Universidade de São Paulo
Instituto de Física

# O spin bias dos halos de matéria escura

Beatriz Tucci Schiewaldt

Orientador: Prof. Dr. Luis Raul Weber Abramo

Dissertação de mestrado apresentada ao Instituto de Física da Universidade de São Paulo, como requisito parcial para a obtenção do título de Mestra em Ciências.

Banca Examinadora:
Prof. Dr. Luis Raul Weber Abramo (Universidade de São Paulo)
Prof. Dr. Ravi Kiran Sheth (University of Pennsylvania)
Andrés Balaguera Antolínez (Instituto de Astrofisica de Canarias)

São Paulo
2021

**FICHA CATALOGRÁFICA**
**Preparada pelo Serviço de Biblioteca e Informação**
**do Instituto de Física da Universidade de São Paulo**

Schiewaldt, Beatriz Tucci

O spin Bias dos halos da matéria escura São Paulo, 2021.

Dissertação (Mestrado) – Universidade de São Paulo. Instituto de Física. Depto. de Física Matemática

Orientador: Prof. Dr. Luis Raul Weber Abramo

Área de Concentração: Física

Unitermos: 1. Cosmologia; 2. Astrofísica; 3. Relatividade (Física).

USP/IF/SBI-074/2021

University of São Paulo
Physics Institute

# The spin bias of dark matter halos

Beatriz Tucci Schiewaldt

Supervisor: Prof. Dr. Luis Raul Weber Abramo

Dissertation submitted to the Physics Institute of the University of São Paulo in partial fulfillment of the requirements for the degree of Master of Science.

Examining Committee:
Prof. Dr. Luis Raul Weber Abramo (Universidade de São Paulo)
Prof. Dr. Ravi Kiran Sheth (University of Pennsylvania)
Andrés Balaguera Antolínez (Instituto de Astrofisica de Canarias)

São Paulo
2021

# Acknowledgements

First I would like to thank my family, who has allowed me to pursue my goals and encouraged my will to do science. I am deeply thankful to my late father Livio Schiewaldt, with whom I used to write physics formulae with coal on the barbecue (that he constructed by his own on our garden, brick by brick). The invaluable and faithful companion of my sister Clara Lis Tucci Schiewaldt, who is always by my side and gives me willingness to face each new step in life. The early youth conversations I had with my brother Yuri Tucci Eterovic, when we discovered the Universe as a unit and debated about society as a whole. The daily support of my mother Lisange Tucci and of Marcelo Jarufe, with whom I know I can sincerely share all my fears and victories. My cats Rá, Om, Ravena and Mimmy, who make my life warmer. My grandparents Zoraide Vinha Schiewaldt, Oscar Schiewaldt, Luisa Tucci and Weber Angelo Tucci, my uncle Mauro Tucci, who gave me all support I needed to reach here.

I am especially thankful for the support of my boyfriend Rodrigo Voivodic, with whom I am delighted to share my life.

I would like to thank all my friends with whom I passed wonderful times. From the Molecular Sciences course, I thank Lucas Magno for guiding me as a physicist from the very beginning of my path to this day, Nickolas Kokron who gave me valuable tips on Cosmology and on doing my PhD abroad, Gabriela Sato Polito for the help in choosing and starting my scientific research. My friends David Berl, Gabriel Couto, José Arthur, Nelson Schuback, Pietro Morgante.

I would also like to thank my friends from the Physics Institute, where I found a new home. André Rodrigues, Breno Agatão, Bruno Eduardo, Danilo Bissoli, Gabriel Golfetti, Hugo Portelinha, Iago Mamede, Lidia Ascon, Mariana Mercucci, Nickolas Alves, Pedro Henrique Tredezini, Theo Meirelles. The friends from the Cosmology group, to whom I owe a lot of my knowledge on Cosmology and with whom I shared great experiences. Caroline Guandalin, Francisco Maion, Natália Villa Nova Rodrigues, Natalí de Santi, Renan Boschetti, Thiago Mergulhão.

Also special thanks to my dear singing teacher and friend Nathalia Lemes, my phsychologist Izabelly Ribeiro and to my fellow traveler Vitor Tonaco.

I am deeply thankful to the learning, inspiration and motivation from the professors and researches who crossed my path. Enrico Bertuzzo, Fabian Schmidt, Fabio Iocco, Facundo Rodriguez, George Matsas, Guilherme Pimentel, Marcello Musso, Marcos Lima, Maria Celeste, Mehrdad Mirbabayi, Neal Dalal, Nelson Padilla, Nima Arkani-Hamed, Paulo Nussenzveig, Raúl Angulo, Ravi Sheth.

The work presented here was developed with the valuable help of my supervisors. I am most grateful to the enlightening conversations with Raul Abramo, who has always given me all the academic and life support, and with Antonio Montero-Dorta, who besides not being my official supervisor has helped me not only to develop my research but also to grow as a professional physicist.

Finally I would like to thank São Paulo Research Foundation (FAPESP) for financial support during my undergraduate research project (grant 2018/17688-9) and also during my Master's (grant 2020/10520-5).

# Resumo

A dependência do *bias* de halos de matéria escura com o seu spin para uma massa fixa é conhecida por spin *bias*. Embora os halos mais massivos de maior spin tenham um *bias* maior do que os seus correspondentes de mesma massa, o efeito se inverte para os halos de menores massas. Nós recentemente esclarecemos este cenário complexo, demonstrando que a inversão do spin *bias* pode ser completamente explicada pela população de halos *splashback* – i.e., halos que em algum momento do passado estiveram no interior do raio de virial de um outro halo (tipicamente muito mais massivo), mas acabaram saindo deste. O fato de que os halos *splashback* causam a inversão do spin *bias* foi primeiramente explicado em Tucci et al. (2021). Nós mostramos que, quando essa população específica é excluída da amostra, somente o spin *bias* intrínseco permanece – i.e., um *bias* maior para os halos de maior spin. Agora que nós entendemos a inversão em massas pequenas, nosso objetivo é investigar as origens físicas do spin *bias* intrínseco. A fim de conectar o spin do halo ao seu *bias* em um único modelo analítico, investigamos as principais teorias para o momento angular e o bias dos halos. Também exploramos como esse mecanismo, que é normalmente estudado em simulações de N-corpos, pode ser estendido para galáxias no contexto da conexão entre halos e galáxias e de formação de galáxias. Finalmente, exploramos as possíveis consequências para a cosmologia observacional.

**Palavras chave:** cosmologia, estruturas em larga escala do Universo, simulações de N-corpos, halos de matéria escura, galáxias, bias

# Abstract

The dependence of dark matter halo bias on spin at fixed mass is known as spin bias. However, although at the high-mass end the higher-spin halos have a higher bias than their lower-spin counterparts, this trend inverts for low-mass halos. We have recently clarified this complex scenario, showing that this inversion of the mass dependence of spin bias can be completely explained by the population of splashback halos – i.e., halos that at some point in the past fell inside the virial radius of another halo (typically, a larger-mass halo), but then exited that parent halo. This dependence of spin bias on splashback halos was first explained in the recent paper Tucci et al. (2021). What we found is that, when this specific population is excluded from the sample, only the intrinsic spin bias signal remains – i.e., a higher bias for higher-spin halos. Now that we understood the low-mass feature, our goal is to shed light onto the physical origins of the high-mass, intrinsic spin bias. In order to connect halo spin to its bias in a single analytical framework, we investigate the main theories for halo angular momentum and halo bias. We also explore how this mechanism, which is more readily studied via halos in N-body simulations, can be extended to galaxies in the context of the halo-galaxy connection framework and galaxy formation. Finally, we study possible consequences for observational cosmology.

**Keywords:** cosmology, large-scale structure of Universe, N-body simulations, dark matter halos, galaxies, bias

# Contents

# List of Figures

# CHAPTER 1

# Introduction

There is ample evidence for the existence of a cold (i.e., non-relativistic) and dark component of matter (CDM). As its own name suggests, dark matter is a kind of matter which interacts gravitationally and, as far as we know, it does not interact through the electromagnetic or other fundamental forces. Together with the cosmological constant $\Lambda$, which is the simplest way to account for the accelerated expansion of the Universe, as well as the other cosmological parameters, CDM forms the basis of the standard $\Lambda$-CDM model (Planck Collaboration et al., 2018).

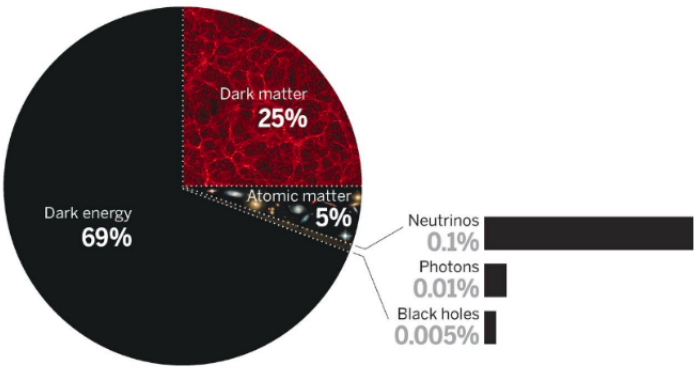


Figure 1.1: Relative abundance of the energy content of the Universe.

Despite its remarkable success, $\Lambda$-CDM does not explain neither the nature of dark matter nor that of $\Lambda$, amongst other thrilling open questions. Is Einstein's General Relativity the complete theory of gravity? Why the Universe is expanding in an accelerated way? At the beginning, were there other particles interacting with the inflaton (the scalar field believed to be responsible for inflation, which is the epoch of accelerated expansion during the first $\sim 10^{-33}$ seconds after the Big Bang, which made the Universe flat and homogeneous)? Why is there more matter than anti-matter in the Universe? What are the Neutrinos masses? Cosmological observations coming up in the near future are expected to shed light into many of these puzzles – and hopefully to find new physics.

A vast amount of data will be collected in the next years by cosmological surveys (DESI Collaboration et al., 2016; LSST Science Collaboration et al., 2009; Amendola et al., 2013), which will map with unprecedented precision the large scale structure (LSS) of the Universe, i.e., the structure formed by the distribution of galaxies and other extragalactic objects on large scales (i.e., over distances larger than $\sim 10^2$ Megaparsecs). Perhaps most surprisingly, the particular way that galaxies are distributed hides information about all the questions raised in the last paragraph.

A key feature of the next generation of surveys is that they will map a huge number of objects, allowing for precise statistics, together with the fact that we will finally be able to see how the Universe looks like on ultra large scales. We do not have accurate tests of General Relativity on such scales yet, and it is there where we expect to see deviations from Einstein's theory of gravity, if they exist at all (Parfrey et al., 2011). Primordial non-Gaussianities, which reflect the existence of non-trivial interactions of the inflation with itself or with other particles during inflation, leave an imprint on the statistics of galaxies on those ultra large scales (Dalal et al., 2008). Neutrinos also affect the clustering of galaxies, leading to a particular scale-dependent signal that could tell us their masses (Castorina et al., 2014).

We cosmologists now have to devote great efforts to develop our understanding of structure formation in the Universe, so that we can maximize our ability to extract information from the upcoming surveys and ultimately improve our knowledge on the fundamental laws governing Nature.

## 1.1 The Large Scale Structure

We believe that inflation gave us the initial conditions of the Universe, such that after inflation and the subsequent reheating phase it is near-homogeneous, and filled with the particles described by the standard model of particle physics, such as baryons, as well as dark matter. The small density fluctuations generated by inflation are characterized by a Gaussian random field, and the peaks in the that initial density field grow as particles are brought together by gravitational interactions, forming deep potential wells. When these overdensities reach a certain threshold, they effectively decouple from the expansion of the Universe and collapse into gravitationally bound structures known as *dark matter halos*. Naturally, halos differ from one to another in terms of properties such as mass, spin, age, concentration, etc.

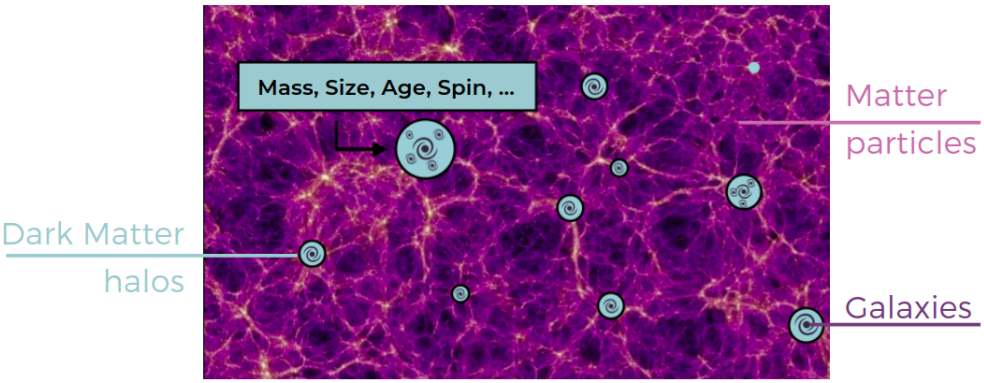


Figure 1.2: Slice of a Millenium N-body cosmological simulation characterizing three different layers of the large scale structure.

The process of halo formation starts early in the Universe for the dark matter component, but baryons are also affected by their tight coupling to radiation, which pushes them out of the heavy and hot overdense regions. This happens until the end of the radiation era, and when baryons finally decouple from photons in the recombination era (approximately 400.00 years after the Big Bang), they are able to fall into the potential wells created by the dark matter halos. This gas cools down and starts to flow towards the dark matter halos, which allows the first galaxies to form. These halos continue to grow at all times, attracting nearby particles or merging with other (smaller) halos, in a process known as the *hierarchical structure formation model* (Dekel and Silk, 1986; Frenk et al., 1988; White and Frenk, 1991). In the present age of the Universe, dark matter halos can be regarded as the building blocks of the large scale structure of the universe. Since the ratio between dark matter and baryonic matter is 5 to 1 (Planck Collaboration et al., 2018), massive, big halos may host one, or many, or sometimes even hundreds of galaxies.

Given this scenario, it is clear that by studying the statistics of the distribution of dark matter halos throughout the universe, one is also studying the statistics of the distribution of galaxies. The *galaxy-halo connection* refers, for the most part, to the full multivariate distribution of properties of halos and the galaxies that form within them (Wechsler and Tinker, 2018; Montero-Dorta et al., 2020). Understanding the connection between halos and galaxies could help us answer open problems in cosmology and astrophysics, such as the determination of cosmological parameters, how galaxies form and evolve, and how galaxies probe the nature and distribution of dark matter.

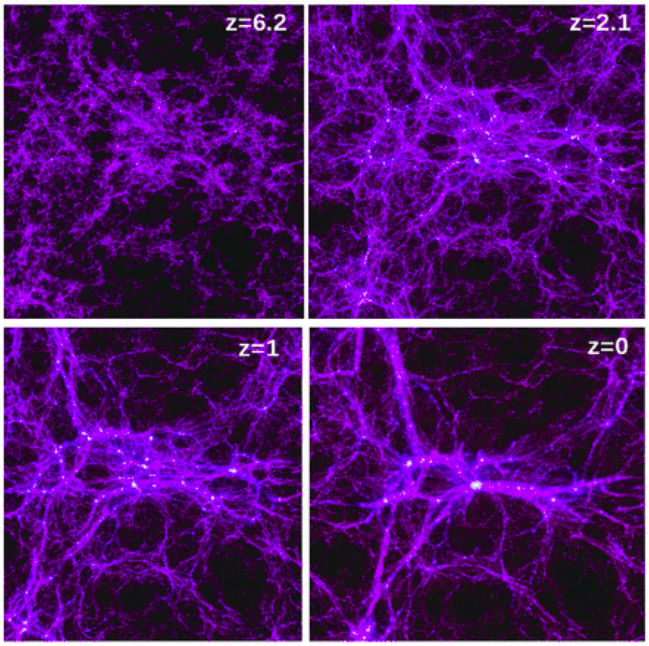


Figure 1.3: Evolution of the (projected) dark matter density field in a slice of 100 $h^{-1}$Mpc and thickness 15 $h^{-1}$Mpc from the Millennium-II simulation (Boylan-Kolchin et al., 2009). The redshift corresponding to each snapshot is shown on the top right. (Zavala and Frenk, 2019)

In the last decade, the rapid development of high-resolution, large-volume numerical simulations allowed us to start studying in detail the distribution of dark matter halos and galaxies with progressively more accurate precision (Efstathiou et al., 1985; Stadel, 2001). These cosmological N-body simulations track the evolution of billions of dark matter particles in the background of an expanding Universe, as they move around through the effect of their mutual gravitational forces. These numerical simulations, together with the advent of a large number of galaxy surveys that could detect and probe the clustering of galaxies, are powerful tools that can help us in our search for understanding the laws of Nature.

