3 + dx^4}, \tag{23}$$

where the best-fit parameter values for a, b, c, and d are listed in Table 2. The fits are found by minimizing the root-mean-squared of relative errors, defined as:

$$\text{Error} \equiv 2\sqrt{\left\langle \left[\frac{P(k)_i^{\text{sim}} - P(k)_i^{\text{model}}}{P(k)_i^{\text{sim}}} \right]^2 \right\rangle_i}, \tag{24}$$

where the average is over the simulated data points in k-space. Meanwhile, $P(k)_i^{\text{sim}}$ is the power spectrum from simulation data, and $P(k)_i^{\text{model}}$ is the

	a	b	c	d
WMAP1	0.018	10.5	10.7	3.03
WMAP3	1.94	20.2	21.6	0.0034
WMAP5	0.453	18.5	19.0	0.0055
WMAP7	0.577	18.4	18.9	0.0286
Average	0.747	16.9	17.55	0.767
Stn Dev.	0.72	3.76	14.10	1.3

Table 2: Table: Fitting parameters for (23) for the different WMAP cosmologies from [35], with the average and standard deviation for each parameter listed in the last two rows.

theoretical power spectrum from either Eqn. (4) or Eqn. (21), depending on whether we are finding the error for the SHM or the AHM.

The first parameter a also has the physical significance of being related to the second moment of the compensated halo profile, i.e. Taylor expanding Eqn. (15) in $\mathbf{k}$, we can see that:

$$a = -\frac{\bar{\rho}}{6M^j r_s^2} \int d^3\mathbf{x}|\mathbf{x}|^2 \delta^j_{\rm AHM}(|\mathbf{x}||M), \tag{25}$$

where $\delta^j_{\rm AHM}$ is the same δ^j from Eqn. (20) above.

Since the compensated halo profile is overdense in the middle, and underdense in the outskirts, we expect the 2nd moment to be negative, and thus $a > 0$, as seen in our best fits in Table 2.

The resulting mean relative errors for different simulations are summarized in Figure 4. Table 3 compares the relative errors with those of HALOFIT, assuming that we use the best-fit parameters from Table 2 for each simulation. We see that, while we achieve smaller errors compared to HALOFIT, we also have more parameters per simulation (4/sim for this work, versus 35/ 16 sims in [35]). If we fix all parameters to their average over 4 simulations, effectively having 1 parameter per simulation, Table 4 shows that we generally get larger errors than HALOFIT. Therefore, as a fitting function, AHM using Eqn. (23) has a comparable performance to HALOFIT, since it is based on a more physical underlying framework. We also see that both the AHM and HALOFIT do far better than the SHM in fitting the simulated data.

We can also use the Akaike Information Criterion (AIC) to compare HALOFIT

	Standard	HALOFIT	Amended
WMAP1	0.23	0.053	0.037
WMAP3	0.29	0.042	0.033
WMAP5	0.25	0.034	0.029
WMAP7	0.25	0.037	0.029

Table 3: Mean relative errors for the SHM, HALOFIT, and the AHM for the different WMAP cosmologies in Figure 4, using each cosmology's individually optimized parameters (Table 1).

	Standard	HALOFIT	Amended
WMAP1	0.23	0.053	0.11
WMAP3	0.29	0.042	0.08
WMAP5	0.25	0.034	0.031
WMAP7	0.25	0.037	0.029

Table 4: Errors for the SHM, HALOFIT, and the AHM if we use the average parameters for all the different WMAP cosmologies, instead of their individually optimized ones.

and AHM. The AIC is given as

$$AIC = 2\kappa - 2\ln(\hat{L}) \tag{26}$$

where κ is the number of parameters a model uses, and $\hat{L} = \exp(-\chi^2_{\rm min}/2)$ is its maximum likelihood. The average AIC for our amended model for the individually fitted parameters is around 17,000, while for the mean parameters the amended AIC is around 70,000. The average AIC for Takahashi et al.'s model is around 19,300 - a lot better than the AHM using mean parameters but slightly worse than AHM's individually fitted parameters. However, this comparison should be taken with a grain of salt as we are using the parameters found to minimize the relative error, not the χ^2, to compute AIC.

To get a more physical picture, we can look at the dark matter density that we obtain from Eqn. (21) by using an inverse Fourier transform. On smaller scales, inside the halos (at distances smaller than the halo's r_{200c}), we should roughly recover the NFW density profile. However, outside the virial radius of a halo, we should expect the amended "compensated" profile density to become negative in order to satisfy mass conservation. From Figure 5, we