In principle, anybody can run their own dark matter-only N-body numerical simulation by using, e.g., the Gadget-4 (Springel et al., 2021) or Ramses codes (Teyssier, 2002), and then apply halo finder algorithms to detect the halos. The more we refine the methods and increase the spatial resolution of those simulations, the greater becomes their computational costs. Most of these simulations are parallelizable, so it is also possible to run them in big clusters. There are a number state-of-the-art simulations run on supercomputers for a long time, with snapshots (i.e., 3D portraits of the distribution of the particles at some given instants) made available for the community to analyse. These simulations typically have only dark-matter particles, such as in the case of MultiDark (Klypin et al., 2016), Quijote (Villaescusa-Navarro et al., 2020) and Uchuu (Ishiyama et al., 2021), but some of them

can also solve hydrodynamical equations in order to account for baryonic physics, such as the case of Illustris (Springel et al., 2005), Millennium (Boylan-Kolchin et al., 2009), Eagle (Schaye et al., 2015) and BAHAMAS (McCarthy et al., 2017). Uchuu, which is one of the latest, has 2 billion dark-matter particles in a box of 2 $h^{-1}$Gpc and was run in the ATERUI II supercomputer in Japan, producing 4PB of raw data. Finally, we should also mention that it is possible to populate the dark-matter only simulations "by hand" with galaxies, using methods such as *halo occupation distribution* (HOD) models (Cooray and Sheth, 2002; Zheng et al., 2005, 2009; Voivodic and Barreira, 2020).

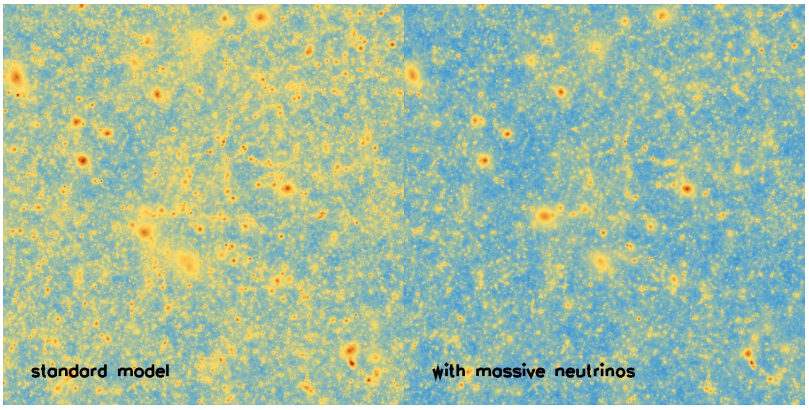


Figure 1.4: Comparision of the BAHAMAS simulations with and without massive neutrinos. http://www.virgo.dur.ac.uk/2016/10/30/BAHAMAS/index.html

## 1.2 Bias

The distribution of matter throughout the Universe holds information about new physics, since it depends sensitively on the precise way in which matter clusters: e.g., the nature of dark matter as well as the theory of gravity both influence how halos form and evolve. However, what we observe in astrophysical surveys are objects such as galaxies, quasars, galaxy groups and clusters, etc., which we call *cosmological tracers* since they trace the distribution of matter particles. Even though we don't observe them directly, dark matter halos can also be thought as tracers of the underlying matter density.

We can define the *bias* as the statistical relation between the distribution of the cosmological tracers to the underlying matter density field (Desjacques et al., 2018). By constructing an expansion over the observables constructed from the density field, where the bias parameters are the coefficients, we are able to infer the matter distribution by looking at the distribution of these cosmological tracers. This connection allows us to extract information about many open questions in physics, such as the nature of dark matter, dark energy, gravity and inflation.

On large scales and at a fixed time, we can say that the overdensity of galaxies (or any cosmological tracer) and of matter is given by the simple relation:

$$\delta_g(\boldsymbol{x}) \equiv \frac{n_g(\boldsymbol{x})}{\overline{n}_g} - 1 = b_1\,\delta(\boldsymbol{x}) = b_1\left(\frac{\rho_m(\boldsymbol{x})}{\overline{\rho}_m} - 1\right), \tag{1.1}$$

where $n_g(\boldsymbol{x})$ is the number of galaxies and $\rho_m(\boldsymbol{x})$ is the matter density at a certain location $\boldsymbol{x}$, while $\overline{n}_g$ and $\overline{\rho}_m$ are their mean over the volume. As we are going to see in more detail later on, we should write a bias expansion defined at each order in perturbation theory, and as we explore progressively smaller scales, we also need to add more terms to the right-hand side of Equation (1.1) in order to take into account those non-linear interactions. We would also have to add a noise term that reflects the stochastic influence of small-scale perturbations on the formation of galaxies, but we will neglect this term here for simplicity.

Cosmological surveys allow us to extract some summary statistics of the galaxy overdensity, such as the *correlation function* (its two-point function in real space) $\xi_{gg}(r) = \langle\delta_g(\boldsymbol{x})\delta_g(\boldsymbol{x}+\boldsymbol{r})\rangle$ or its Fourier transform, the *power spectrum* $P_{gg}(k) = \langle\delta_g(k)\delta_g^*(k)\rangle$. Given the observed statistics of galaxies, we can infer those of the matter density field by using the appropriate scaled bias parameters, e.g., $P_{gg}(k) = b_1^2\langle\delta(k)\delta^*(k)\rangle$. Since the distribution of matter potentially contains information on new physics, the importance of a good understanding of galaxy bias becomes quite evident, since it allow us to make the connection between observations and theory.

As we already saw in the context of halo-galaxy connection, studying the bias of dark-matter halos, which are the building blocks of the large-scale structure, is an essential step to understand the bias of galaxies. However, to this day we do not completely understand its biasing. It is well known that rare, massive halos tend to be found in denser regions in the universe (i.e., they have a larger bias), while the less massive ones are more usually found in regions with lower density (i.e., tey have a lower bias) (Kaiser, 1984; Sheth and Tormen, 1999). In other words, massive halos are more tightly clustered than less massive ones.

Let us see an example of this fact. If we calculate the matter power spectrum $P_{mm}(k)$ and the halo-matter cross spectrum $P_{hm}(k) = \langle\delta_h(k)\delta^*(k)\rangle$ for halos of different masses in MDR1 (one of the cosmological simulations of MultiDark), the bias of such halos on large scales would then be given by $b_1 = \lim_{k\to 0}\frac{P_{hm}(k)}{P_{mm}(k)}$. As Figure 1.5 shows, the more massive the halo, the larger is its bias. There is a fitting function provided by Tinker et al. (2010), through which we can obtain the bias of halos as a function of their mass.

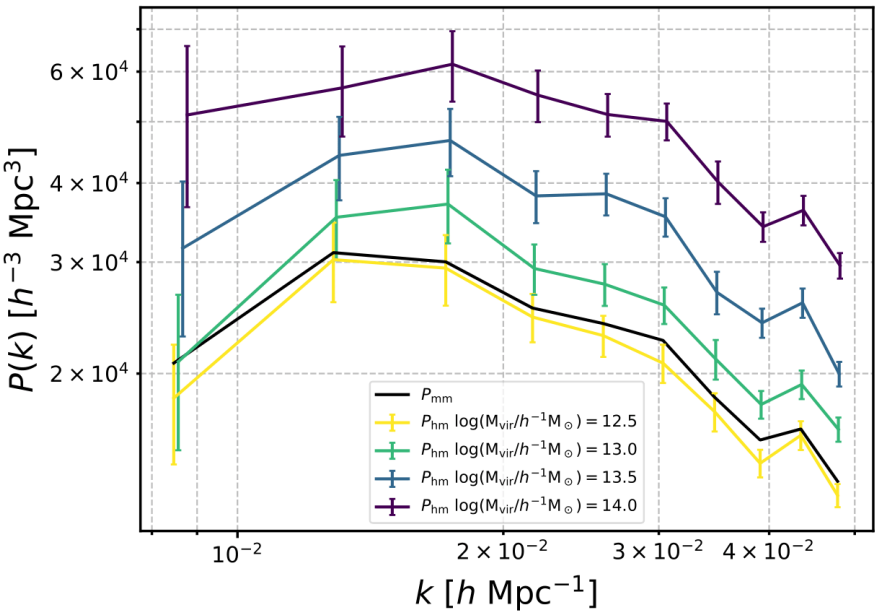


Figure 1.5: Matter power spectrum and halo-matter cross spectrum for halos of different bias calculated in MDR1.

The observation that halo bias depends mostly on halo mass is key to several (semi-)analytical models for halo and galaxy statistics, such as the HOD model. Indeed, we can say that halo mass is the primary bias parameter – but hardly the only parameter. Halos of a given mass can differ in their formation history, as well in their properties at any given time. Moreover, they can live in different large-scale environments, which also affects their bias. Therefore, a natural question arises: how do these different circumstances are connected to halo clustering? Recent numerical simulations revealed that, for a fixed mass, halo clustering depends sensitively on several secondary halo properties (i.e., factors besides mass, which is the primary bias parameter) such as age, spin and concentration (Gao et al., 2005; Wechsler et al., 2006; Gao and White, 2007; Sato-Polito et al., 2018).

## 1.3 Secondary Bias

The first evidence for other dependencies of halo bias besides mass came from numerical simulations, which found that, *at fixed halo mass*, lower-mass halos that assemble a significant portion of their mass early on (i.e., older halos) are more tightly clustered that those that assemble at later times (younger halos) (Sheth and Tormen, 1999; Gao et al., 2005). This effect dictated that halo clustering (bias) is not exclusively determined by halo mass, but also depends on the assembly history of halos. The dependence of halo bias on other halo properties beside mass has collectively been termed *assembly bias* because of this first result. However, since then it has been found that halo bias depends on several

*secondary* halo properties, at levels comparable to halo assembly bias, so that a more appropriate name to this effect is *secondary bias*.

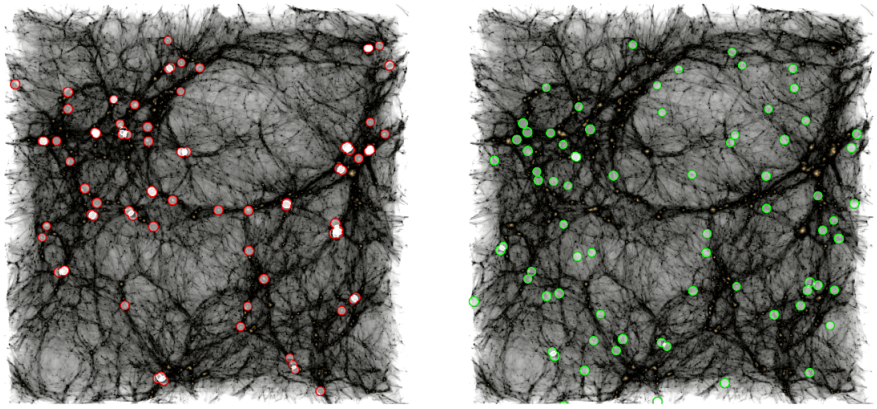

Figure 1.6: 90x90x30 Mpc slice of a cosmological simulation at $z = 0$ for halos with mass $\log M = 10.8\, h^{-1} M_{\odot}$. Red halos are the 5% more concentrated halos and green halos are the 5% less concentrated ones (Li et al., 2008). We see that these two population of halos of same mass but distinct concentrations display different clustering features and thus have different bias.

Since galaxies also differ from one another in term of properties such as color, star formation rate, metallicity, morphology, etc., and they reside in halos of different masses, it is natural to expect that secondary bias also extends to galaxies. *Secondary galaxy bias* concerns the effect of how galaxy clustering depends on secondary galaxy properties, at *fixed host halo mass*.

To this day we do not have neither conclusive observational evidence nor a complete analytical framework to explain all the secondary bias trends. Our work focuses on measuring these effects with high precision in cosmological N-body simulations, and attempting some theoretical explanations for the results. We are also interested in developing and implementing new techniques to find observational signatures of secondary bias.

Amongst our projects, Montero-Dorta et al. (2020, 2021) are two first-author publications by Prof. Dr. Antonio Montero-Dorta and co-authored by Prof. Dr. Raul Abramo and Beatriz Tucci. In Montero-Dorta et al. (2020) we study the galaxy-halo connection and secondary galaxy bias (see also Section 4.3), while in Montero-Dorta et al. (2021) we have proposed a new route to probe spin bias in observations, using the kinetic Sunyaev-Zel'Dovich effect (see Section 7.3).

Amongst those secondary bias effects, *halo spin bias*, i.e., the dependence of halo clustering on spin (intrinsic angular momentum) at fixed halo mass, was widely thought to result in a higher bias for higher-spin halos, over the entire halo mass range (Gao and

White, 2007; Villarreal et al., 2017; Salcedo et al., 2018). However, in Sato-Polito et al. (2018) (the result of a project also developed along with the same supervisors at the Physics Institute of the University of São Paulo), our group used the high-resolution MultiDark numerical simulations to show that this trend actually inverts at the low-mass end, with lower-spin halos being more tightly clustered than their higher-spin counterparts. This surprising result was later confirmed by Johnson et al. (2019) using two other simulations, Vishnu and Consuelo.

We have recently explained this effect in a first-author publication by Beatriz Tucci and co-authored by her advisors (Tucci et al., 2021), where we demonstrated how splashback halos can fully explain the low-mass spin bias inversion (see Section 7.1). Splashback halos are distinct halos which at some previous time were classified as subhalos. This means that a splashback halo at a given redshift passed through the virial radius of an eventual host halo at an earlier time, but then exited that halo. It is well known that splashback halos are responsible for part of the low-mass secondary bias with respect to halo assembly history (Dalal et al., 2008; Sunayama et al., 2016; Mansfield and Kravtsov, 2020), therefore we unveiled a link between two different secondary bias trends: spin bias and assembly bias.

In Tucci et al. (2021) we found that when the population of splashback halos is removed, the spin bias inversion at low-masses vanishes, and only the *intrinsic* spin bias signal remains, i.e., the dependence of halo clustering at fixed mass on spin for high-mass halos. Since we were able to successfully clean the statistical signal of spin bias at low-masses, isolating what seemed to be a complex effect, we now expect to shed light into the physical mechanisms behind the high-mass trend, with the ultimate goal of providing a complete picture for the origin of halo spin bias.

Developing an analytical framework for high-mass spin bias analytically is a step further in complexity, compared with what we have done for low-mass spin bias. We have gone through the main theory for the origin of halo spin, Tidal Torque Theory (TTT) (White, 1984; Doroshkevich, 1970; Catelan and Theuns, 1996), which states that halo spin arises due to misalignment between the initial shear and inertia tensors (see Section 5.1). We also studied the two main analytical models of halo formation which are capable of making predictions for halo bias, the excursion set approach and the peak formalism (see Chapter 3).

During this process we came up with two hypothesis. First, TTT in the peak formalism (Catelan and Theuns, 1996) can give us the relation between halo angular momentum and peak curvature, which is related to large-scale halo bias. Second, a model for the barrier of halos that depends on the initial shear (Castorina et al., 2016) could be implemented in principle, if we could use TTT to relate halo spin to the shear-dependent quantity appearing in the barrier. In order to understand these ideas, we will develop the necessary

background of the aforementioned theories in the first part of this thesis, while the second part will introduce secondary halo bias in more detail and what we have investigated so far regarding halo spin bias.

## 1.4 Outline

The rest of the text is organised as follows:

**Chapter 2** introduces some important quantities in the study of structure formation and provides the necessary theoretical background.

**Chapter 3** discusses the main analytical models of halo formation, such as the peak formalism, excursion sets and ellipsoidal collapse model.

**Chapter 4** demonstrates how we can study dark matter halos though cosmological N-body numerical simulations.

**Chapter 5** features the theories we have so far for the origin of the angular momentum of dark matter halos.

**Chapter 6** consists of a brief review on secondary halo bias, showing how we can measure it in simulations and focusing on what we know so far about assembly bias.

**Chapter 7** explains the physical origins of low-mass spin bias and displays our ideas for high-mass spin bias.

# PART I

# Dark Matter Halos

# CHAPTER 2

# How do halos cluster?

This chapter and the next one will be quite technical, since we will learn all the necessary tools to tackle the spin bias puzzle. For that, we will explore many of the important theoretical tools developed in the last years to explain structure formation in the Universe. Equipped with them, we are able to provide a consistent theoretical framework for LSS and to efficiently extract cosmological information from the upcoming surveys.

## 2.1 Setting the stage

The first step is to develop a model for the clustering of matter, which takes place under gravitational collapse given the initial density perturbations (see Coles 2001 for a review). Unlike the case of the CMB, for which a linear perturbation theory description is sufficient, the matter perturbations become nonlinear at low/intermediate redshifts on scales of relevance for cosmology. Strictly speaking, the nonlinearity depends on the scale we are looking at, and processes can only be studied with numerical N-body simulations or heuristic arguments. At large scales and initial times we can make use of perturbation theory to accurately predict matter clustering, and this can be extended to quasi-linear scales as long as the results converge at sufficient high order up a maximum scale of validity (see Bernardeau et al. 2002 for a review).

Notably, as our observations give us the statistics of the distribution of galaxies and not of matter, the second step is to connect our theory of matter clustering to that of galaxy[1] clustering. The *bias parameters* will provide us the statistical relation between the two, as we are going to discuss in Section 2.5. As stated in Desjacques et al. (2018), "the perturbative theory of galaxy clustering, valid on quasi-linear scales, is based on two key ingredients: (i) a perturbation theory prediction for the matter density and tidal field; (ii) a complete parametrization of galaxy bias at each order in perturbation theory".

[1] In the context of the bias expansion, "galaxy" can also refer to any cosmological tracer, such as dark-matter halos, quasars, etc.

### Notation and assumptions

In this thesis we will assume a spatially flat background described by the Friedmann-Robertson-Walker (FRW) metric, as supported by cosmological observations (Dodelson and Schmidt, 2020). We will use *comoving coordinates*[2], such that in this background the matter particles will remain at a fixed position while the Universe expands. Further, as we add perturbations to the FRW metric, we also need to specify the gauge, since the General Relativity field equations are invariant under a change of coordinates, i.e., diffeomorfisms. We will choose to attach the observers to the points in the unperturbed frame in comoving coordinates, and work in the so-called *conformal-Newtonian gauge*, in which the equations for dark matter reduce to the Newtonian case (Luca Amendola, 2010). We will often use as our time variable the conformal time[3]. Finally, we will be dealing with scalar perturbations only, e.g., matter density scalar perturbations, and not vector or tensor (gravitational waves) perturbations. We will also assume that the initial fluctuations are purely adiabatic (density) perturbations, and therefore we will neglect entropy (isocurvature) perturbations.

## 2.2 Cosmological perturbations

Since the initial conditions of the LSS are given by inflation, which is originated by quantum mechanical vacuum fluctuations, we ought to work with Gaussian random fields in order to describe the precise initial conditions of our Universe, as they are not deterministic (for an overview on inflation, we refer the reader to Mukhanov 2005; Baumann 2009).

In agreement with the *cosmological principle*, we will assume that all the cosmological random fields are statistically homogeneous and isotropic. We will use the fair sample hypothesis, which states that samples extracted from regions of the Universe that are sufficiently distant from each other are independent realizations of the same physical process. Thus, we have the ergodic hypothesis, which allows us to trade ensemble averages with spatial averages if the volumes are large enough.

### Gaussian random fields and their moments

As we already saw in Chapter 1, the cosmological matter perturbations:

$$\delta(\boldsymbol{x}) \equiv \frac{\rho(\boldsymbol{x}) - \overline{\rho}}{\overline{\rho}} \,, \tag{2.1}$$

[2]The physical coordinate $\boldsymbol{x}_{\rm ph}(t)$ and comoving one $\boldsymbol{x}$ are related by the scale factor $a(t)$, such that $\boldsymbol{x}_{\rm ph}(t) = a(t)\boldsymbol{x}$.

[3]The conformal time $\tau$ and physical time $t$ are related by the scale factor $a$ such that $d\tau \equiv a^{-1}t$.

or its Fourier transform:

$$\delta(\boldsymbol{k}) = \int d^3\boldsymbol{x}\, \delta(\boldsymbol{x})\, e^{-i\boldsymbol{k}\cdot\boldsymbol{x}} \tag{2.2}$$

can be written in terms of a Gaussian random field with zero mean, i.e., $\langle\delta(\boldsymbol{x})\rangle = \langle\delta(\boldsymbol{k})\rangle = 0$. If we are able to measure the density fluctuations at their actual positions (in *real space*), then we can compute their two-point *correlation function* $\xi(r)$ as:

$$\xi(r) \equiv \xi(|\boldsymbol{x}_1 - \boldsymbol{x}_2|) \equiv \langle\delta(\boldsymbol{x}_1)\delta(\boldsymbol{x}_2)\rangle\,, \tag{2.3}$$

since *statistical homogeneity* imposes that the correlation functions $\xi(\boldsymbol{x}_1, \boldsymbol{x}_2)$ are invariant under global translations on a fixed time slice, such that it can be written as $\xi(\boldsymbol{x}_1 - \boldsymbol{x}_2)$. Moreover, *statistical isotropy* requires this quantity to be invariant under global rotations, allowing us to write the correlation function as $\xi(|\boldsymbol{x}_1 - \boldsymbol{x}_2|)$. In Fourier space, we obtain the power spectrum $P(k)$:

$$\langle\delta(\boldsymbol{k})\delta^*(\boldsymbol{k'})\rangle = P(k)\,(2\pi)^3\,\delta_D(\boldsymbol{k} - \boldsymbol{k'})\,, \tag{2.4}$$

where statistical homogeneity requires the Dirac delta and statistical isotropy requires $P(\boldsymbol{k})$ to depend only on $k = |\boldsymbol{k}|$. Its amplitude given by:

$$\mathrm{PDF}(|\delta(\boldsymbol{k})|) = \frac{1}{\sqrt{2\pi\sigma_P^2(\boldsymbol{k})}}\,\exp\left\{-\frac{1}{2}\frac{|\delta(\boldsymbol{k})|^2}{\sigma_P^2(\boldsymbol{k})}\right\}\,, \tag{2.5}$$

where $\sigma_P^2(\boldsymbol{k}) = (2\pi)^3 P(k)/\tilde{V}(\boldsymbol{k})$, with $\tilde{V}(\boldsymbol{k})$ expressing the phase space volume over which we define the mode $\boldsymbol{k}$. Another related quantity is the variance of an isotropic random field, which is also the value of the correlation function in the limit of zero separation:

$$\xi(0) = \frac{1}{2\pi^2}\int_0^\infty dk\, k^2\, P(k) = \int_0^\infty \frac{dk}{k}\Delta^2(k)\,, \tag{2.6}$$

where we have defined the dimensionless power spectrum $\Delta^2(k) \equiv k^3 P(k)/(2\pi^2)$.

## Smoothed Fields

We often deal with the smoothed version of the matter density field, $\delta_R$, which is obtained by convolving the density field with a spherically symmetric kernel $W_R$ at a given smoothing filter scale $R$. The kernels are normalized as $\int d^3\boldsymbol{x}\, W_R(|\boldsymbol{x}|) = 1$, such that $\lim_{k\to 0} W_R(k) = 1$. We list below the most used filtering kernels and their Fourier transform.

- Sharp-$k$ filter

$$W_R(x) = \frac{3}{4\pi R^3}\left[3\,\frac{j_1(x/R)}{x/R}\right], \quad W_R(k) = \Theta_H(1-kR) \tag{2.7}$$

- Gaussian filter

$$W_R(x) = \frac{1}{[2\pi R^2]^{3/2}}\, e^{-\frac{1}{2}x^2/R^2} \quad W_R(k) = e^{-\frac{1}{2}R^2k^2} \tag{2.8}$$

- Tophat filter

$$W_R(x) = \frac{3}{4\pi R^3}\Theta_H\left(1-\frac{x}{R}\right), \quad W_R(k) = 3\,\frac{j_1(kR)}{kR} \tag{2.9}$$

Here, $j_1(x)$ is the spherical Bessel function:

$$j_1(x) = \frac{\sin x - x\cos x}{x^2}\,. \tag{2.10}$$

The shape of the filter in real space should reflect in some sense the Lagrangian density profile of halos. Therefore, the sharp-$k$ filter is not physical, as the corresponding real-space filtering kernel $W_R(x)$ is not always positive. Since convolution in real space is a multiplication in Fourier space, we have that the variance of the matter field on a scale $R$ is

$$S(R) \equiv \sigma^2(R) = \langle\delta_R^2\rangle = \int_0^\infty d\,\ln k\;\Delta^2(k)\,|W_R(k)|^2\,. \tag{2.11}$$

We can generalize this definition to obtain the *spectral moments* of the matter power spectrum smoothed on a scale $R$,

$$\sigma_n^2(R) \equiv \int_0^\infty d\,\ln k\;\Delta^2(k)\,k^{2n}\,|W_R(k)|^2\,, \tag{2.12}$$

from which we recover $S(R) = \sigma_0^2(R)$.

## The two-point correlation function

Since $\xi(r)$ will represent a fundamental tool for the measurement of halo clustering in this project, we will devote some time to better understand its interpretation, following the pedagogical arguments presented in Luca Amendola (2010).

If $\rho_0 dV$ is the average number of particles in the infinitesimal volume and $dN_{ab} = \langle n_a n_b\rangle$ is the average number of pairs in the volumes $dV_a$ and $dV_b$ (i.e. the product of the number of particles in one volume times the number in another volume), for volumes separated by $r_{ab}$, then the 2-point correlation function $\xi(r_{ab})$ is defined for $r_{ab} > 0$ as

$$dN_{ab} = \langle n_a n_b \rangle = \rho_0^2 dV_a dV_b [1 + \xi(r_{ab})] \ . \tag{2.13}$$

If the distribution has been obtained by throwing the N particles at random (i.e. without any preference with respect to the place), then there is no reason for $dN_{ab}$ to depend on the location, then

$$\langle n_a n_b \rangle = \langle n_a \rangle \langle n_b \rangle = \rho_0^2 dV_a dV_b \ . \tag{2.14}$$

If $\xi \neq 0$, the particles are correlated. Then the correlation function can be written as a spatial average of the product of the density contrast $\delta(r_a) = n_a/(\rho_0 dV_a) - 1$ at two different points:

$$\xi(r_{ab}) = \frac{dN_{ab}}{\rho^2 dV_a dV_b} - 1 = \langle \delta(r_a)\delta(r_b) \rangle \ . \tag{2.15}$$

If this average is taken to be the sample average, then it means we have to average over all possible positions:

$$\xi(\mathbf{r}) = \frac{1}{V} \int \delta(\mathbf{y})\delta(\mathbf{y} + \mathbf{r}) dV_y \ . \tag{2.16}$$

In practice it is easier to derive the correlation function as the average density of particles at a distance r from another particle, i.e. by choosing the volume $dV_a$ so that $\rho_0 dV_a = 1$. Then the number of pairs is given by the number of particles in the volume $dV_b$

$$dN_{ab} = \rho_0 dV_b [1 + \xi(r_b)] \ , \tag{2.17}$$

$$\xi(r) = \frac{dN(r)}{\rho_0 dV} - 1 = \frac{\langle \rho_c \rangle}{\rho_0} - 1 \ , \tag{2.18}$$

where $\langle \rho_c \rangle$ is the average number of particles at distance $r$ from any given particle and $\rho_0$ is the expected number of particles at the same distance in a uniform distribution. If $\xi(r) > 0$, there are more particles than in a uniform distribution.

In practice, a direct estimation of the shell density (Eq. 2.18) is difficult because of the complicated boundary and selection procedure that a real survey often has. The simplest way to measure $\xi$ is to compare the real catalog to an artificial random catalog with exactly the same boundaries and the same selection function (the selection function here is the mean number of objects). Then the estimator can be written as (Davis and Peebles, 1983)

$$\xi = \frac{DD}{DR} - 1, \tag{2.19}$$

where DD means the number of galaxies at distance $r$ counted by an observer centered on a real galaxy (data D). This is divided by the number of galaxies DR at the same distance but in the random catalog.

In other words, instead of calculating the volume of the shell (which is a difficult task in realistic cases), we estimate $\xi$ by counting the galaxies in the Monte Carlo realization. In this way all possible boundaries and selection function effects can be taken into account through the random catalog, since they will affect DD and DR in the same way.

The most commonly-used estimator (also used in this work) is from Landy and Szalay (1993),

$$\xi = \frac{1}{RR}\left[DD\left(\frac{n_R}{n_D}\right)^2 - 2\,DR\left(\frac{n_R}{n_D}\right) + RR\right] , \tag{2.20}$$

where $n_D$ and $n_R$ are the mean number densities of galaxies in the data and random catalogs. While this estimator requires more computational time it is less sensitive to the size of the random catalog and handles edge corrections well, which can affect clustering measurements on large scales (Coil, 2013).

## 2.3 Large Scale Structure fields

This section will give us useful definitions of the fields defined on the study of structure formation in the Universe and with which we will be dealing throughout this thesis. The notation presented here is a mixture of Hahn and Paranjape (2014); Castorina et al. (2016); Desjacques et al. (2018).

### Tidal field

The first field (or tidal tensor, as it is a second-rank tensor) we can define is the local gravitational observable of General Relativity, the second derivative of the gravitational potential $\Phi$, which is related to $\delta$ via the Poisson equation,

$$\Phi = \mathcal{F}^{-1}\{-k^{-2}\delta(\boldsymbol{k})\} , \tag{2.21}$$

such that the tidal field reads

$$\psi_{ij} = \partial_{ij}\Phi = \mathcal{F}^{-1}\left\{\frac{k_i k_j}{k^2}\,\delta(\boldsymbol{k})\right\} , \tag{2.22}$$

where $\mathcal{F}^{-1}$ denotes the inverse Fourier transform. Its eigenvalues are $\lambda_1 \leq \lambda_2 \leq \lambda_3$. and we define $(\psi_1, \psi_2, \psi_3, \psi_4, \psi_5, \psi_6) = (\psi_{11}, \psi_{22}, \psi_{33}, \psi_{12}, \psi_{13}, \psi_{23})$. In an arbitrary basis, we can define the following useful combination of matrix elements

$$\sigma\,\nu \equiv \psi_1 + \psi_2 + \psi_3 \tag{2.23}$$
$$\sigma\,l_2 \equiv \sqrt{15}\,(\psi_1 - \psi_3)/2 \tag{2.24}$$
$$\sigma\,l_3 \equiv \sqrt{5}\,(\psi_1 - 2\psi_2 + \psi_3)/2 \tag{2.25}$$
$$\sigma\,l_A \equiv \sqrt{15}\,\psi_A\,,\ A = 4, 5, 6, \tag{2.26}$$

where $\sigma$ is the smoothed variance filtered with a tophat (see Section 2.2).

## Shear field

The shear is the traceless part of the tidal field

$$K_{ij} = \mathcal{F}^{-1}\left\{\left[\frac{k_i k_j}{k^2} - \frac{1}{3}\delta_{ij}\right]\delta(\boldsymbol{k})\right\} \tag{2.27}$$

with eigenvalues $\kappa_1 \leq \kappa_2 \leq \kappa_3$.

## Delta

The matter fluctuation $\delta$ is the trace of the tidal field, i.e.,

$$\delta = \mathrm{Tr}\,\partial_{ij}\psi = \lambda_1 + \lambda_2 + \lambda_3\,. \tag{2.28}$$

## Nabla of Delta

We can define the first derivative of $\delta$ as

$$\boldsymbol{\eta} = \boldsymbol{\nabla}\delta = \mathcal{F}^{-1}\{i\boldsymbol{k}\delta(\boldsymbol{k})\}\,. \tag{2.29}$$

## Hessian of Delta

The second derivative of $\delta$ in turn is

$$H_{ij} = \zeta_{ij} = \partial_{ij}\delta = \mathcal{F}^{-1}\{-k_i k_j \delta(\boldsymbol{k})\}\,, \tag{2.30}$$

with eigenvalues $\zeta_1 \leq \zeta_2 \leq \zeta_3$.

## Curvature

The curvature

$$x = -(\zeta_1 + \zeta_2 + \zeta_3)/\sigma_2 \tag{2.31}$$

will be an important quantity in the peak formalism (Section 3.5), as it reflects the "sharpness" or "shalloweness" of the peak. The highest the curvature $x$, the sharper the peak. The spectral moment $\sigma_2$ is defined in Equation (2.12) and smoothed with a Gaussian filter.