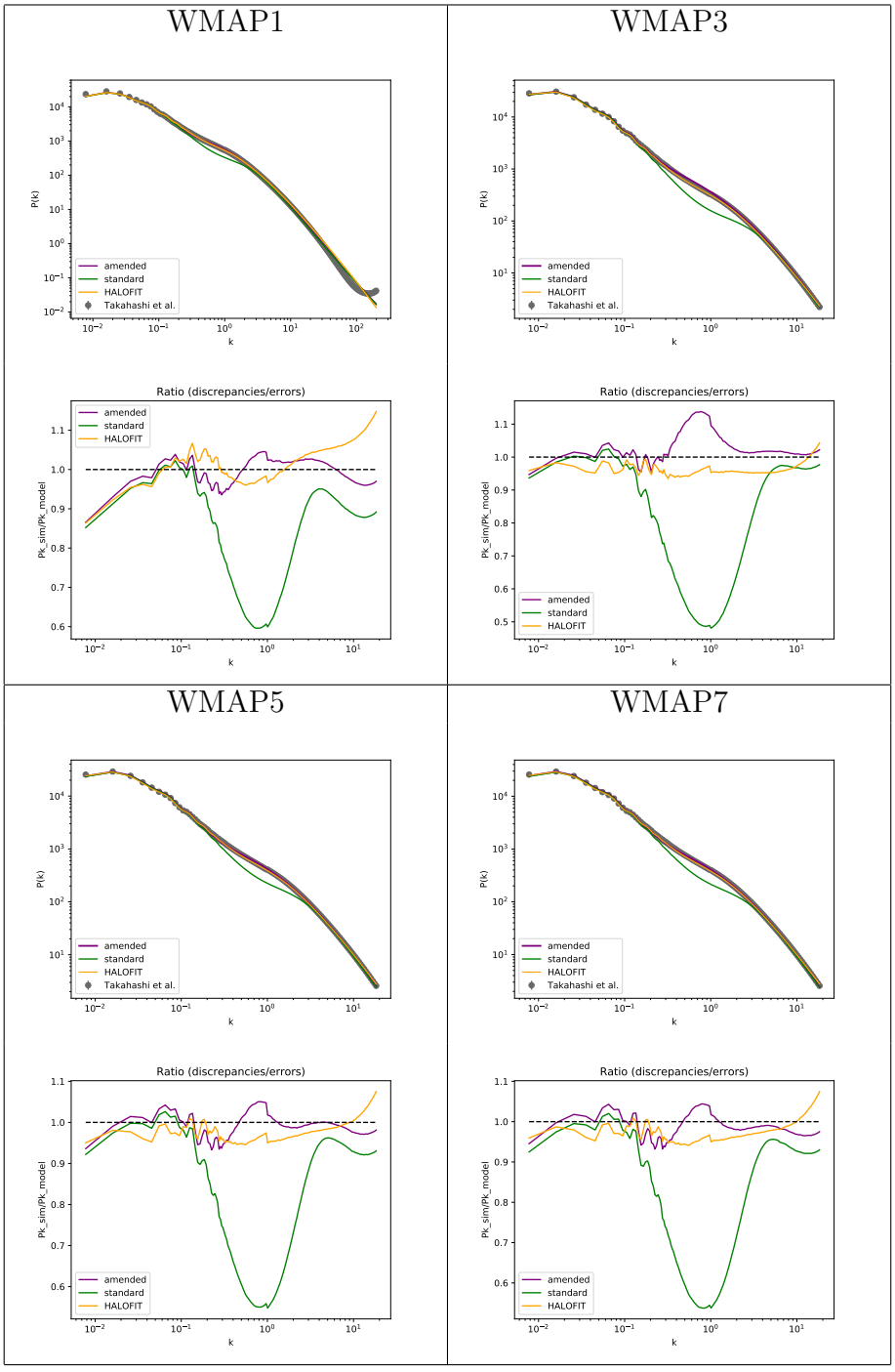

Figure 4: Comparison of Takahashi et. al's simulation data [35] with the standard halo model (SHM), amended halo model (AHM), and HALOFIT. The panels below each $P(k)$ plot show the ratios of model to simulated spectra.

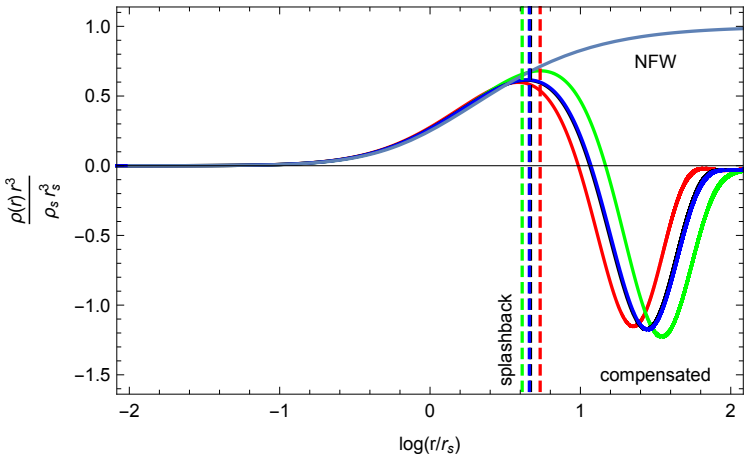

Figure 5: The best-fit compensated density profiles for the amended model, in different cosmologies, compared to the NFW profile. The light blue curve at the top represents the NFW density profile, while the red, green, black, and blue curves are the densities obtained from our amended model for the WMAP1, WMAP3, WMAP5, and WMAP7 cosmologies respectively. The vertical lines are for the splashback radius from More, Diemer, and Kratsov [13, 21], which is close to where density starts decreasing rapidly in our model as well.

see that our density profile matches NFW up to the mean $r_{200c} \sim 6 \times r_s$, which is roughly what we would expect from theory given that the average concentration c is around 6-7. However, it crosses zero and becomes negative at roughly $2 \times r_{200c}$, although the exact value appears to depend on cosmology. The AHM results and comparison with simulations are summarized in Figure 4 and Tables 3-4.

From Figure 4, it can be seen that the deviation from simulation data (from [35]) resulting from our amended model is significantly smaller than the deviation from the standard halo model, indicating that this new modified power spectrum is a better fit for dark matter density in general. When k is a large enough number - k $\geq$ 5 h/Mpc - the amended model, the standard halo model, and the numerical HALOFIT all produce similar results, as we should expect given that HALOFIT and the amended model are supposed to replicate SHM on large k's. However, as seen in Figure 4, AHM and HALOFIT are significantly more accurate than SHM on intermediate, scales with $k \sim 1$ h/Mpc. Physically, this indicates that the current halo model profile does require some compensation to fit the data, similar to what we proposed in Eqn. (17)-(18). This new halo model also conserves mass and fits the simulated dark

matter density better than previous models did, resulting in a new physical model for dark matter clustering on large scales.

Another advantage of the AHM is that, unlike the SHM, it has little sensitivity to including small halos. The reason is that in the SHM, it is assumed that all the mass is included in halos, and therefore convergence of integrals over halo mass requires including relatively small halos. However, in the AHM the halos are compensated (i.e. have zero mass), and thus small halos do not contribute to large scale observables.

2.6 CMB Power Spectra: Amended Model Lensing Predictions

We provide an example of how mass non-conservation can impact observational predictions. In this section we study the weak lensing of CMB maps, that is being measured with unprecedented precision using current and future experiments [4, 32, 25]. To see what power AHM would predict, and show that SHM overpredicts the actual lensing power, we calculate the weak lensing power that should be observed from the CMB [4, 2] using AHM and the extended Limber approximation [20] (also see Appendix B).

This lensing power, as can be seen from Figure 6, matches the measurements from 2018 Planck results [4] fairly well. The standard model seems to overpredict the power (see Figures 6 and 7), especially at small L's, as a result of the large 1-halo contribution to power spectrum at high redshifts. We believe this is primarily because of a lack of mechanism for mass conservation in the SHM. At larger L's in Figure 6, it can be seen that the SHM also does not predict lensing power as well as AHM and HALOFIT, likely due to this being in the transition region where SHM is not very accurate to begin with.

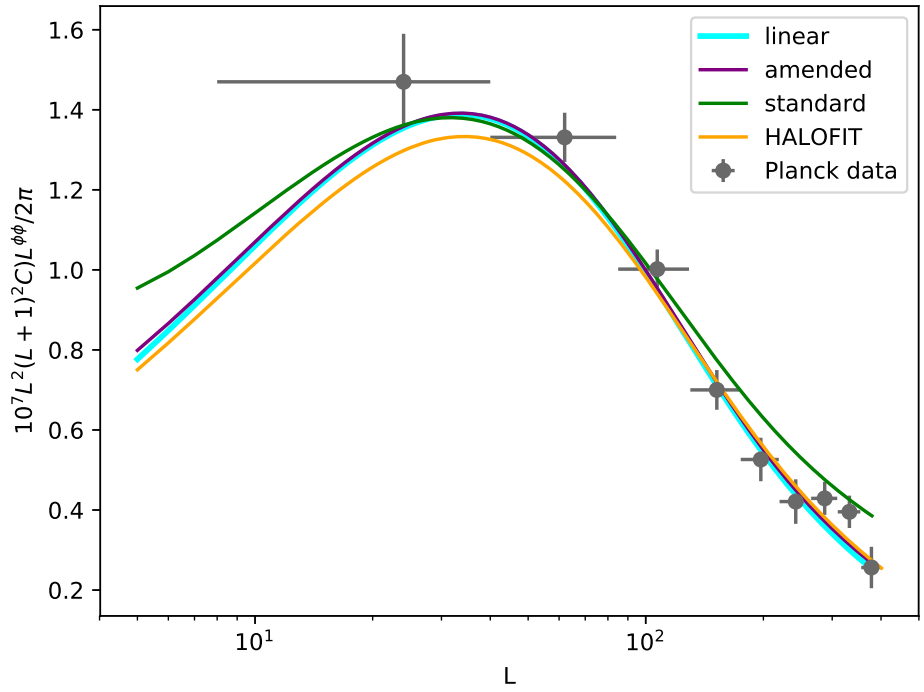

Figure 6: Predictions for CMB gravitational lensing power spectra for our AHM, the SHM, and linear power (while the linear power spectrum/light blue curve is obtained from HALOFIT, this is the initial power spectrum from a high redshift, and not the same one as the nonlinear HALOFIT which is the solid orange curve). The cosmology used here is WMAP7 since it has the closest parameters to the 2018 Planck cosmological parameters [3]. For comparison, we show the measurement of CMB power spectrum from Planck 2018 (plotted as the grey errorbars) [4]. It can be seen that the SHM generally overpredicts power compared HALOFIT and AHM, on larger L's, where L is inverse angular distance (unitless). This corresponds to the transition region in real space, where SHM does not do a good job of predicting the power, and where AHM and HALOFIT work better.

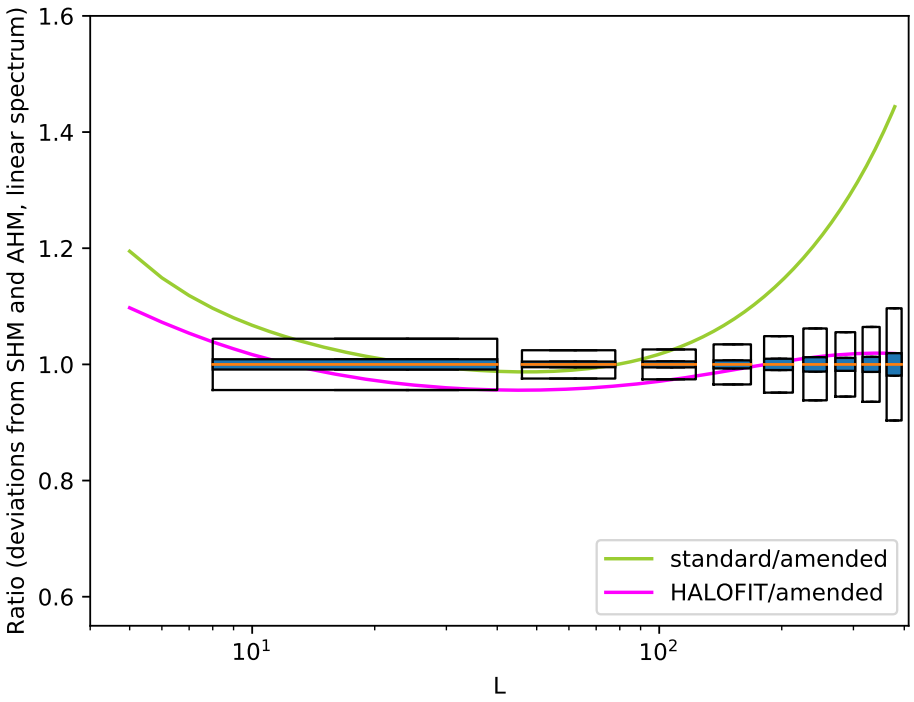

Figure 7: Ratio of the predicted lensing power of the SHM, the AHM, and linear model to each other. It can be seen that the largest discrepancy SHM shows is at $L \lesssim 20$ and $120 \lesssim L$. From Table 1 of [4], if we average the uncertainties, we see that at small L's ($\leq$40), the error should be around 10%, so the SHM errors become too large at $L \lesssim 10$. For large L's, at $250 \lesssim L$, the errors become larger than 20% which is larger than the error predicted for CMB lensing. This shows a significant discrepancy between SHM and AHM, and that SHM would not predict the right observed CMB power at large and small L's.

3 Beyond Mass Density Power Spectrum

3.1 Motivation

So far, we have fitted the AHM to Takahashi et al.'s simulation data [35] to model dark matter density on large scales, in the form of the power spectrum. Now, we want to test this model more thoroughly to see whether it is compatible with other observables, and if it can be used to make predictions about power spectra beyond that of dark matter mass density. To do this, we used simulation data from another source, the Illustris-TNG simulations [1]. We compare the simulation data from TNG100-3-Dark, which uses Planck 2015 cosmology [2], with AHM to see how accurately AHM models the mass power spectrum. Furthermore, we compare the data to the halo power spectrum and the cross correlated power spectrum, which was not done previously.

3.2 Derivation of Cross Power Spectrum from Mass and Halo Spectra

Dark matter halos are related to mass by:

$$\delta_{\mathbf{k}} = \delta_{L,\mathbf{k}} + \sum_j \frac{M^j}{\bar{\rho}} u_{\mathrm{AHM}}(k, M^j) \exp(i\mathbf{k}\cdot\mathbf{x}_j)$$
$$\simeq \left[1 + \frac{1}{\bar{\rho}} \int dM M n(M) b(M) u_{\mathrm{AHM}}(k, M)\right] \delta_L(k), \quad (27)$$

and

$$\delta_{\mathbf{k}}^{\mathrm{halo}} = \frac{1}{\bar{n}_{\mathrm{halo}}} \sum_j \exp(i\mathbf{k}\cdot\mathbf{x}_j) \simeq \frac{\delta_L(k)}{\bar{n}_{\mathrm{halo}}} \int dM n(M) b(M), \quad (28)$$

on large scales.

Now, from our splitting of linear and nonlinear terms for AHM:

$$\langle \delta_{\mathbf{k}} \delta^*_{\mathbf{k}'} \rangle = (2\pi)^3 \delta^3(\mathbf{k} - \mathbf{k}') P(k). \quad (29)$$

Also,

$$\int d^3x \exp(i(\mathbf{k}-\mathbf{k}')\cdot\mathbf{x}) = (2\pi)^3\delta^3(\mathbf{k}-\mathbf{k}'). \tag{30}$$

We revised the matter power spectrum in AHM to be of the form in Eqn. (22):

$$P_{\rm m}(k) = \frac{1}{V}\langle|\delta^2_{\mathbf{k}}|\rangle = \frac{1}{\bar{\rho}^2}\int dMn(M)M^2|u_{\rm AHM}(k|M)|^2 + \left[1+\frac{1}{\bar{\rho}}\int dMMn(M)b(M)u_{\rm AHM}(k|M)\right]^2 P_L(k). \tag{31}$$

Note that theoretically, if the simulation data contains all the halos:

$$\bar{n}_{\rm halo} = \int dMn(M), \tag{32}$$

where $\bar{n}_{\rm halo}$ is the average halo density over all space. However, since generally we cannot resolve all halos, Eqn. (32) usually becomes $\bar{n}_{\rm halo} < \int dMn(M)$ instead.