## Ellipicity and Prolateness

The ellipicity $e$ and prolateness $p$ associated to the tidal tensor are

$$Y = e\delta = (\lambda_3 - \lambda_1)/2 \tag{2.32}$$
$$Z = p\delta = (\lambda_3 - 2\lambda_2 + \lambda_1)/2\,, \tag{2.33}$$

characterize the asymmetry of the isodensity profile (Catelan and Theuns, 1996), while

$$y = -(\zeta_3 - \zeta_1)/(2\sigma_2) \tag{2.34}$$
$$z = -(\zeta_3 - 2\zeta_2 + \zeta_1)/(2\sigma_2)\,, \tag{2.35}$$

describe the shape of the peak (Bardeen et al., 1986).

## Tidal Anisotropy

We can note that, since

$$K_{ij} = \left(\frac{\partial_{ij}}{\nabla^2} - \delta_{ij}\right)\delta \tag{2.36}$$

and thus $K_{ij} = \mathrm{diag}(\lambda_i - \frac{1}{3}(\lambda_1 + \lambda_2 + \lambda_3))$ in the eigenbasis of the tidal field, we get that

$$\mathrm{tr}(K_{ij})^2 = (\lambda_1^2 + \lambda_2^2 + \lambda_3^2) - \frac{2}{3}(\lambda_1 + \lambda_2 + \lambda_3)^2 + \frac{3}{9}(\lambda_1 + \lambda_2 + \lambda_3)^2 \tag{2.37}$$
$$= \left(1 - \frac{2}{3} + \frac{1}{3}\right)(\lambda_1 + \lambda_2 + \lambda_3)^2 - 2(\lambda_1\lambda_2 + \lambda_1\lambda_3 + \lambda_2\lambda_3) \tag{2.38}$$
$$= \frac{2}{3}(\lambda_1 + \lambda_2 + \lambda_3)^2 - 2(\lambda_1\lambda_2 + \lambda_1\lambda_3 + \lambda_2\lambda_3)\,. \tag{2.39}$$

We can define the tidal anisotropy as $\alpha$ (Paranjape et al., 2018)

$$\alpha = \frac{1}{2}[(\lambda_1 - \lambda_2)^2 + (\lambda_1 - \lambda_3)^2 + (\lambda_2 - \lambda_3)^2] \tag{2.40}$$

$$= (\lambda_1^2 + \lambda_2^2 + \lambda_3^2) - (\lambda_1\lambda_2 + \lambda_1\lambda_3 + \lambda_2\lambda_3) \tag{2.41}$$

$$= (\lambda_1 + \lambda_2 + \lambda_3)^2 - 3(\lambda_1\lambda_2 + \lambda_1\lambda_3 + \lambda_2\lambda_3) \tag{2.42}$$

such that

$$\alpha = \frac{3}{2}\,\mathrm{tr}(K_{ij})^2\,. \tag{2.43}$$

## 2.4 The Cosmic Web

We can use the value of the eigenvalues of the tidal field at each region of the LSS to define its different components. These structures are composed by *nodes* ($\lambda_1 < 0$), a region to which matter is infalling from all directions, *filaments* ($\lambda_1 < 0$ & $\lambda_2 > 0$), where matter is only receding along its axis, *sheets* ($\lambda_2 < 0$ & $\lambda_3 > 0$), where the infall happens only in one direction and finally *voids* ($\lambda_3 < 0$), a structure from which matter is receding in all directions. It is claimed that the position of halos in the cosmic web highly affects its internal and clustering properties (Musso et al., 2018; Paranjape et al., 2018; Ramakrishnan et al., 2019).

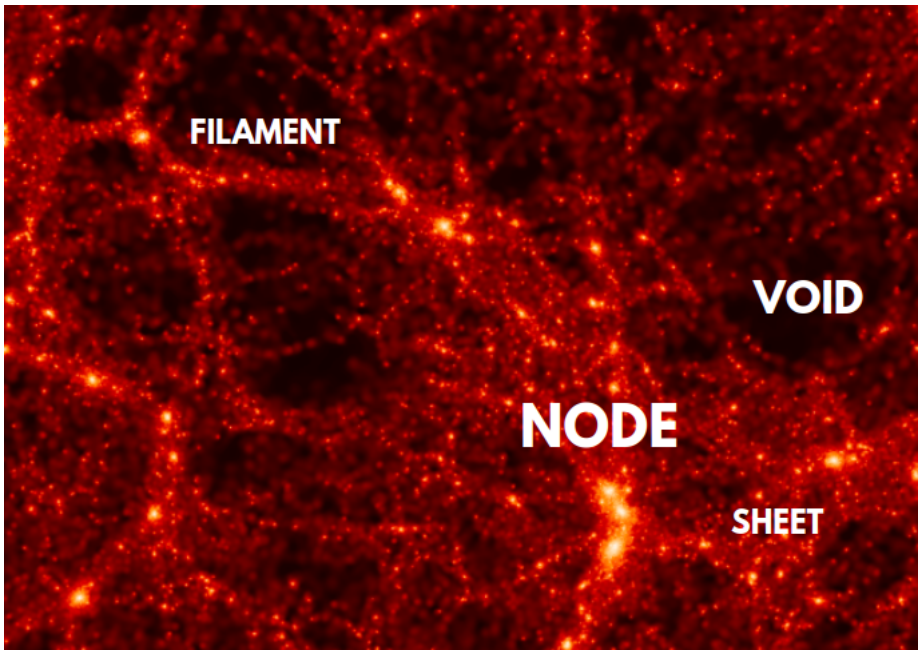


Figure 2.1: Slice of a MultiDark simulation indicating the different components of the cosmic web.

## 2.5 Bias

The bias can be broadly defined as the statistical relation between the distribution of the cosmological tracers (galaxies, dark matter halos, quasars, etc.) to their underlying matter density field. This connection allows us to extract information about many open questions in physics, such as the nature of dark matter, dark energy, gravity and inflation.

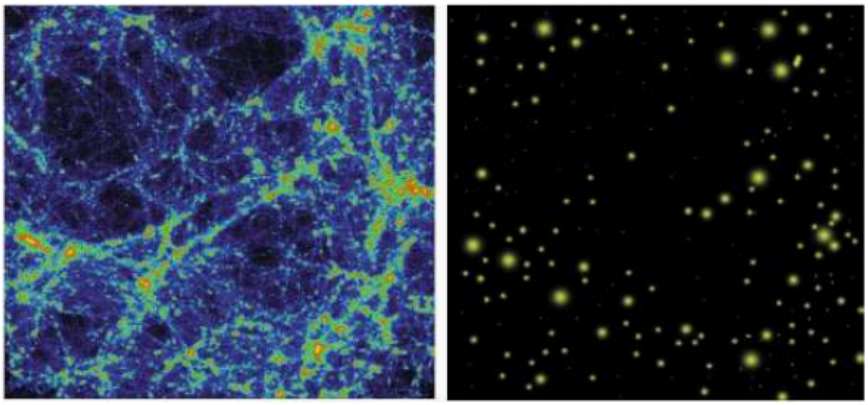

Figure 2.2: Slice of a Numerical Simulation. Left: dark matter density field. Right: distribution of dark matter halos. Comparing two images, we can visually see how halos are biased tracers of the underlying density field. Cooray and Sheth (2002)

### The Bias Expansion

The bias expansion is built from the bias parameters, which are defined at each order in perturbation theory and at fixed time. Assuming that on large scales structure formation is completely driven by gravity, these constitute a finite set of parameters on quasi-linear scales. We restrict ourselves to these scales where highly non-linear effects do not play an important role, such as complicated and yet not understood processes of galaxy formation and evolution (see Desjacques et al. 2018 for a review).

In summary, we can write the overdensity of the cosmological tracers, such as galaxies, as a function of certain operators $O(\mathbf{x}, \tau)$, where the bias parameters are defined as their coefficients, i.e.,

$$\delta_g(\mathbf{x}, \tau) = \sum_O b_O(\tau) O(\mathbf{x}, \tau) \,. \tag{2.44}$$

In order to characterize galaxy clustering on the scales where the bias expansion is valid, we have then to write down all operators[4] at each order in perturbation theory allowed by general covariance, as for instance powers of the tidal field operator.

[4]"Operators" here are not in the same sense of usual Quantum Field Theory (QFT), where the fields act on a Hilbert space. Presumably use such word because these will be renormalized as we do for fields in QFT.

Note that this perturbative bias expansion is ill-defined, considering that on large scales contributions of higher-order bias terms are not necessarily suppressed and some physical quantities might depend on an arbitrary smoothing scale which should be irrelevant. We can then redefine these operators by introducing local counterterms that absorb the large contributions. In this way, we naturally obtain an effective theory, where the physical bias parameters are the coefficients of renormalized operators and do not depend on any smoothing scale. For more information, see e.g. Assassi et al. (2014).

## The Basis of Operators

The allowed operators and corresponding appearing in the expansion can be broken in three categories. First, the local observables of long-wavelength spacetime perturbations. As the equivalence principle states, these are the operators related to second derivatives of the gravitational potential $\psi$, such as $\delta$ which is related to $\nabla^2\psi$ through the Poisson equation (see Section 2.3) or the shear field[5] (see Section 2.3). The second category is the one of higher derivative ones, which are also gravitational observables but involve more than two derivatives of $\psi$, as for instance $\nabla^2\delta$. Lastly we have the *stochastic contributions*, which besides the leading order one which takes into account the small-scale effects of galaxy formation processes, also introduces a stochastic contribution in the expansion for each operator. These bias terms can be seen as a "scatter" in the deterministic bias parameters $b_O$.

Here, we will briefly discuss which are the operators appearing in the bias expansion at first and second orders, i.e.,

$$\delta_g = b_1\delta + b_{\nabla^2\delta}\nabla^2\delta + \varepsilon \qquad \text{1st order} \tag{2.45}$$

$$+ b_2\delta^2 + b_{K_2}K_2 + \varepsilon_\delta\delta\,. \qquad \text{2nd order} \tag{2.46}$$

Since the tidal field $\partial_{ij}\psi$ is our gravitational observable, naturally the first operator appearing in the expansion is its trace, $\delta$, and its square at second order, $\delta^2$. From the remaining traceless part of the tidal field, the shear field $K_{ij}$, we have at second order the invariant $K_2 \equiv (K_{ij})^2$ (see Section 2.3) which arises to account for non-spherically symmetric processes of galaxy formation.

The *non-local*, higher-derivative bias $b_{\nabla^2\delta}$ appears when we take into consideration that the cosmological tracers are not point-like, but they actually have an extension which is characterized by $R_*$. In the case of galaxies, for example, this is usually the size of its host halo, but it can be considerably larger. If e.g. flux of ionizing UV radiation affects galaxy

[5]Note that $\partial_{ij}/\nabla^2$ counts as zero net spatial derivatives.

formation, then $R_*$ is of the order of several hundred Mpc. To account for this non-locality, our simple, local operator $\delta$ has to be re-scaled to a functional with an appropriate kernel, and if we expand the re-scaled $\delta(\boldsymbol{x})$ around $\boldsymbol{x}$ the first term appearing will be the operator $\nabla^2\delta$ (terms involving its first derivatives vanish due to isotropy). The higher-derivative term is suppressed on large scales, since $\nabla^2\delta(\mathbf{x})$ becomes $-k^2\delta(\mathbf{k})$ in Fourier space – see Lazeyras and Schmidt (2019) for its actual measurement.

Our theory relies on the study of how large-scale perturbations affect galaxy clustering, so we do not consider how small-scale perturbations influence galaxy formation. However, we can treat them as stochastic contributions, in a Universe without primordial non-gaussianities (which is the case we usually assume) and introduce stochastic fields in our expansion, such as $\varepsilon$. The higher-order stochastic term $\epsilon_\delta\delta$ arises due to the fact that the presence of stochasticity at a given time couples to gravitational evolution. This contribution do not correlate with the long wavelengths perturbations and its correlation with $\varepsilon$ vanishes on large scales, such that its contribution can be neglected for two-point statistics analysis.

# CHAPTER 3

# How do halos form?

The next sections will be devoted to give us an overview of the developed models to understand structure formation in the Universe, the *peak formalism* and *excursion sets*, which are constructed in *Lagrangian space*, that is, in the initial matter density field extrapolated to a desired reference time using the linear growth. The peak formalism claims that halos are formed by peaks in the initial density field, while excursion sets assumes that a halo of radius $R$ is formed when the smoothed density at $R$ first reaches a density threshold $\delta_{\rm cr}$. Both assumptions can be combined in the so-called *excursion-set peaks* approach.

These analytical models are able to give us a predictions which match reasonably well with numerical N-body simulations (see Chapter 4), despite their simple assumptions. The main observable predicted by these models is the *halo mass function*. This quantity gives us the number of halos of a given mass, so that it can be readily tested against N-body simulations, and from it we can obtain some of the bias parameters.

This chapter is intended to give a brief summary on the contents of Chapters 5 and 6 from Desjacques et al. (2018) which will be important for developing our hypothesis on the high-mass spin bias.

## 3.1 Spherical Collapse

The prediction of the density threshold $\delta_{\rm cr}$ relies on the *spherical collapse* approximation, which treats the collapse of halos as spherically-symmetric homogeneous density perturbations in a FLRW metric (Gunn and Gott, 1972). Such perturbations are know to evolve independently of its surroundings and thus depend only on the matter at its interior (Einstein and Straus, 1945; Carrera and Giulini, 2010), such that the equations of motion can be easily solved for the evolution of its radius $R$,

$$\frac{d^2R}{dt^2} = -\frac{GM(<R)}{R^2} = -\frac{4\pi G}{3}\overline{\rho}_{m,i}(1+\Delta_i)\frac{R_i^3}{R^2}\,, \tag{3.1}$$

where $R_i$, $\Delta_i$ and $\overline{\rho}_{m,i}$ are the initial radius, average density contrast and comoving background matter density, respectively. For an Einstein de-Sitter Universe, the solution is given by linearly growing modes at first, as expected, until a *turn around* (when the radius is maximum). The turn-around radius can be determined by integrating Equation (3.1), setting $dR/dt = 0$ and solving the resulting energy conservation equation for $R$. After turn around, the halo radius decreases until it collapses to a singularity at $R = 0$. This unphysical solution does not happen in Nature, where gradient instabilities due to shell cross break spherical symmetry and lead to a complex bound structure as we see in N-body simulations.

Assuming that virialization is perfect and instantaneous, it is possible to estimate the density of the virialized halo, which in linear theory and at the time of collapse is the critical density $\delta_{\rm cr} \simeq 1.686$ (Desjacques et al., 2018). Of course, this is a rough estimate which also neglects the effects of shear, angular momentum or small-scale perturbations in the gravitational collapse to bounds structures. Spherical symmetry is indeed a good approximation for the formation of massive halos, but the collapse of small halos tend to depend more on the shear field, leading to an *ellipsoidal collapse* where the axes aligned with different eigenvalues of the shear field collapse in different times. This effect can be modeled by extending the constant barrier $\delta_{\rm cr}$ to a stochastic barrier that depends on halo mass (see Section 3.4).

## 3.2 Press-Schechter

The Press-Schechter halo mass function (Press and Schechter, 1974) can be derived based on the assumption that the (linear and smoothed) matter perturbations are Gaussian and that a Lagrangian volume $V = (4\pi/3)R^3$ collapses to form a halo of $M = \bar{\rho}_m V$ when the linearly extrapolated, smoothed density perturbation $\delta_R$ exceeds the critical density threshold $\delta_{\rm cr}$ of spherical collapse.

The last, naïve assumption lead to the halo mass function

$$\bar{n}_{h,\rm PS}(M) = \frac{\overline{\rho}_m}{M}\,\nu_c\, f_{\rm PS}(\nu_c)\left|\frac{d\ln\sigma(R)}{d\ln M}\right|\,, \tag{3.2}$$

where

$$f_{\rm PS}(\nu) = \sqrt{\frac{2}{\pi}}\,\nu_c \exp\left(-\frac{\nu_c^2}{2}\right)\,, \tag{3.3}$$

which is not properly normalized, since when integrated over all halo masses, the resulting matter contained in halos corresponds to only half of the matter density available. At first, Press and Schechter (1974) added a "fudge factor of 2" to fix their prediction, but later this incorrect normalization was realized to be a consequence of the so-called *cloud-in-cloud*

problem (Peacock and Heavens, 1990). The nature of this problem is that the Press-Schechter formalism does not take into account the possibility that an underdense region identified on a scale $R$ may be inside a halo on a bigger scale $R' > R$ (Bond et al., 1991) (see Figure 3.1). This can be solved by the *excursion set* approach with the *first-crossing* constrain.