In addition to the mass power spectrum, we now also look at the halo auto-power spectrum in AHM:

$$P_{\rm halo}(k) = \frac{1}{\bar{n}_{\rm halo}} + \left[\frac{1}{\bar{n}_{\rm halo}}\int_{\text{sim. halos}} dMn(M)b(M)\right]^2 P_L(k), \tag{33}$$

which predicts how dark matter halos are distributed throughout space.

We also look at the matter-halo cross-power spectrum in AHM, which predicts the correlation between mass density distribution and halo density distribution. This takes the form:

$$P_{\rm m\times halo}(k) = \frac{1}{\bar{\rho}\bar{n}_{\rm halo}}\int_{\text{sim. halos}} dMn(M)Mu_{\rm AHM}(k|M) + \left[\frac{1}{\bar{n}_{\rm halo}}\int_{\text{sim. halos}} dMn(M)b(M)\right]\left[1+\frac{1}{\bar{\rho}}\int dMn(M)Mb(M)u_{\rm AHM}(k|M)\right]P_L(k), \tag{34}$$

with wavenumber k spacing:

$$k = \frac{2\pi}{\sqrt[3]{V}}\sqrt{i^2 + j^2 + l^2}, \tag{35}$$

where V is the simulation box volume and i, j, and l run over the dimensions of the grid size of the simulation.

Eqn. (31)-(34) are what we hope to match the simulation data from Illustris-TNG to in the sections below.

3.3 Simulation Data and Preliminary Results

To check the simulation data for reliability, we first take a preliminary look at the halo power spectrum which does not depend on our amended model. We did this by Fourier transforming the halo density from the simulations and matching this density to the k spacing in Eqn. (35). The preliminary results can be seen in Figure 8, which shows the scatter points as matching fairly well with the theoretical model, also showing that we can proceed with the simulation's mass power spectrum and cross power spectrum. Further agreeing with this is Figure 9, which shows the the halo power spectrum deviations mostly fall within the error range of the simulation data. This error range is the variance of the data points for each wavenumber k, as the power is calculated as an average of several data points within each k-bin. There are a couple of outlier points that go outside the error range, but due to the resolution of the data, that is a reasonable occurrence.

Now, the mass power spectrum from AHM, SHM, and HALOFIT are used to match the Illustris-TNG data. The preliminary results are shown in Figures 10 and 11. It can be seen that at small k's, there are a couple of points that go significantly above what the theoretical models would predict - however, we expect around 67% of our predictions to fall within the allowed error range - which the data approximately does.

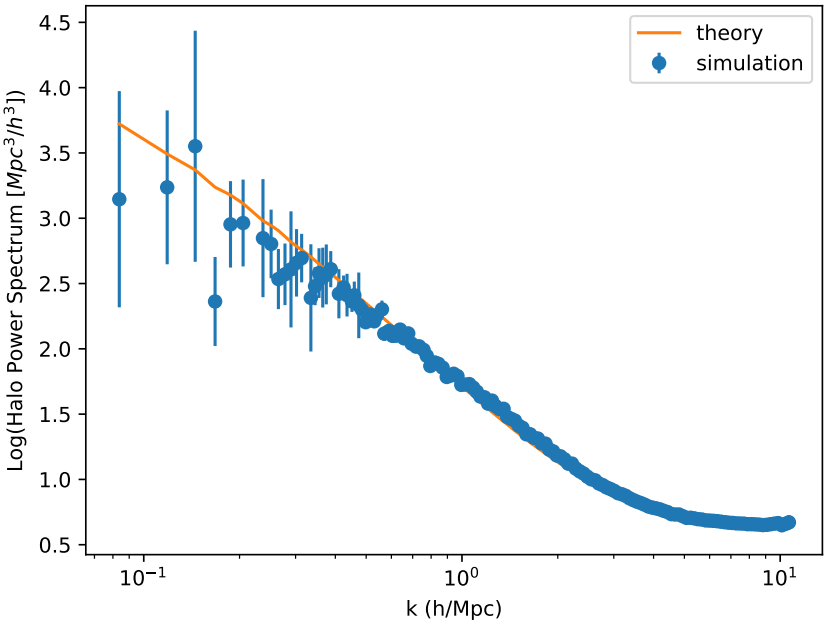

Figure 8: The simulation data of halos compared with the theoretical power in Eqn. (33). The scatter points are from the simulation, with the vertical lines being the errors as the sample variance, and the solid line is the theoretical power.

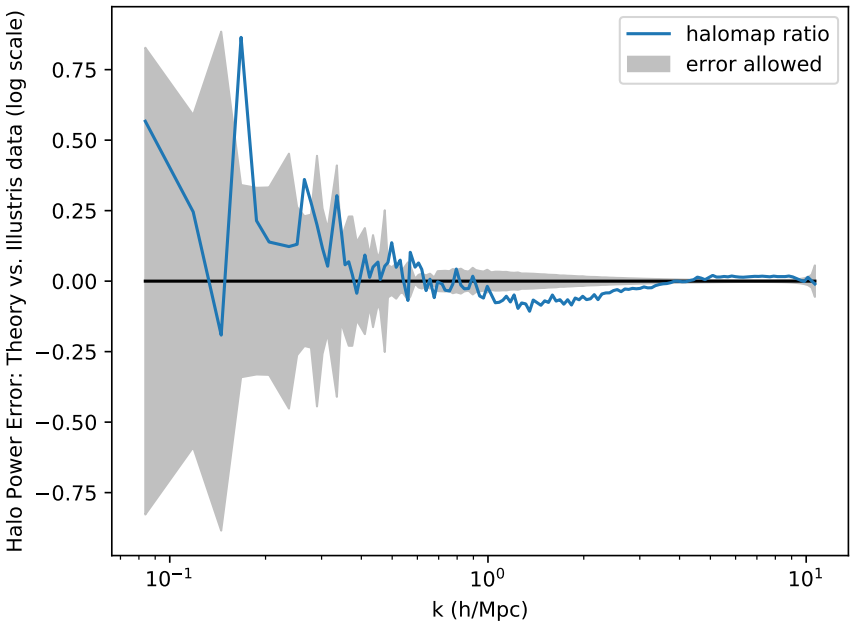

Figure 9: The difference of the log of predicted (Eqn. 33) and the simulated halo spectrum from Illustris.