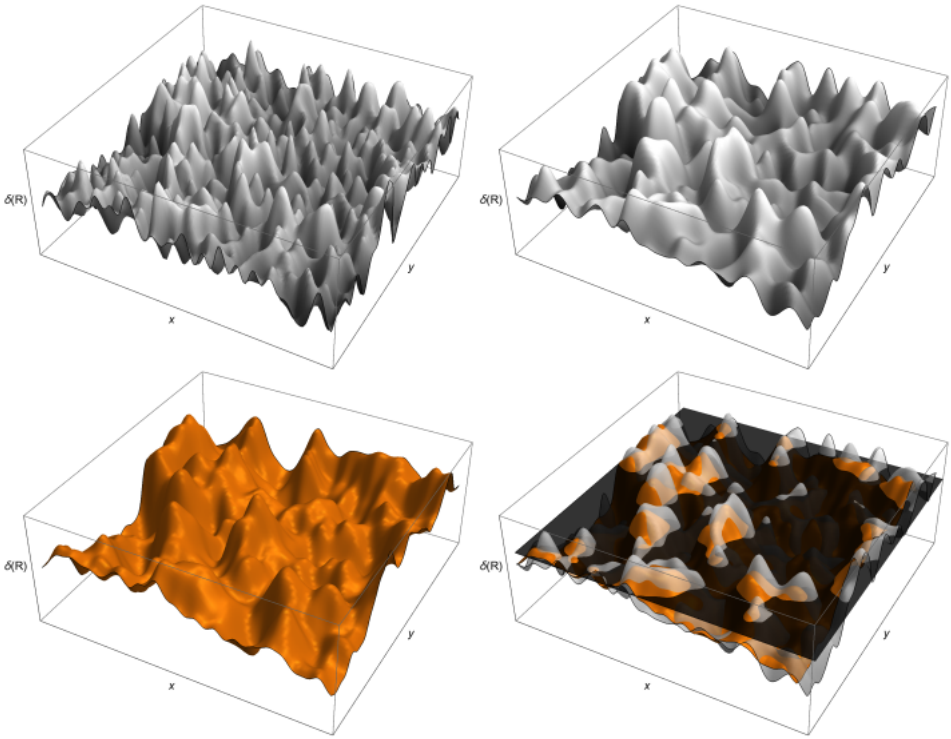


Figure 3.1: Illustration of a two-dimensional Gaussian density field $\delta(R)$ in a side length of 50 $h^{-1}$ Mpc smoothed on different scales $R$: 0.4 $h^{-1}$Mpc (top left), 1 $h^{-1}$Mpc (top right), 2 $h^{-1}$Mpc (bottom left). The fields with R = 1 $h^{-1}$Mpc and R = 2 $h^{-1}$Mpc, along with the spherical collapse threshold ($\delta_{\rm cr} = 1.686$, black plane), are superimposed in the bottom-right panel. Note that there are regions which are above threshold at the larger smoothing scale R = 2 $h^{-1}$Mpc but below the threshold at R = 1 $h^{-1}$Mpc, what illustrates the cloud-in-cloud problem. Desjacques et al. (2018)

## 3.3 Excursion Sets

The main assumption of the excursion set formalism is that a Lagrangian point $\mathbf{q}$ belongs to a halo of radius $R$ if its radius is the *maximum smoothing scale* at which the smoothed linear overdensity

$$\delta(R) \equiv \delta_R(\mathbf{q}) = \int d^3x\, W_R(|\boldsymbol{x}|)\, \delta^{(1)}(\mathbf{q} + \boldsymbol{x}) \tag{3.4}$$

exceeds the critical overdensity $\delta_{\rm cr}$ at the time considered (Desjacques et al., 2018). The most common filtering kernels $W_R$ are described in Section 2.2. Note that, since the shape of the filters in real space reflect the Lagrangian density profile of halos, the sharp-$k$ filter is not physically motivated, as it can have negative values in real space.

To find the halo in which a given point in the Universe belongs to, we consider how $\delta(R)$ centered at this point changes as a function of smoothing scale $R$. We begin at $R \to \infty$, where the density fluctuations are zero, and decrease the scale evaluating the overdensity at each $R$. The density contrast at any given smoothing scale fluctuates around zero with corresponding variance $S(R) \equiv \sigma^2(R)$. The smoothing scale at which the density contrast *first reaches* $\delta_{\rm cr}$ give us the mass scale $M(R)$ of the halo to which the given point belongs. The first-crossing conditions assures that we are treating isolate objects and keeping track of the underdensities inside them, thus naturally solving the cloud-in-cloud problem. The excursion set approach thus yields a mass function which is properly normalized, allowing us to estimate the value of some bias parameters of dark-matter halos, although in general they under-predict the formation of high-mass halos.

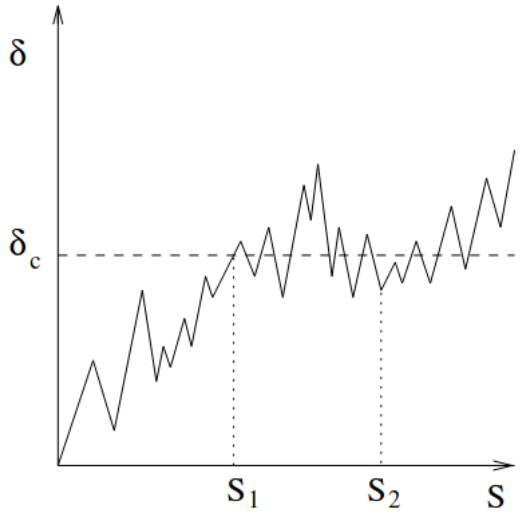


Figure 3.2: Random walk trajectory for $\delta$ as a function of $S \equiv S(R)$. An object has crossed the critical threshold at $S_1 \equiv S(R_1)$, but at a certain point $S_2 \equiv S(R_2)$ (where $R_2 < R_1$) it lays below the threshold, what again characterizes the cloud-in-cloud problem. In the excursion set approach, this trajectory will be associated to an object of $R_1$, as it represents the "time" at which the trajectory *first crosses* the threshold. (Maggiore and Riotto, 2010a)

The overdensity trajectories as a function of smoothing scale can be thought of as random walks, where $R$ is the time variable. Since $S(R)$ is a monotonic decreasing function of $R$, we can use it alternatively as the time variable. The solutions to the excursion set equations can be obtained numerically using Monte-Carlo techniques (Bond et al.,

1991), while analytically we have exact solutions in two scenarios. First, when the steps of the random walk are completely independent, i.e. the Markovian case with sharp-$k$ filter (Bond et al., 1991), or when the steps are completely correlated. However, as we discussed before, these do not represent the realistic cases of halo formation, even though it is possible to expand around these two filters in order to obtain solutions for more physically motivated filters, such as a top-hat (see Section 2.2).

However, any of them do not agree well with the bias parameters measured from numerical simulations (Desjacques et al., 2018), leading us to the need of a more realistic model of halo collapse. In the next section, we are going to extend the assumption that halo collapse has spherical symmetry, through which we obtained the constant barrier $\delta_{\rm cr}$ in the spherical collapse model, to the *ellipsoidal collapse model* using *stochastic barriers*.

## 3.4 Ellipsoidal Collapse Model

Overdense regions of the Gaussian density field are in general triaxal (Doroshkevich, 1970; Bardeen et al., 1986), and not spherical as the spherical collapse model assumes. Also, this models does not take into account the role of the tidal shear $K_{ij}$, although this quantity has been show influence structure formation (Hoffman, 1986; Del Popolo and Gambera, 1998).

Despite the complicated shape of actual protohalo patches seen in N-body simulations, we can study them through the collapse of an isolated homogeneous ellipsoid and approximate the dynamics of a ellipsoidal collapse into a *moving excursion-set barrier*. Since shear tends to slows down the collapse of low-mass halos, its effect can be translated into a barrier $B(M)$ that decreases monotonically with $M$. In this model, the relative abundance of high-mass halos increases, in such a way that the halo mass function is in better agreement with the results from N-body simulations when compared to the usual excursion set approach (Sheth and Tormen, 2002).

While the spherical collapse has a dependence only on one scalar quantity of the tidal field, which is its trace $\delta$ (see Section 2.3), the critical density of the non-spherical collapse also depends on the other two invariants constructed from the tidal field: $\mathrm{tr}[(K_{ij})^2]$ and $\mathrm{tr}[(K_{ij})^3]$. The three invariants of the tidal field can be parametrized as

$$\delta, \quad K_2 \equiv \frac{3}{2}\mathrm{tr}(K_{ij}^2), \quad K_3 \equiv \frac{9}{2}\mathrm{tr}(K_{ij}^3)\,. \tag{3.5}$$

Note that $K_2$ is the tidal anisotropy $\alpha$ that we obtained in Section 2.3, so that $K_2 = 3e^2 + p^2$, where $e$ and $p$ are the shear ellipticity and prolateness. The factor of $3/2$ ensures that $\langle K_2 \rangle = S$, implying that for Gaussian initial conditions, $5K_2$ is a $\chi^2$-distributed with 5 d.o.f., while $9/2$ is added such that $|K_3| \leq K_2^{3/2}$.

In order to calculate the tidal bias parameters, we have to develop a model for the barrier $B$ characterizing the first-crossing of multi-dimensional random walks $(\delta, K_2, \ldots)$. However, we do not have yet a complete theory of Lagrangian bias with the dependence of the galaxy number density on $K_{2\ell}$ and $K_{3\ell}$ separately. While the bias parameters such as $b_1$ and $b_2$ can be considered as a response of a long wave-length density perturbation $\delta$, the tidal bias can be seen as a response to the long wave-length perturbation $K_{2\ell}$ and $K_{3\ell}$. Sheth et al. (2013) considered though a simple model where the barrier is given by

$$B(S, K_2) = \delta_{\rm cr} \left(1 + \sqrt{\frac{K_2}{K_c}}\right), \tag{3.6}$$

where $K_c$ is the characteristic scale for the effect of the tidal shear. The dependence of the barrier on $S$ is only given by $K_2 \propto S$. Due to the fact that $K_2 \geq 0$, we have that $B \geq \delta_{\rm cr}$. Now, we have a first-crossing distribution of 2-dimensional walks that start from a given $\delta_\ell$ and $K_{2\ell}$.

## Fuzzy Barrier

Numerical studies of the ellipsoidal collapse and Lagrangian halos show that the collapse barrier is "fuzzy" (Sheth et al., 2001), in the sense that, for a fixed $S$, there is a large range of possible $\delta_{\rm cr}$, as one can see in Figure 3.3. For each dark matter halo, one can trace back its dark matter particles to the initial position and calculate its mean density with a tophat sphere centered at the initial center-of-mass. These overdensities linearly extrapolated to $z = 0$ indicate the snapshot of the collapse barrier. The mean barriers increase with decreasing halo mass, as expected from the ellipsoidal collapse model. The scatter $\Sigma$ around the mean barrier is generated by other quantities beside the density such as the tidal shear, which correlate with the large-scale environment.

This scatter in the barrier is actually a consequence from the fact that we are projecting the actual, multi-dimensional collapse barrier onto a simpler sub-space in which $S$ is the only variable. The dependence of halo collapse on the tidal shear tensor $K_{ij}$ can only be fully captured by the moving barrier $B(S)$ if the shear is explicitly accounted for.

An alternative to introducing the tidal shear in the barrier explicitly is thus to use a "fuzzy" moving barrier that includes the scatter around the mean such as the square-root barrier (Paranjape et al., 2012)

$$B(S) = \delta_{\rm cr} + \beta\sqrt{S}\,, \tag{3.7}$$

where $\beta$ can be calibrated in simulations. This approximation turns to be more convenient than introducing an explicit dependence on the tidal shear $K_{ij}$ and other fields for the purposes of calculating the halo mass function and the bias parameters. This ansatz

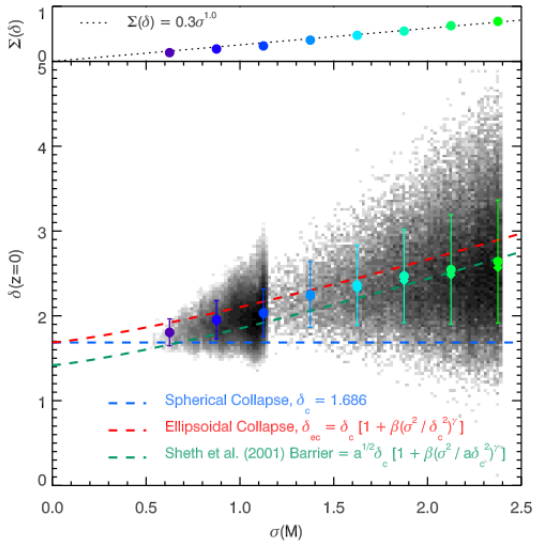


Figure 3.3: Smoothed linear overdensity, extrapolated to $z = 0$ for the Lagrangian regions that collapse and form halos by $z = 0$, as a function of halo mass parametrized through $\sigma(M)$ . Circles correspond to the mean overdensities, whereas diamonds show the median overdensities. Shown for comparison are the spherical collapse barrier (blue dashed line), the ellipsoidal collapse barrier of (Sheth et al., 2001) (red dashed line), and the collapse barrier associated with the Sheth-Tormen mass function (Sheth and Tormen, 1999) (green dashed line). The upper panel shows the scatter $\Sigma$ in the barrier height as a function of $\sigma(M)$. From Robertson et al. (2009). (Desjacques et al., 2018)

is adopted in the excursion-set peak approach. Alternatively, the scatter has also been modeled as a "diffusive barrier" in the context of the path integral approach (Maggiore and Riotto, 2010b). The resulting mass functions are in good agreements with simulations (Achitouv et al., 2013).

## 3.5 The Peak Formalism

To understand the formation of dark matter haloes, we will start by studying how they are connected to their progenitors in the initial density field.

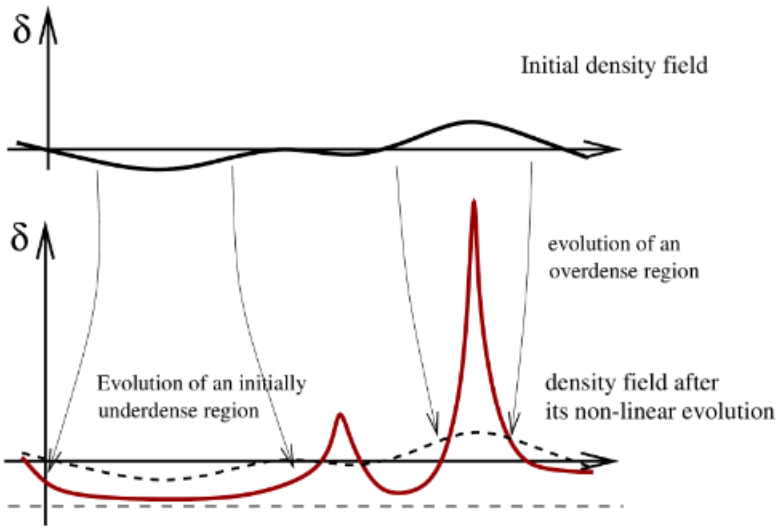


Figure 3.4: Evolution of the initial density field. (Bernardeau et al., 2002)

First, we will naturally assume that the initial smoothed density field extrapolated to the collapse epoch using linear theory, $\delta_R(\mathbf{q}) \equiv \delta_R^{(1)}(\mathbf{q})$, to be a homogeneous Gaussian random field of zero mean. If we denote $\{\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_p, \ldots\}$ by the positions of a set of points referring to e.g. the positions of halo centers, we can write their Lagrangian comoving number density as a sum of Dirac distributions

$$n_g(\mathbf{q}) = \sum_p \delta_D(\mathbf{q} - \mathbf{q}_p)\,. \tag{3.8}$$

As shown by Kac, Rice and BBKS, since we are assuming $\delta_R$ to be a homogeneous Gaussian random field of zero mean, the number density of their *stationary points* $n_{\rm sp}(\mathbf{q})$ can be entirely expressed in terms of the normalized dimensionless variables $\nu \equiv \delta_R/\sigma_0(R)$ (note that this differs from the peak significance $\nu_c \equiv \delta_{\rm cr}/\sigma_0(R)$ ) and its derivatives, $\eta_i \equiv \partial_i\delta_R/\sigma_1(R)$ and $\zeta_{ij} \equiv \partial_{ij}\delta_R/\sigma_2(R)$, where the spectral moments are defined as

$$\nu \equiv \frac{\delta_R}{\sigma_0(R)} \left(\neq \nu_c \equiv \frac{\delta_{\rm cr}}{\sigma_0(R)}\right) \quad \eta_i \equiv \frac{\partial_i\delta_R}{\sigma_1(R)} \quad \zeta_{ij} \equiv \frac{\partial_{ij}\delta_R}{\sigma_2(R)} \tag{3.9}$$

$$\sigma_n^2(R) \equiv \int_{\mathbf{k}} k^{2n} W_R^2(k) P(k), \tag{3.10}$$

where $W_R(k)$ is a spherically symmetric filter and $P(k)$ is the linear matter power spectrum. In order for the peak constrain be consistent, we have to use a filtering kernel $W_R(k)$ (see Section 2.2) such that $\sigma_n(R)$ converges at least up to $n = 2$. While this is not true for a top-hat filter, this is the case for a Gaussian filter and, apparently, for the actual filter $W_R(k)$ which can be measured by tracing back the halos of N-body simulations to the initial conditions. For example, the Lagrangian density profile of halos Dalal et al. (2008)

indicates that the actual filter is more extended than a tophat, but more compact than a Gaussian. Here we will assume a Gaussian filter for simplicity.