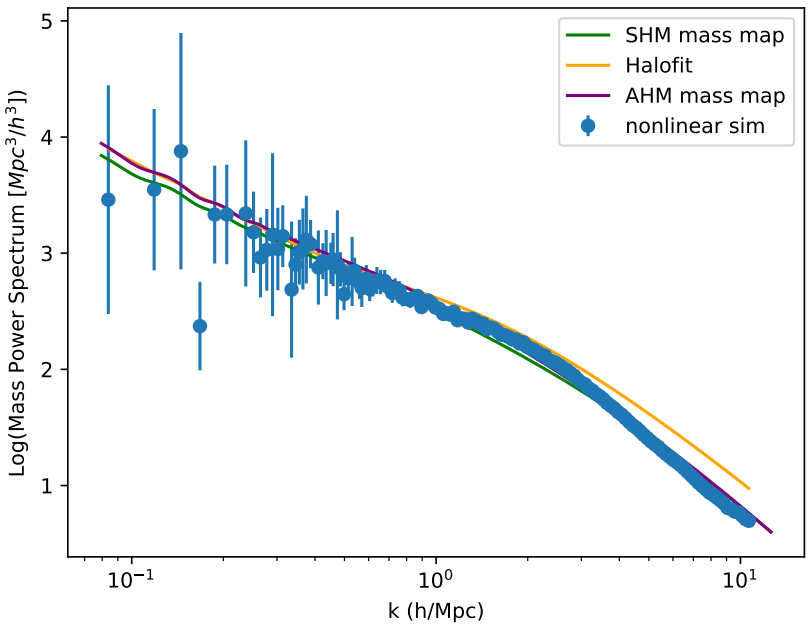

Figure 10: The simulation data of dark matter mass distribution compared with the theoretical power from AHM (Eqn. 22), SHM (Eqn. 4), and HALOFIT (CAMB). The scatter points are from the simulation, with the vertical lines being the errors as the sample variance, and the solid line is the theoretical power.

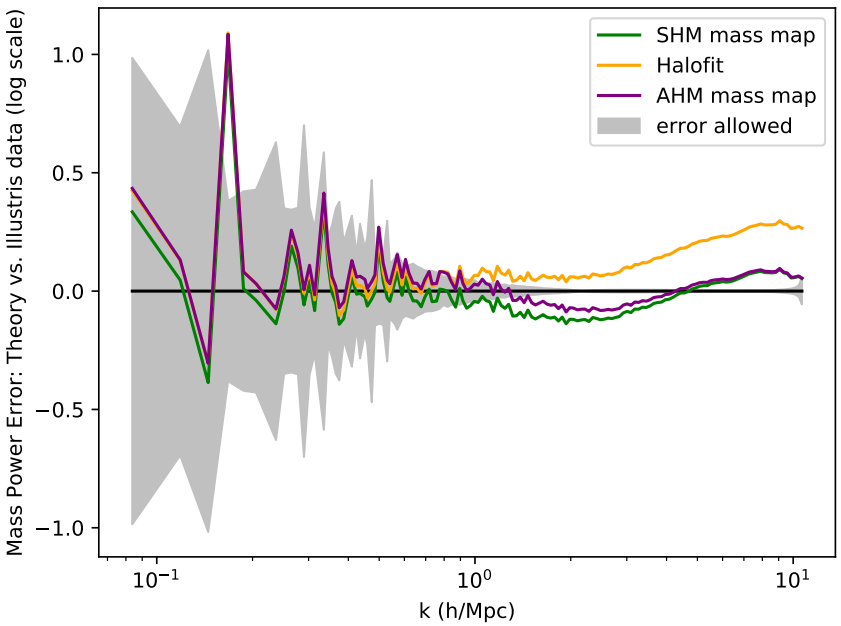

Figure 11: The difference between logs of the predicted (Eqn. 22)) and the simulated mass power spectrum from Illustris. AHM and HALOFIT seem to underestimate the power as $k \geq 2.0$ - this is likely to due the parameters in the fitting function Eqn. (23) not being optimized for this particular cosmology in TNG-3 Dark. This may also be due to the fact that the data is noisy, and the errors may be underestimated.

We now take a first look at the cross-correlated power spectrum, to see how well matter density distribution follows halo distribution. The preliminary results are shown in Figure 12. It can be seen that at around $k \leq 12$, the cross-correlated power starts to overestimate the simulation data, showing that matter density and halo density do not correlate with each other much. From Figure 12, it can be seen that the 2-halo term is suppressed at large $k's$, which would require a modification term to the AHM in Eqn. 34. The modification to the power spectrum to include this effect and its physical explanation is still currently a work in progress.

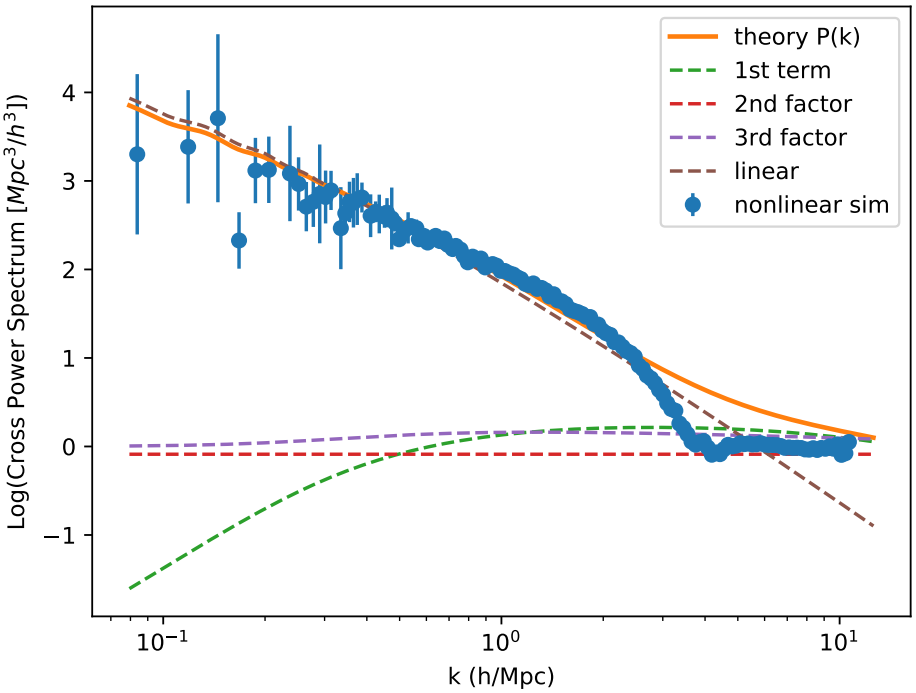

Figure 12: The ratio between Eqn. (34) and the simulation data from Illustris-TNG.

4 Conclusions and Future Prospects

Here, we introduced the AHM of structure formation, which fixes the problem of problem of mass non-conservation in the SHM using a simple and user-friendly framework. The compensated halo profiles in the AHM provide predictions for the matter power spectrum in Eqn. (22) that fit N-body simulation data as well as the parametrized HALOFIT model used in CAMB, and better than the standard halo model for mid to small values of k ($k \leq$ 5 h/Mpc). This leads to an accurate and physical halo model that conserves mass, and fits simulations and theoretical expectations, on both small and large scales.

While the current model introduces the AHM, several future steps can be anticipated:

1. Subhalo structure could in theory be added to the AHM, by adding a subhalo compensation term that is smaller and less significant than our current compensation term (since subhalo effects are smaller). A preliminary comparison of the current AHM (without subhalo terms) to simulation data containing subhalos [1] does show that AHM still models the data fairly well even without a subhalo term.

2. Since our halos are compensated, the lowest order multipole moment is a *dipole*. While the mean dipole would vanish for an average profile, it could have a scatter that contributes to the matter power spectrum on large quasilinear scales. It would be interesting to hunt for this dipole signal in simulations or weak lensing observations.

3. More generally, should (co-)variance of halo profiles be included in the AHM framework, and if so, how?

4. Can we use match-filter methods to directly measure mean compensated halo profiles from N-body simulations?

5. The AHM framework can be further fine-tuned and/or tested using larger boxes, as well as neutrino and/or baryonic effects.

6. Similar to CMB lensing studied here, it would be interesting to see how predictions for the kinematic Sunyaev-Zel'dovich effect and 21-cm intensity mapping at high-z, that may be sensitive to momentum and hydrogen mass conservation, might be impacted.

7. Another potential application of the AHM is capturing environmental dependence of halo properties through cross-terms such as:

$$\sum_j \langle \delta_L(\mathbf{x}')\delta^j(\mathbf{x}-\mathbf{x}^j)\rangle, \tag{36}$$

that contribute to 2-point correlation function (or the power spectrum). This could be further generalized to other tracers, such as galaxies or hot gas, by quantifying how profiles of individual halos may be different for environments with different linear overdensities.

8. The AHM can be used to model 1-point probability distribution function for conserved observables, such as weak lensing convergence, or kinetic Sunayev-Zel'dovich effect (e.g., extending treatment introduced in [37]).

9. Momentum conservation here is a result of mass conservation, but halos can have their own compensated momentum profiles, which is not addressed here. Exploring how this momentum profile will fit into the AHM can give further physical evidence and insights for the need of halo model amendments.

10. So far, we have only looked at amending the power spectrum of the standard halo model, but not at any higher order statistics of this model. Can the AHM be extrapolated to find the bispectrum and trispectrum?

11. In Chapter 3, we introduced the cross-correlated power spectrum, and found that the AHM didn't predict well how dark matter halo density correlated with mass density at large $k's$. Modifying AHM in a way that preserves its success in predicting mass density power but also models cross-correlation power more accurately is a next step.

These points are all aspects of the AHM that can be further explored, and with these we would have a better understanding of how to probe large scale dark matter structures. Further fine-tuning can help make predictions about the characteristics of dark matter itself, depending on how accurately this model can predict dark matter structures.

References

[1] *Illustris TNG100-3-Dark*, 2019.

[2] P. A. R. Ade et al. Planck 2015 results. XV. Gravitational lensing. *Astron. Astrophys.*, 594:A15, 2016.

[3] N. Aghanim et al. Planck 2018 results. VI. Cosmological parameters. 2018.

[4] N. Aghanim et al. Planck 2018 results. VIII. Gravitational lensing. 2018.

[5] Giorgio Arcadi, Maíra Dutra, Pradipta Ghosh, Manfred Lindner, Yann Mambrini, Mathias Pierre, Stefano Profumo, and Farinaldo S. Queiroz. The waning of the WIMP? A review of models, searches, and constraints. *Eur. Phys. J. C*, 78(3):203, 2018.

[6] Kenath Arun, S.B. Gudennavar, and C. Sivaram. Dark matter, dark energy, and alternate models: A review. *Advances in Space Research*, 60(1):166–186, Jul 2017.

[7] J. M. Bardeen, J. R. Bond, N. Kaiser, and A. S. Szalay. The Statistics of Peaks of Gaussian Random Fields. , 304:15, May 1986.

[8] Gianfranco Bertone. *Particle Dark Matter : Observations, Models and Searches*. 2010.

[9] Pasquale Blasi, Rainer Dick, and Edward W. Kolb. Ultra-High Energy Cosmic Rays from Annihilation of Superheavy Dark Matter. *Astropart. Phys.*, 18:57–66, 2002.

[10] Latham Boyle, Kieran Finn, and Neil Turok. CPT-Symmetric Universe. *Phys. Rev. Lett.*, 121(25):251301, 2018.

[11] Bernard Carr and Florian Kuhnel. Primordial Black Holes as Dark Matter: Recent Developments. 6 2020.

[12] Asantha Cooray and Ravi K. Sheth. Halo Models of Large Scale Structure. *Phys. Rept.*, 372:1–129, 2002.

[13] Benedikt Diemer and Andrey V. Kravtsov. Dependence of the outer density profiles of halos on their mass accretion rate. *Astrophys. J.*, 789:1, 2014.

[14] Dimitry Ginzburg, Vincent Desjacques, and Kwan Chuen Chan. Shot noise and biased tracers: A new look at the halo model. , 96(8):083528, Oct 2017.

[15] Alister W. Graham, David Merritt, Ben Moore, Juerg Diemand, and Balsa Terzic. Empirical models for Dark Matter Halos. I. Nonparametric Construction of Density Profiles and Comparison with Parametric Models. *Astron. J.*, 132:2685–2700, 2006.

[16] Daniel Grin, Mustafa A. Amin, Vera Gluscevic, Renée Hl˘ ozek, David J.E. Marsh, Vivian Poulin, Chanda Prescod-Weinstein, and Tristan L. Smith. Gravitational probes of ultra-light axions. 4 2019.

[17] Boryana Hadzhiyska, Sownak Bose, Daniel Eisenstein, Lars Hernquist, and David N. Spergel. Limitations to the "basic" HOD model and beyond. 2019.

[18] Nick Hand, Uros Seljak, Florian Beutler, and Zvonimir Vlah. Extending the modeling of the anisotropic galaxy power spectrum to $k = 0.4\ h\mathrm{Mpc}^{-1}$. *JCAP*, 1710(10):009, 2017.

[19] Antony Lewis, Anthony Challinor, and Anthony Lasenby. Efficient computation of CMB anisotropies in closed FRW models. , 538:473–476, 2000.

[20] Marilena LoVerde and Niayesh Afshordi. Extended Limber Approximation. *Phys. Rev.*, D78:123506, 2008.

[21] Surhud More, Benedikt Diemer, and Andrey Kravtsov. The splashback radius as a physical halo boundary and the growth of halo mass. *Astrophys. J.*, 810(1):36, 2015.

[22] Julio F. Navarro, Carlos S. Frenk, and Simon D. M. White. The Structure of cold dark matter halos. *Astrophys. J.*, 462:563–575, 1996.

[23] Hiroko Niikura et al. Microlensing constraints on primordial black holes with Subaru/HSC Andromeda observations. *Nature Astron.*, 3(6):524–534, 2019.