The number density stationary points, which are defined as points $\mathbf{q}_p$ satisfying $\boldsymbol{\eta}_i(\mathbf{q}_p) = 0$, can be written as

$$n_{\rm sp}(\mathbf{q}) = \delta_D[\boldsymbol{\eta}(\mathbf{q})] \left| \frac{\partial \boldsymbol{\eta}(\mathbf{q}_p)}{\partial \mathbf{q}_p} \right| . \tag{3.11}$$

In the neighbourhood of a stationary point, we have that

$$\eta_i(\mathbf{q}) = \eta_i(\mathbf{q}_p) + \partial_j \eta_i(\mathbf{q}_p)(q - q_p)^j + \mathcal{O}\left((\mathbf{q} - \mathbf{q}_p)^2\right) \tag{3.12}$$

$$= \frac{\sigma_2}{\sigma_1} \zeta_{ij}(\mathbf{q}_p)(q - q_p)^j + \mathcal{O}\left((\mathbf{q} - \mathbf{q}_p)^2\right) \tag{3.13}$$

$$= \frac{\sqrt{3}}{R_1(R)} \zeta_{ij}(\mathbf{q}_p)(q - q_p)^j + \mathcal{O}\left((\mathbf{q} - \mathbf{q}_p)^2\right) , \tag{3.14}$$

where $R_1$ is the mean distance between stationary point defined from

$$R_n(R) \equiv \sqrt{3} \frac{\sigma_n(R)}{\sigma_{n+1}(R)} , \tag{3.15}$$

which gives an ordered sequence of characteristic lengths $R_{n+1} \geq R_n$, where $R_0$ is the typical separation between zero-crossings of the density field. $R_0$ and $R_1$ are the only scales involved in the calculation of the peak correlation functions. If the Hessian $\zeta_{ij}$ is invertible, we can thus write the *Kac-Rice formula*

$$n_{\rm sp}(\mathbf{q}) = \sum_p \delta_D(\mathbf{q} - \mathbf{q}_p) = \frac{3^{3/2}}{R_1^3} |\det \zeta_{ij}(\mathbf{q})| \, \delta_D[\boldsymbol{\eta}(\mathbf{q})] , \tag{3.16}$$

or

$$n_{\rm sp}(\mathbf{q}) = \left( \frac{\sigma_1}{\sigma_2} \right)^3 |\det \zeta_{ij}(\mathbf{q})| \, \delta_D[\boldsymbol{\eta}(\mathbf{q})] , \tag{3.17}$$

which holds for any smooth random field. In order to restrict our counting to density maxima, we have that $\zeta_{ij}(\mathbf{q}_p)$ must be negative definite and impose a certain threshold height $\nu_c$, such that the localized number density of "BBKS peaks" (Bardeen et al., 1986) of height $\nu_c$ then reads

$$n_{\rm pk}(\mathbf{q}) = \frac{3^{3/2}}{R_1^3} |\det \zeta_{ij}(\mathbf{q})| \, \delta_D[\boldsymbol{\eta}(\mathbf{q})] \, \Theta_H(\lambda_3) \, \delta_D[\nu(\mathbf{q}) - \nu_c] , \tag{3.18}$$

where $\lambda_3$ is the smallest eigenvalue of $-\zeta_{ij}$.

## 3.6 Excursion-Set Peaks

The ideas of excursion sets, which rely on random-walks at all points of space, can be conciliated with the peak formalism approach, which only rely on a subset of points where initial density maxima are located. The combination of the BBKS peak constraint with the up-crossing condition has been dubbed excursion-set peaks (ESP) (Musso and Sheth, 2012; Paranjape and Sheth, 2012). The up-crossing condition guarantees that peaks smoothed on a scale $R$ are not embedded in a peak associated to a larger smoothing scale, i.e., it solves the peak-in-peak effect (with analogy to the cloud-in-cloud problem, see Section 3.2).

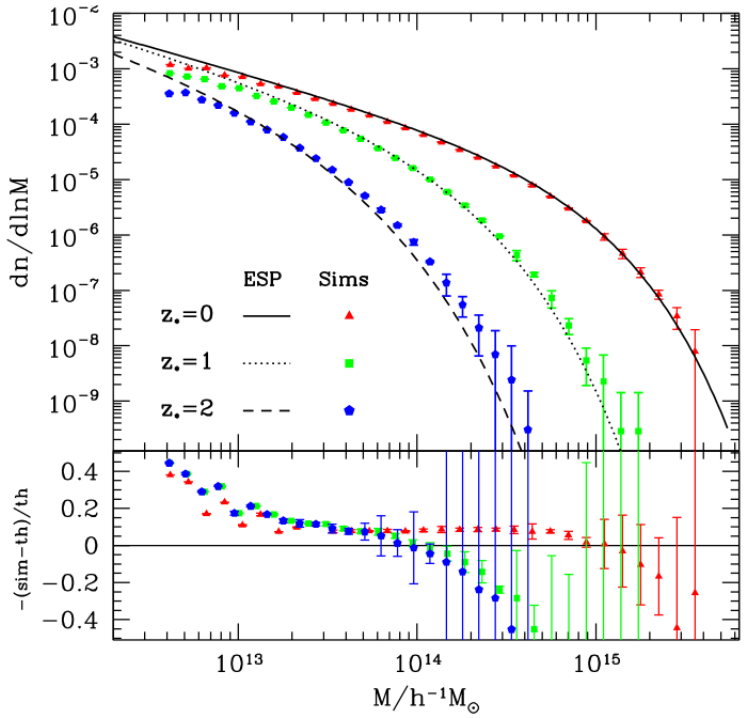


Figure 3.5: Top panel: Logarithmic mass function of SO halos (with $\Delta$ SO = 200) extracted from N-body simulations. Different symbols refer to different redshifts as indicated in the figure. The solid, dotted and dashed curves represent the ESP prediction at z = 0, 1 and 2. Bottom panel: Fractional deviation of the simulations from the ESP prediction. In both panels, error bars denote the scatter among realizations. From Moradinezhad Dizgah et al. (2016).

### ESP$\tau$

Castorina et al. (2016) incorporated the effect of shear in the ESP model by introducing a dependence of a fuzzy barrier on $\tau$,

$$\tau^2 \equiv \frac{3q_{(3)}^2 + (l_2 - \gamma\sqrt{15}y)^2 + (l_3 - \gamma\sqrt{5}z)^2}{(1-\gamma^2)}\,, \tag{3.19}$$

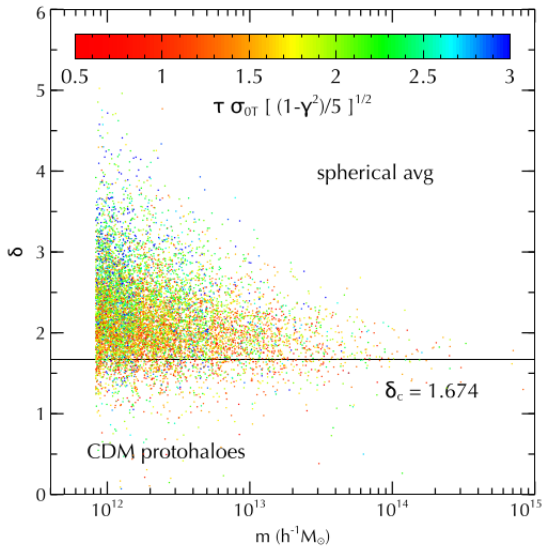


Figure 3.6: Distribution of the $z = 0$ protohalo overdensity in the WMAP3 simulation as a function of mass, coloured by the measured values of $\tau\sigma_{0T}\sqrt{(1-\gamma^2)/5}$. Horizontal line shows the spherical collapse value $\delta_{\rm cr} = 1.674$ (the spherical collapse threshold for the cosmology considered). The dots correspond to $10^4$ randomly chosen halos with more than 200 particles.

where

$$q^2_{(3)} \equiv \frac{l_4^2 + l_5^2 + l_6^2}{3}\,, \tag{3.20}$$

which captures the *misalignment between the tidal shear and shape tensor* as it depends only on off-diagonal shear components in the shape eigenbasis. They thus consider the barrier

$$B(S) = \delta_{\rm cr} + \beta\sqrt{S}\,\tau\,\sqrt{(1-\gamma^2)/5}\,, \tag{3.21}$$

where, as in Equation (3.7), $\beta$ is a free parameter to be fitted from N-body simulations at $z = 0$. We can see in Figure 3.6 that there is a correlation indeed in the scatter of halo overdensities with the $\tau$ parameter. The idea is that high-sheared regions, which are expected to have a higher $\tau$, tend to slow down the collapse of halos, as the different directions of the shear eigenvectors will have different collapse times. Halos of same mass in high-sheared regions have thus to be formed in high-density regions.

As we are going to see in Section 7.2, this model, which is known as ESP$\tau$, will constitute one of our hypothesis for the origins of high-mass spin bias.

# CHAPTER 4

# How do we study halos?

The goal of this chapter is to introduce one of the main tools for the study of dark matter halos and structure formation in the Universe, which is through the use of numerical N-body cosmological simulations. These are powerful tools which allow us to investigate the clustering, properties and evolution of halos up to non-linear scales, where our theoretical frameworks usually break down. Furthermore, with the advent of hydrodynamical N-body simulations, we are able to study the dynamics of galaxies and the halo-galaxy connection. This leads us to a better understanding of the rules of structure formation in the Universe, therefore allowing us to efficiently extract cosmological information from the upcoming surveys.

For example, the comparison of the analytical models presented in Chapter 3 with numerical simulations can be done by tracing back the dark-matter particles containing at a halo to the initial conditions, which can be easily done as particles carry a unique ID across the entire simulation.

Here, we will explore some of the state-of-the-art numerical simulations which were used throughout this work, discuss about how we can find halos in these simulations, analyse which are their main properties and finally see how we can use these tools to learn about the connection between halos and galaxies.

While the last chapter was devoted to introduce the theoretical tools that we have so far, here we will become acquainted with numerical analysis in N-body simulations, so that we will be able not only to develop an analytical framework for spin bias but also to test it in numerical simulations.

## 4.1 N-body Simulations

N-body simulations are numerical tools which consists on distributing a large number of particles inside a box according to the initial conditions given by inflation and then evolving those particles in time by solving the Poisson equation for their mutual gravitational

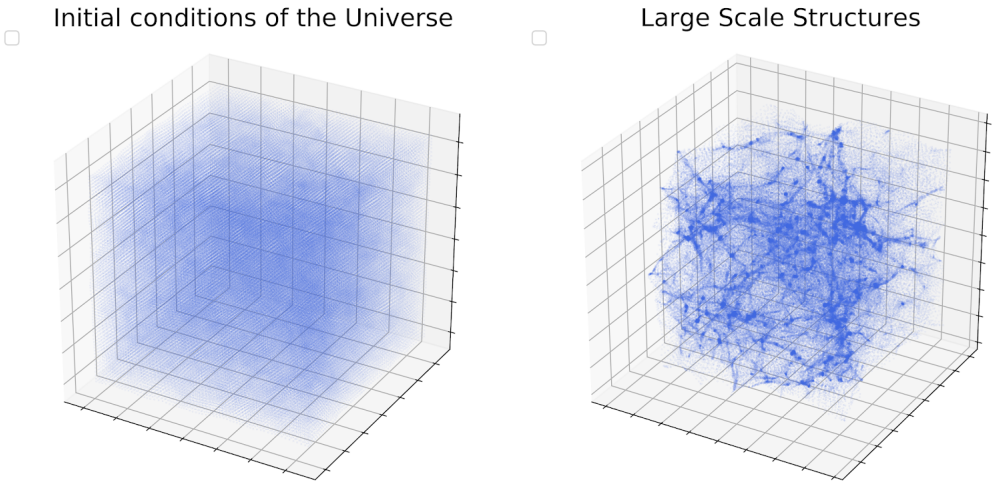


Figure 4.1: (Left) The initial distribution of matter at the beginning of the N-Body simulation. (Right) The final distribution of matter at the final snapshot of the simulation after structure formation. TensorFlow Blog.

interaction. Each time step of the simulation produces a snapshot, and the final snapshot corresponding to $z = 0$ reproduces the patterns seen in the LSS of our Universe.

Of course, this final distribution may differ one from another depending on the cosmological parameters (besides possible issues related to numerical resolution and equation solvers, for example). It is also possible to add extra interaction between the particles, such as a cross-section between the dark matter particles motivated by particle physics models of self-interacting dark matter (SIDM). These modifications are very useful for us to test how the LSS observables, such as the power spectrum, (sub)halo mass function, splashback radius, etc, would change if we change the physics behind the simulation. Therefore, these modified N-body numerical simulations can give us clues on how signs of new physics would manifest themselves in the data extracted from observations, thus helping us understanding the aforementioned open questions in cosmology.

## MultiDark

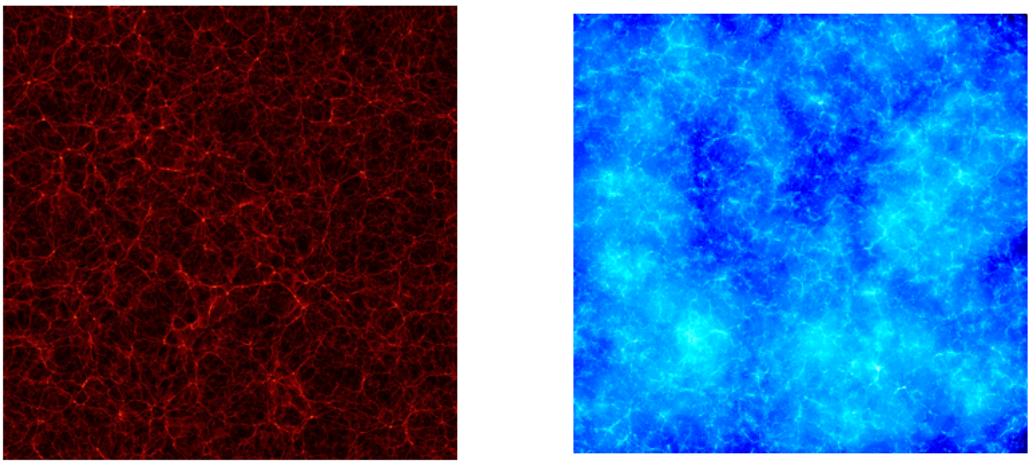

Figure 4.2: (Left): Slice through SMDPL at redshift z=1 (Stefan Gottlöber, IDL). (Right): Slice through MDPL2 at redshift z=0. The size of points is correlated with their mass, while overlaying dots are emphasized by brighter color. (Kristin Riebe, Topcat)

The MultiDark suite of cosmological dark matter only simulations spans boxes of different sizes. As we can see in the table below, with them we are able to reach a large range of scales and masses. Therefore, they constitute a very useful tool to study e.g. halo clustering and the Baryon Acoustic Oscillations. For us, they will be valuable tools for us to analyse secondary halo bias, since they allow us to measure the properties of halos and their clustering. The catalogs used in this work follow a Planck cosmology are downloaded from the CosmoSim database[1].

In order to calculate the correlation function for a large mass range and thus estimate the secondary bias of dark matter halos, a combination of 5 MultiDark boxes with different sizes is made (see the methodology in Section 6.1). As Figure 4.3 displays, each box has a different mass function and thus covers a different mass range.