[24] Chiamaka Okoli and Niayesh Afshordi. Concentration, Ellipsoidal Collapse, and the Densest Dark Matter haloes. *Mon. Not. Roy. Astron. Soc.*, 456(3):3068–3078, 2016.

[25] Y. Omori et al. A 2500 deg^2 CMB Lensing Map from Combined South Pole Telescope and Planck Data. *Astrophys. J.*, 849(2):124, 2017.

[26] J. A. Peacock and R. E. Smith. Halo occupation numbers and galaxy bias. *Mon. Not. Roy. Astron. Soc.*, 318:1144, 2000.

[27] Oliver H.E. Philcox, David N. Spergel, and Francisco Villaescusa-Navarro. The Effective Halo Model: Creating a Physical and Accurate Model of the Matter Power Spectrum and Cluster Counts. 4 2020.

[28] B. Ryden. *Introduction to Cosmology*. Addison Wesley, 2003.

[29] Fabian Schmidt. Towards a self-consistent halo model for the nonlinear large-scale structure. *Phys. Rev.*, D93(6):063512, 2016.

[30] Uroš Seljak. Analytic model for galaxy and dark matter clustering. *Monthly Notices of the RAS*, 318(1):203–213, Oct 2000.

[31] Uroš Seljak and Zvonimir Vlah. Halo Zel'dovich model and perturbation theory: Dark matter power spectrum and correlation function. *Phys. Rev.*, D91(12):123516, 2015.

[32] Blake D. Sherwin et al. Two-season Atacama Cosmology Telescope polarimeter lensing power spectrum. *Phys. Rev.*, D95(12):123529, 2017.

[33] Manodeep Sinha, Andreas A. Berlind, Cameron K. McBride, Roman Scoccimarro, Jennifer A. Piscionere, and Benjamin D. Wibking. Towards accurate modelling of galaxy clustering on small scales: testing the standard ΛCDM + halo model. *Monthly Notices of the RAS*, 478(1):1042–1064, Jul 2018.

[34] H. Sugai, P. A. R. Ade, Y. Akiba, D. Alonso, K. Arnold, J. Aumont, J. Austermann, C. Baccigalupi, A. J. Banday, R. Banerji, and et al. Updated design of the cmb polarization experiment satellite litebird. *Journal of Low Temperature Physics*, 199(3-4):1107–1117, Jan 2020.

[35] Ryuichi Takahashi, Masanori Sato, Takahiro Nishimichi, Atsushi Taruya, and Masamune Oguri. Revising the Halofit Model for the Nonlinear Matter Power Spectrum. *Astrophys. J.*, 761:152, 2012.

[36] Hugo Tavio, Antonio J. Cuesta, Francisco Prada, Anatoly A. Klypin, and Miguel A. Sanchez-Conde. The dark outside: the density profile of dark matter haloes beyond the virial radius. 2008.

[37] Leander Thiele, J. Colin Hill, and Kendrick M. Smith. Accurate analytic model for the thermal Sunyaev-Zel'dovich one-point probability distribution function. *Phys. Rev.*, D99(10):103511, 2019.

[38] Jeremy L. Tinker, Andrey V. Kravtsov, Anatoly Klypin, Kevork Abazajian, Michael S. Warren, Gustavo Yepes, Stefan Gottlober, and Daniel E. Holz. Toward a halo mass function for precision cosmology: The Limits of universality. *Astrophys. J.*, 688:709–728, 2008.

[39] Jeremy L. Tinker, Brant E. Robertson, Andrey V. Kravtsov, Anatoly Klypin, Michael S. Warren, Gustavo Yepes, and Stefan Gottlöber. The Large-scale Bias of Dark Matter Halos: Numerical Calibration and Model Tests. , 724(2):878–886, Dec 2010.

[40] Diego F. Torres and Luis A. Anchordoqui. Astrophysical origins of ultrahigh energy cosmic rays. *Rept. Prog. Phys.*, 67:1663–1730, 2004.

Appendix A Power Spectrum Derivation

The standard halo model power spectrum is given by Eq. (4) while the amended model spectrum is given by Eq. (22). Both of these equations are dependent on functions such as the bias, number density, and linear power spectrum.

The linear power spectrum is a measure of the matter density at early times, when matter was distributed fairly uniformly. It is most accurately calculated using the numerical CAMB package in Python [26], but can also be approximated by [7],

$$P_{lin}(x) = 2\pi\omega_N \frac{k^{n_s}}{H_0^{n_s+3}} T(x)^2, \tag{37}$$

where n_s is the spectral index, H_0 is the Hubble constant, ω_N is a normalization constant, and T(x) is the Bardeen-Bond-Kaiser-Szalay (BBKS) transfer function given by

$$T(x) = \frac{ln(1+0.171x)}{0.171x}\big(1{+}0.284x{+}(1.18x)^2{+}(0.399x)^3{+}(0.490x)^4\big)^{-1/4}. \tag{38}$$

The number density and bias functions from Eqn. (4), (22), and (34) are from Tinker et al.'s papers [38, 39], and they can be seen quantitatively in Figures 13 and 14. The number density is given by

$$\frac{dn}{dM} = f(\sigma)\frac{\rho_m}{M}\frac{ln\sigma^{-1}}{dM}, \tag{39}$$

where M is the halo mass range, σ is a function defined as

$$\sigma = \int P(k)_{lin}\hat{W}(kR)k^2 dk, \tag{40}$$

where $P(k)_{lin}$ is the linear power spectrum and $\hat{W}(kR)$ is the Fourier transform of the real- space top-hat window function of a halo of radius R [38], given by

$$\hat{W}(kR) = \frac{3}{k^3R^3}[\sin(kR) - kR\cos(kR)]. \tag{41}$$

The function u$(k|m)$ is the Fourier transform of the NFW profile in Eq. (5), which is given mathematically by Eq. (16).

$f(\sigma)$ in Eq. (39) is a universal parametric function subject to changes in redshift and cosmological parameters, and it is of the form

$$f(\sigma) = A\left[\left(\frac{\sigma}{b}\right)^{-a} + 1\right]e^{-\frac{c}{\sigma^2}}, \tag{42}$$

where the values of A, a, b, and c are given in Table 2 of Tinker et al. [38].

The bias function from [39], which predicts where peaks in the overdense regions are mostly likely to be found, is of the form

$$b(\nu) = 1 - A_{bias}\frac{\nu^{a_{bias}}}{\nu^{a_{bias}} + \delta_c^{a_{bias}}} + B\nu^{b_{bias}} + C\nu^{c_{bias}}, \tag{43}$$

where the parameters A_{bias}, a_{bias}, B, b_{bias}, C and c_{bias} are given by Table 2 in [39].

Note that A_{bias}, a_{bias}, b_{bias}, and c_{bias} here are **different** from the number density parameters A, a, b, c above.

For the AHM, we used an overdensity of $200/\Omega_m$, or approximately 680, to find the corresponding fitting parameters for the bias and number density functions. The results are summarized in Figure 13 and 14 below.