[1] https://www.cosmosim.org/cms/files/rockstar-data/

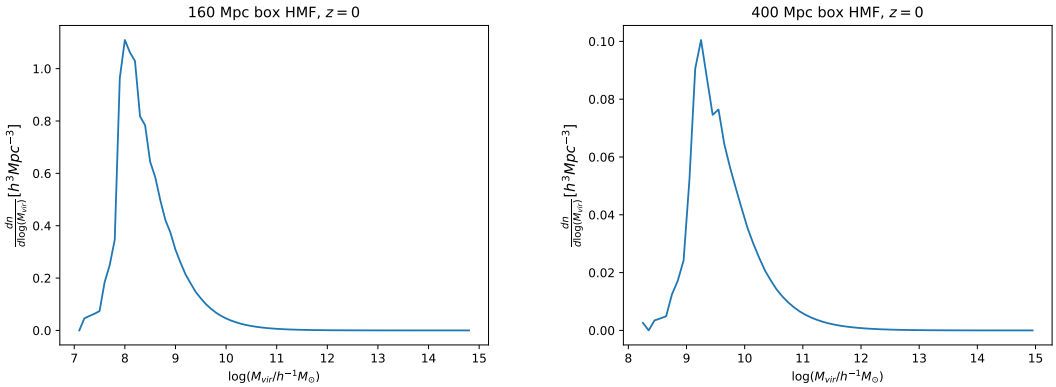


Figure 4.3: (Left): Halo mass function (HMF) of VSMDPL at $z = 0$. (Right): HMF of SMDPL at $z = 0$.

Dark matter halos are basically a collection of dark matter particles in the simulation which have collapsed when reaching a certain density threshold. There are several halo finders which we can use to detect these halos in the simulation, given the distribution of dark matter particles in the box. The halo catalogues of MultiDark are produced using the ROCKSTAR (Robust Overdensity Calculation using K-Space Topologically Adaptive Refinement) halo finder (Behroozi et al., 2012), which is based on a refinement of friends-of-friends groups (see Section 4.1) in six phase-space dimensions and one time dimension. ROCKSTAR identifies dark-matter halos and their substructures, as well as tidal features. We can see in Section 4.1 how the halo catalog produced by ROCKSTAR looks like. It consists on a huge file for each snapshot ($\sim$ 100 GB) with several halo properties, allowing us to analyze many of the halo internal and environmental properties.

```
#scale(0) id(1) desc_scale(2) desc_id(3) num_prog(4) pid(5) upid(6) desc_pid(7) pha
ntom(8) sam_Mvir(9) Mvir(10) Rvir(11) rs(12) vrms(13) mmp?(14) scale_of_last_MM(15)
 vmax(16) x(17) y(18) z(19) vx(20) vy(21) vz(22) Jx(23) Jy(24) Jz(25) Spin(26) Brea
dth_first_ID(27) Depth_first_ID(28) Tree_root_ID(29) Orig_halo_ID(30) Snap_num(31)
Next_coprogenitor_depthfirst_ID(32) Last_progenitor_depthfirst_ID(33) Last_mainleaf
_depthfirst_ID(34) Tidal_Force(35) Tidal_ID(36) Rs_Klypin(37) Mmvir_all(38) M200b(3
9) M200c(40) M500c(41) M2500c(42) Xoff(43) Voff(44) Spin_Bullock(45) b_to_a(46) c_t
o_a(47) A[x](48) A[y](49) A[z](50) b_to_a(500c)(51) c_to_a(500c)(52) A[x](500c)(53)
 A[y](500c)(54) A[z](500c)(55) T/|U|(56) M_pe_Behroozi(57) M_pe_Diemer(58) Halfmass
_Radius(59) Macc(60) Mpeak(61) Vacc(62) Vpeak(63) Halfmass_Scale(64) Acc_Rate_Inst(
65) Acc_Rate_100Myr(66) Acc_Rate_1*Tdyn(67) Acc_Rate_2*Tdyn(68) Acc_Rate_Mpeak(69)
Acc_Log_Vmax_Inst(70) Acc_Log_Vmax_1*Tdyn(71) Mpeak_Scale(72) Acc_Scale(73) First_A
cc_Scale(74) First_Acc_Mvir(75) First_Acc_Vmax(76) Vmax\@Mpeak(77) Tidal_Force_Tdyn
(78) Log_(Vmax/Vmax_max(Tdyn;Tmpeak))(79) Time_to_future_merger(80) Future_merger_M
MP_ID(81)
```

The properties of halos however need a large amount of dark-matter particles for them to be well defined. Thus, in order to maintain a reasonable numerical resolution, a mass cut is made, in such a way that only halos with more than $\sim$500 particles are taken into account. The size, number of particles, mass resolution and the mass cut (in $h^{-1}M_{sun}$ and log) are shown below for the MultiDark boxes used.

| | VSMDPL | SMDPL | MDPL2 | BigMDPL | HugeMDPL |
|---|---|---|---|---|---|
| size Mpc/$h$ | 160 | 400 | 1000 | 2500 | 4000 |
| particles | $3840^3$ | $3840^3$ | $3840^3$ | $3840^3$ | $4096^3$ |
| res $M_{sun}/h$ | $6.2 \times 10^6$ | $9.6 \times 10^7$ | $1.5 \times 10^9$ | $2.4 \times 10^{10}$ | $7.9 \times 10^{10}$ |
| cut $M_{sun}/h$ | $> 3.1 \times 10^9$ | $> 4.8 \times 10^{10}$ | $> 7.5 \times 10^{11}$ | $> 1.2 \times 10^{13}$ | $> 3.9 \times 10^{13}$ |
| log $M_{sun}/h$ | $> 9.49$ | $> 10.68$ | $> 11.88$ | $> 13.08$ | $> 13.60$ |

## Illustris

The Illustris Project are magneto-hydrodynamical simulations with dark matter only and also dark matter with baryonic content simulations, where it is implemented state-of-the art physical models of galaxy formation and evolution. They represent an excellent tool to study the physical mechanisms behind the galaxy-halo connection.

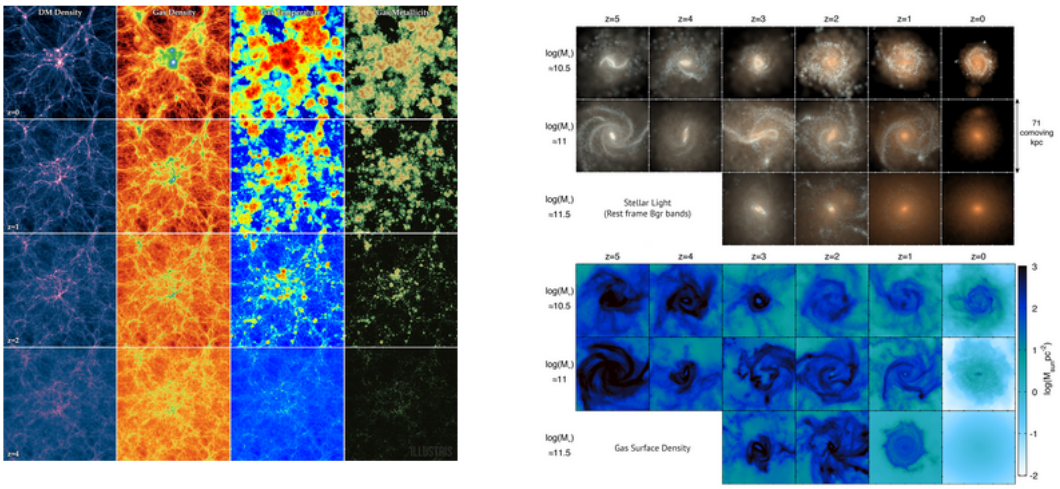


Figure 4.4: (Left): Redshift evolution of a whole box slice from $z = 4$ to $z = 0$, showing four projections: dark matter density, gas density, gas temperature, and gas metallicity. (Right): Sample of massive galaxies from $z = 5$ (left) to $z = 0$ (right), showing their stellar light distribution (rest-frame B,g,r band composites) and their gas surface densities. Illustris Collaboration.

Several boxes with different sizes and particle resolution are available as well. In order to solve the physics of galaxy formation, the resolution of hydrodynamical simulations has to be considerably higher than a dark matter only simulation, allowing for the study of small-scale physics such as star formation, AGN feedback, etc. However, the drawback is that these boxes usually have a small volume due its high computational cost.

The recent project of IllustrisTNG has made available some hydrodynamical boxes on larger scales, being the TNG300 (300 Mpc side length) the largest one. Such boxes allow us for the study of the large-scale clustering of halos and galaxies, besides the connection between their small-scale properties, as we have show in Montero-Dorta et al. (2020) (see

Section 4.3). Furthermore, together with TNG100 (100 Mpc side length), we were able to probe how the kinetic Sunyaev-Zeldovich effect, which arises due to the scattering of photons on the electrons of the gas in galaxy clusters, can be used as a observational probe of halo spin bias (see Section 7.3).

Dark matter halos are defined in Illustris with the Friends-of-Friends algorithm (Knebe, 2011), where particles are linked together if their distance lies below a certain threshold, called linking length. This means that the distances of particles at the boundary of such a linked object (a "FOF group") are smaller or equal than the linking length, corresponding to a density threshold. FOF groups cannot intersect, so that a particle can be assigned uniquely to just one FOF group (for a given linking length). Further, there are small "sub-halos" living inside the dark-matter halo, which are substructures defined with smaller linking lengths lying completely within their host.

### Uchuu

The Uchuu[2] dark matter only simulations (Ishiyama et al., 2021) can be downloaded from http://skiesanduniverses.org/Simulations/Uchuu/ and uses a Planck2015 cosmology. It was run by the supercomputer ATERUI II in Japan.

Uchuu has impressive particle resolution when compared to other simulations of the same type, allowing us to analyze the clustering of low-mass halos on very large scales. For example, the 2 $h^{-1}$Gpc Uchuu box and the 2.5 $h^{-1}$Gpc BMDPL MultiDark box have $12800^3$ and $3840^3$ particles, respectively. We can reach the range of halo masses $\log(M_{\rm vir}/h^{-1}M_\odot) \sim 11.0 - -14.0$. Its halo catalogs are also produced using `ROCKSTAR`.

## 4.2 Halo Properties

### Virial Mass and Radius

As one can imagine, there are several properties we can define for halos, as we do for galaxies. The most essential one is halo mass. However, this apparently simple property is not trivially defined. In numerical simulations, for example, we can define it as the sum of all particles bounded to the halo (e.g., inside a FOF group as in Illustris). A practical and perhaps the most used definition of halo mass is the virial mass, i.e., the mass inside the radius in which the virial theorem applies. In order to understand it, we have to remember that the critical density of the universe is defined as

$$\rho_{\rm cr} = \frac{3\,H^2(t)}{8\pi G}\,, \tag{4.1}$$

[2] "Uchuu" is a Japanese word meaning "Universe".

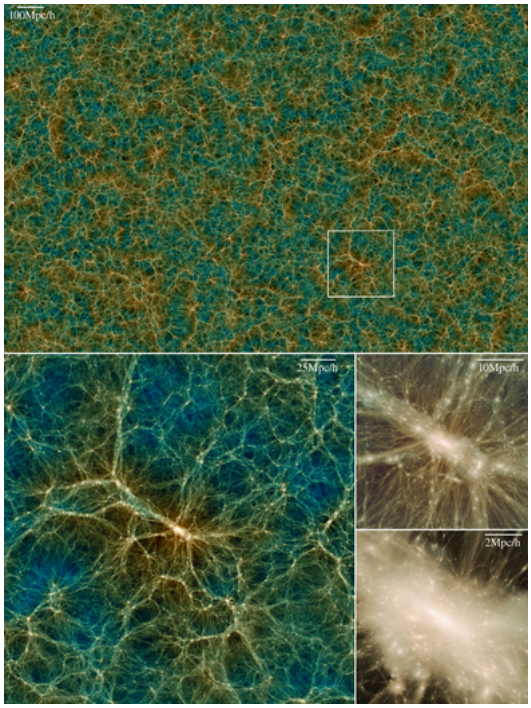


Figure 4.5: Dark matter distribution in Uchuu at $z = 0$. The image shows a 2000 $h^{-1}$Mpc $\times$ 2000 $h^{-1}$Mpc projected volume with a thickness of 25 $h^{-1}$Mpc. The white box in the top panel is the same region visualized in the left-bottom panel, in which the spatial volume is equivalent to the Bolshoi simulation. Uchuu Images.

where $H(t)$ is given by the Friedmann equations according to the considered cosmology (Dodelson, 2003).

If we define the radius $R_\Delta$ at which the mean density of the halo is equal to a certain constant $\Delta$ times the critical density of the universe, i.e., $\bar{\rho}(R_\Delta) = \Delta\,\rho_{\rm cr}$, we have that the mass $M_\Delta$ of this halo is given by

$$M_\Delta = \frac{4}{3}\pi R_\Delta^3\,\Delta\,\rho_{\rm cr}\,. \tag{4.2}$$

If we want to estimate $\Delta$ to be the density threshold inside which the virial theorem holds, then using the spherical collapse model (see Chapter 3) for an Einsten de-Sitter universe we obtain that $\Delta_{\rm vir} \approx 178$ (Bryan and Norman, 1998), from which we can define the *virial mass* of halos $M_{\rm vir}$. It is very common in the literature for people to use simply $M_{200}$ for the virial mass.

## Splashback Radius

The virial radius of halos may not be a good definition of the halo boundary. Sometimes there are some objects which are still gravitationally bounded to the object, but that

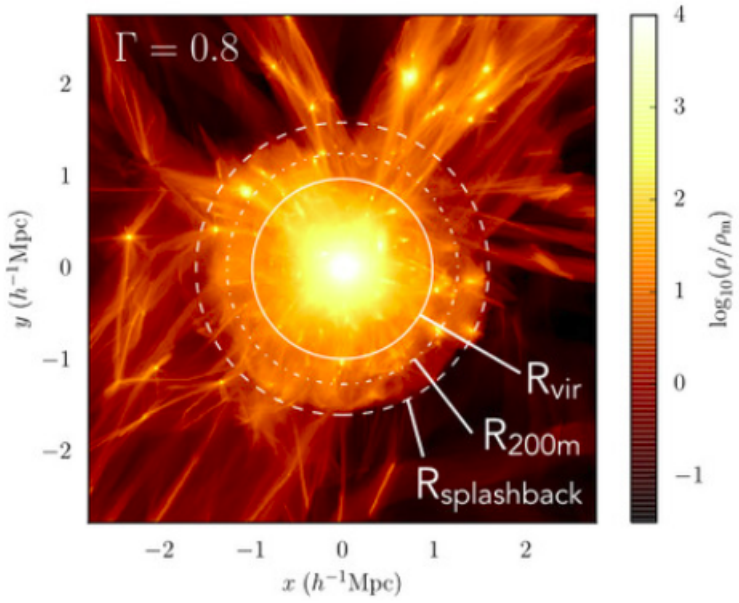


Figure 4.6: Comparision between the splashback radius $R_{\rm sp}$, $R_{200}$ and $R_{\rm vir}$ in a simulation. From http://www.benediktdiemer.com/research/splashback/.

happen to be in orbit outside the virial radius of their host halo. These "ex-subhalos" are known as *splashback* (or backsplash) halos, living in the so-called *splashback radius* $R_{\rm sp}$, so that we can also define a corresponding mass $M_{\rm sp}$. The splashback radius covers a few times the virial radius of the host (Wang et al., 2009; Adhikari et al., 2014) and is claimed to be a "more physical" boundary of halos (More et al., 2015). Its location can be in principle be detected in real data (More et al., 2016) and it has been shown to depend on accretion rate, redshift, cosmology (Diemer et al., 2017), cosmic expansion and gravity (Adhikari et al., 2018).

Besides mass, there are several other halo properties we can define. As we discussed in Chapter 1, halo mass is the *primary* halo property in the sense that it is the property to which halo bias is most sensitive. However, at fixed halo mass, bias also depend the *secondary* halo properties, such as spin, concentration and age.

## Halo Spin

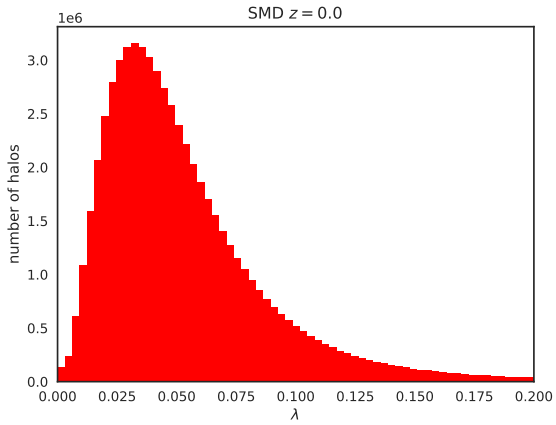


Figure 4.7: Spin distribution in SMDPL at $z = 0$.