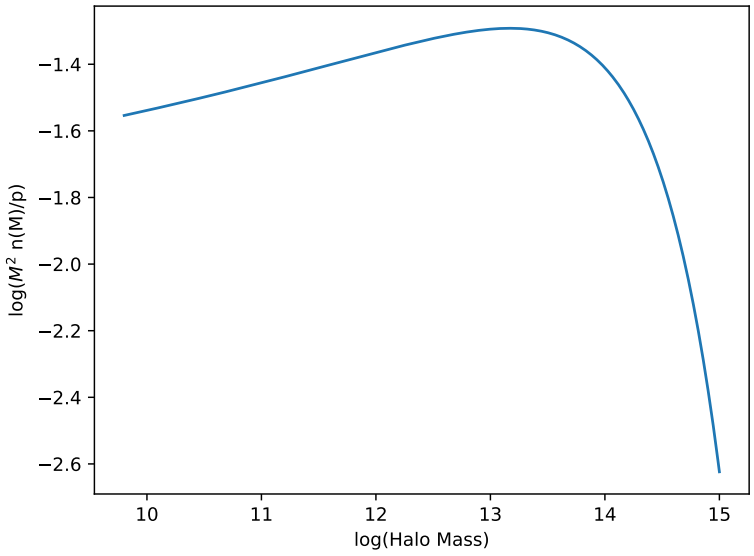

Figure 13: The number density using the fitting function in [38] and the overdensity mentioned above.

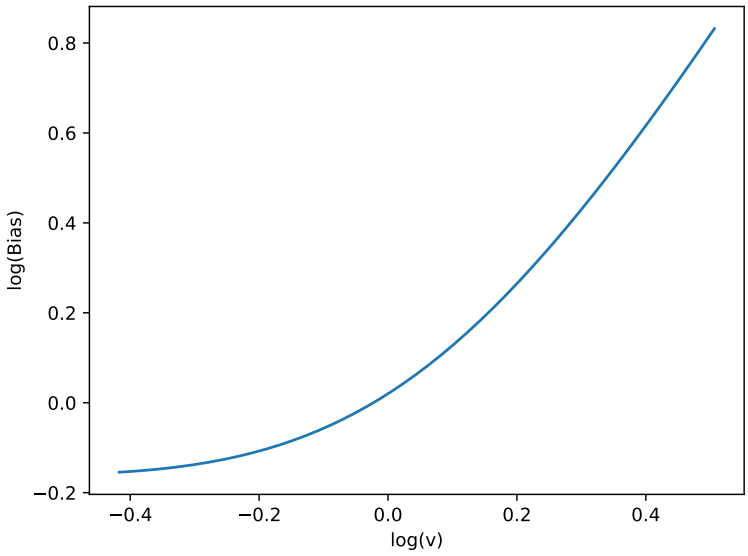

Figure 14: The bias function using the fitting function in [39] and the overdensity mentioned above.

Appendix B Limber Approximation

One relatively quick and simple way to estimate the power of CMB lensing is by using the extended Limber approximation from LoVerde and Afshordi (2008) [20]. This approximation calculates the angular power spectrum as a series expansion in $(l+\frac{1}{2})^{-1}$, with the underlying assumptions that the angular separations are small and the power spectrum is isotropic [20]. Additionally, we assume a flat metric, as it approximates the observable universe.

Mathematical Derivation

We first start with the gravitational potential:

$$\phi(\hat{n}) = -2\int_0^{r_*} \Psi[r\hat{n}, \eta_0 - r](-z)dr, \tag{44}$$

where we use Poisson's equation

$$\nabla^2\Psi = \frac{4\pi G\rho_m}{a}\delta_m, \tag{45}$$

with overdensity

$$\delta_m = \frac{\rho_m}{\bar{\rho}_m} - 1. \tag{46}$$

As following the notation used in the rest of this thesis, ρ_m is the matter density at a certain location in the sky while $\bar{\rho}_m$ is the approximated average matter density of the universe (or in simulations, average density in the simulation volume), given by

$$\bar{\rho}_m = \Omega_{m0}\frac{3H_0^2}{8\pi G}, \tag{47}$$

where H_0 is the Hubble constant at redshift z=0 and Ω_{m0} is the matter density paramter at z=0. Also in Eqn. (45) and (47), G is the gravitational constant, $6.647\times10^{-11}\mathrm{m}^3/\mathrm{kg}\cdot\mathrm{s}^2$ in SI units, and a is the relation between scale factor and redshift, a=$\frac{1}{z+1}$.

Then, applying the Limber approximation [20]:

$$k = \frac{L+\frac{1}{2}}{r} \tag{48}$$

with $k \sim 1/r$ we get,

$$\frac{[L(L+1)]^2}{2\pi} C_L^{\phi\phi} = 4\frac{[L(L+1)]^2}{2\pi} \int_0^{r_*} \frac{(r_* - r)^2}{r^4 r_*^2} P_{\psi\psi} dr, \tag{49}$$

where $C^{\phi\phi}$ is the observed angular lensing power and the CMB lensing power $P_{\psi\psi}$ is defined as

$$P_{\psi\psi} = \frac{(4\pi G\bar{\rho}_m)^2}{k^4 a^2} P_{nonlinear}(k), \tag{50}$$

and

$$r = \frac{c}{H_0} \int \frac{dz}{\sqrt{\Omega_{m0}(1+z)^3 + 1 - \Omega_{m0}}}, \tag{51}$$

r_* is the radius at a redshift of z=1000. It should be noted here that there is no one-to-one correspondence between inverse angular distance L and wavenumber k, given that r can be the same for different coordinates ($r = \sqrt{x^2 + y^2 + z^2}$, degeneracy in r).

The constants in Eqn. (50) can be simplified using Eqn. (47) as

$$4\pi G\bar{\rho_m} = 4\pi G\left(\Omega_{m0}\frac{3H_0^2}{8\pi G}\right) = \frac{3}{2}H_0^2\Omega_{m0}, \tag{52}$$

and subsequently,

$$(4\pi G\bar{\rho_m})^2 = \frac{9}{4}H_0^4\Omega_{m0}^2 \tag{53}$$

As a result, Eqn. (49) becomes

$$\frac{[L(L+1)]^2}{2\pi} C_L^{\phi\phi} = \frac{4}{2\pi} \int_0^{r_*} \frac{(r_* - r)^2}{r^4 r_*^2} \frac{9H_0^4\Omega_{m0}^2}{4k^4 a^2} P_{nl}(k) dr, \tag{54}$$

which can then be simplified to,

$$\frac{[L(L+1)]^2}{2\pi} C_L^{\phi\phi} = \frac{9}{2\pi} H_0^4\Omega_{m0}^2 \int_0^{r_*} (1+z)^2 \frac{(r_* - r)^2}{r_*^2} P_{nl}(k) dr \tag{55}$$

from Eqn. (50), (51), and (53). The constant $\frac{c}{H_0}$ is equal to 3000 Mpc/h in the model that we used (AHM), and agrees within the range of Planck data observations, as seen above in Fig. 6.