Spin, a dimensionless quantity which characterizes the angular momentum of halos, is given by

$$\lambda = \frac{|J|}{\sqrt{2} M_{\rm vir} V_{\rm vir} R_{\rm vir}} , \tag{4.3}$$

where $J$ is the halo angular momentum calculated inside the virial radius $R_{\rm vir}$ and $V_{\rm vir}$ is its circular velocity at $R_{\rm vir}$. Naturally, we could also calculate this quantity with other density threshold besides the virial one.

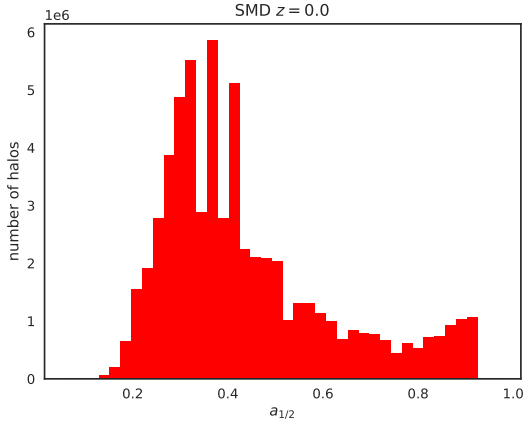


Figure 4.8: Spin distribution in SMDPL at $z = 0$.

## Halo Age

Age, $a_{1/2}$, is defined as the scale factor at which half of the peak mass of the halo was accreted.

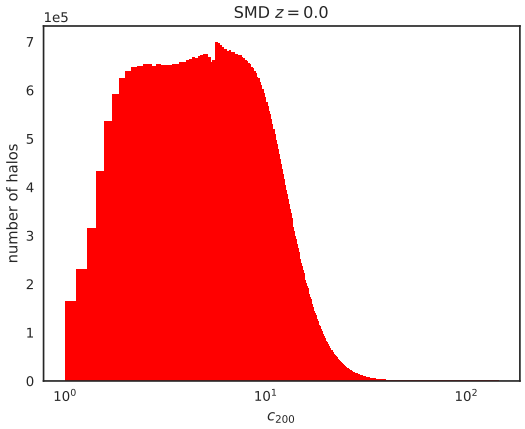


Figure 4.9: Concentration distribution in Illustris TNG300 at $z = 0$.

## Halo Concentration

We know that halo age correlates with halo concentration, which is defined as

$$c_{\text{vir}} = \frac{R_{\text{vir}}}{R_s} \tag{4.4}$$

where $R_s$ is the Kyplin scale radius (Klypin et al., 2011) from the NFW density profile Navarro et al. (1996),

$$\rho(r) = \frac{\rho_0}{(r/R_s)(1 + r/R_s)^2}\,, \tag{4.5}$$

a fitting function for the density profile of halos that was found to fit well to every dark matter halo.

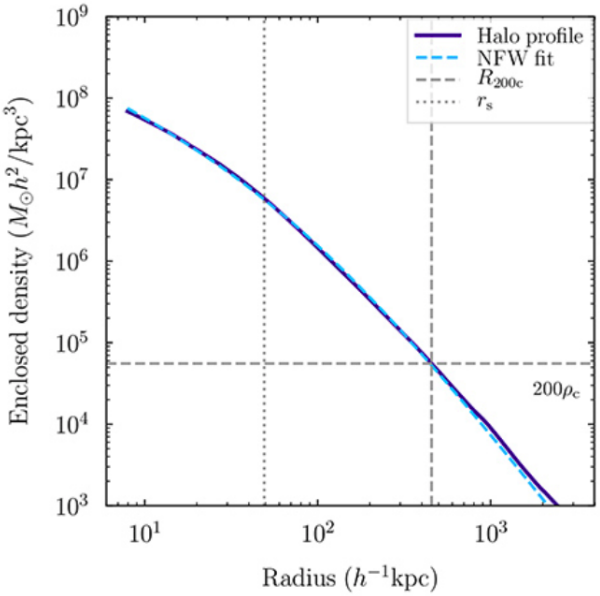


Figure 4.10: Halo density profile (purple) and NFW fit (blue), where the dashed lines show the locations of $R_s$ and $R_{200}$. Taken from http://www.benediktdiemer.com/visualization/theory-toy-models/.

## 4.3 Galaxy-Halo Connection

As we saw in Chapter 1, according to our cosmological observations the ratio between the amount of baryons to dark matter in the Universe is 1/5. In our current cosmological models, basically every galaxy form within a dark matter halo, in such a way that their properties and distribution throughout the Universe are intimately connected.

Currently we am interested in better understanding such connections and how they can give us hints on galaxy clustering, formation and evolution models. At the end of the day, most of our data comes from galaxies (and will come in the near future). Then, making this connection can help us testing our analytical models of the large scale structure, besides from extracting information on cosmological parameters and e.g. the nature of dark matter.

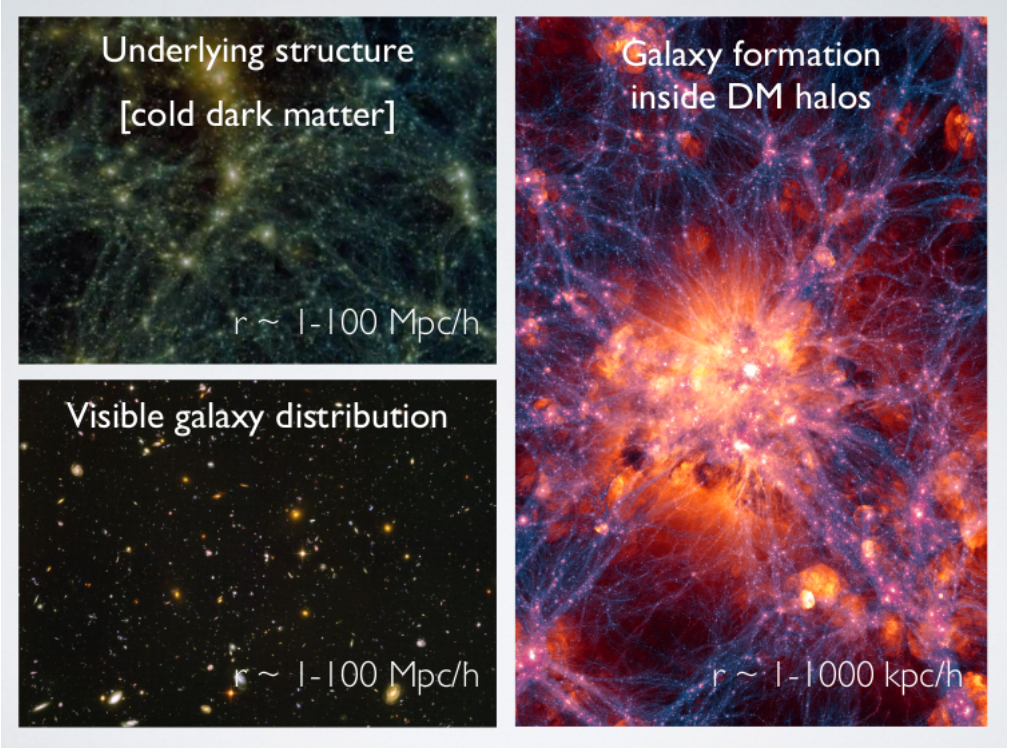


Figure 4.11: (Top left): Dark matter distribution in the Uchuu N-body simulation. (Bottom left): Galaxies in the Hubble Deep Field from HST. (Right): Illustris hydrodynamical simulation. (Diagram borrowed from Antonio Montero-Dorta)

In a recent work (Montero-Dorta et al., 2020), we measured with high precision and analyzed the properties and clustering of galaxies with respect to dark matter halos in the magneto-hydrodynamical IllustrisTNG300 simulation. In order to illustrate the galaxy-halo connection, we can see for example in the left hand plot of Figure 4.12 how the (g-i) colour of galaxies is related to the mass and age of the halo where the galaxy lives. A higher (lower) value of (g-i) colour corresponds to a redder (bluer) galaxy. High-mass halos are only inhabited by very red galaxies and, at fixed halo mass, older halos tend to host redder galaxies, but the correlation seems stronger at the low-mass end. In the right hand plot we see that halo assembly bias is manifested in the clustering of galaxies: at fixed halo mass, older halos and redder galaxies are more clustered than younger halos and bluer galaxies, but the signal vanishes at the high-mass end.

We did similar analyses for star formation rate, stellar radius, galaxy spin and velocity dispersion, illustrating how secondary halo bias manifests itself on galaxies as well.

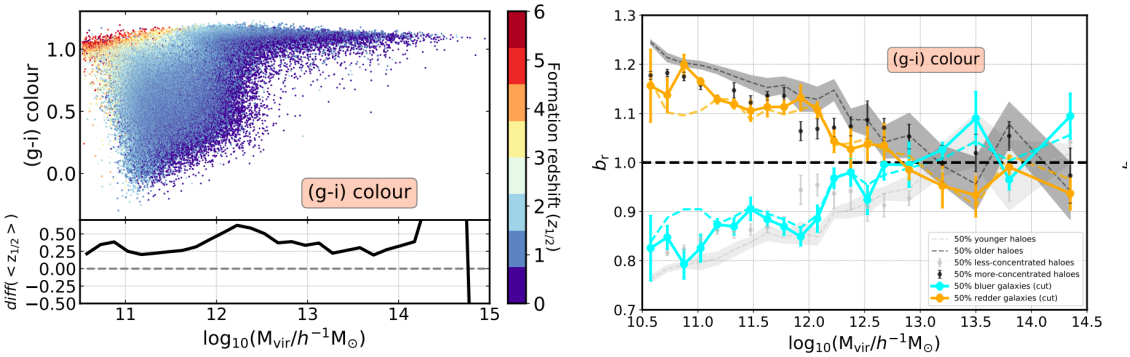


Figure 4.12: (Left): Distribution of central galaxy (g-i) colour as a function of halo mass. The colour code indicates the age of the halo where each galaxy lives, which is described by the halo formation redshift. (Right): Galaxy assembly bias with 50% redder (bluer) galaxies in orange (blue); halo assembly bias with 50% older (younger) halos in darker (lighter) grey tones. Montero-Dorta et al. (2020)

# CHAPTER 5

# How do halos spin?

Halo spin is a dimensionless quantity which characterizes the angular momentum of dark matter halos. In the definition of Bullock et al. (2001), halo spin is given by:

$$\lambda = \frac{|J|}{\sqrt{2} M_{\rm vir} V_{\rm vir} R_{\rm vir}} , \tag{5.1}$$

where $J$ is the halo angular momentum defined in the usual way, and $V_{\rm vir}$ is its circular velocity at the virial radius $R_{\rm vir}$. For a spherically symmetric object the spin is basically the ratio between its angular velocity and the velocity needed for it to be rotationally supported. In the next sections, we will learn how we can predict the angular momentum of halos, and especially TTT will represent an important tool for the development of our theory for the origin of high-mass spin bias.

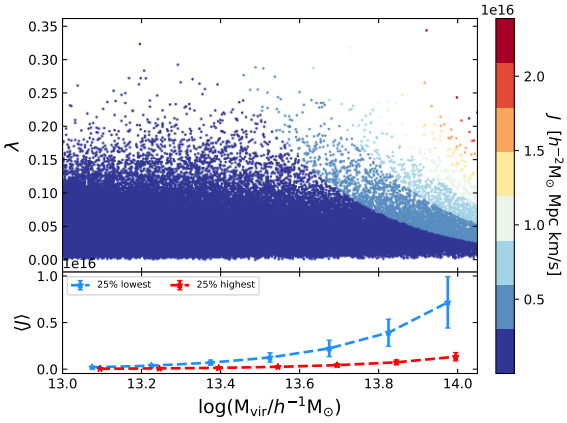


Figure 5.1: (*Upper*): Relation between spin and halo mass colored by halo angular momentum in HMDPL at $z = 0$. (*Lower*): mean angular momentum as a function of halo mass.

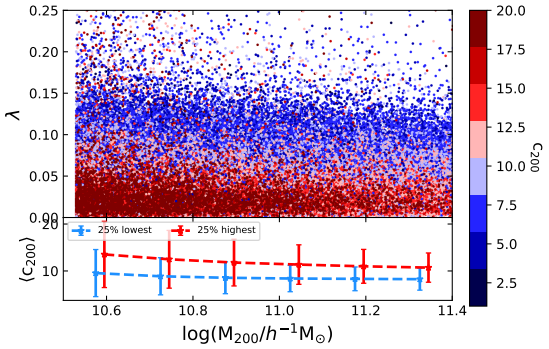


Figure 5.2: Relation between halo spin and halo mass colored by halo concentration in TNG300 at $z = 0$.

## 5.1 Tidal Torque Theory

Peebles (1969) was the first to estimate the angular momentum acquisition of protogalaxies, where they showed to be a second-order effect by using first-order approximations for the density and velocity fields to calculate the growth of a spherical region in an expanding universe. Doroshkevich (1970) showed however that by relaxing the spherical symmetry condition that the angular momentum grows at first order. These calculations were further developed by White (1984), who improved the model and discussed some of its limitations. They obtained that the coupling between the first-order tidal field with the zero-order anisotropy of a collapsing object produces a first-order torque, which results in the angular momentum growing to first-order and proportional to time (Heavens and Peacock, 1988).

These studies gave rise to the *Tidal Torque Theory* (TTT), the most well accepted theory for the origin of angular momentum. In this theory, halo spin arises due to the misalignment between its shape tensor and the shear field exerted by the surrounding matter distribution. TTT accurately predicts the evolution of angular momentum only during linear and quasi-linear regimes of structure formation, while at late times non-linear processes such as mergers tend to also affect its evolution (Porciani et al., 2002; López et al., 2019).

We are going to consider the evolution of the angular momentum of a protohalo in a FLRW Universe, where dark matter is considered a presurelles cold fluid. The origin of the Cartesian coordinate system is assumed to coincide with the centre of mass. Since we are interested in the intrinsic angular motion, we disregard the centre–of–mass motion. Thus, inside an Eulerian volume $V$ it is is given by

$$\mathbf{L}(t) = \int_{a^3V} d\mathbf{r}\, \rho(t)\, \mathbf{r} \times \mathbf{v}\,, \tag{5.2}$$

where $\mathbf{r}(t) = a(t)\mathbf{x}$ is the physical distance and $\mathbf{x}$ is the comoving one, while $\mathbf{v}(t) = d\mathbf{r}/dt$

is the velocity field and $\mathbf{u}(t) = ad\mathbf{x}/dt$ is the peculiar one. Writing it as a function of comoving coordinates, we obtain

$$\mathbf{L}(t) = \overline{\rho}\, a^4(t) \int_V d\mathbf{x}\, (1+\delta)\, \mathbf{x} \times \mathbf{u}\,, \tag{5.3}$$

where we wrote the density as a function of its fluctuations $\rho = \overline{\rho}(1+\delta)$. We can also write this expression in Lagrangian coordinates, where the evolution of angular momentum of a spherical region of volume $V_L$ is

$$\mathbf{L}(t) = \overline{\rho}\, a^3(t) \int_{V_L} d\mathbf{q}\, (\mathbf{q} + \mathbf{S}) \times \frac{d\mathbf{S}}{dt}\,. \tag{5.4}$$

As stated in the Zel'Dovich approximation (Zel'Dovich, 1970), in linear regime matter motion can be described by $\mathbf{x}(t) = \mathbf{q} + \mathbf{S}(\mathbf{q}, t)$, where $\mathbf{S}(\mathbf{q}, t)$ is the displacement vector, which can be written with its spatial and temporal components decoupled, i.e. $\mathbf{S}(\mathbf{q}, t) = \mathbf{f}(\mathbf{q})D(t)$. Moreover, due to the relation between the velocity field and the gravitational potential in linear regime, we have that its spatial part is due only to the initial matter surroundings. We can thus write $\mathbf{x}(\mathbf{q}, t) \approx \mathbf{q} + D(t)\nabla\psi(\mathbf{q})$ and obtain

$$\mathbf{L}(t) = a^2(t)\dot{D}(t) \int_{V_L} d\mathbf{q}\, \bar{\rho}(t) a^3(t)\, \mathbf{q} \times \nabla\psi(\mathbf{q})\,. \tag{5.5}$$

Assuming that the gradient $\nabla\psi(\boldsymbol{q})$ doesn't change much inside the Lagrangian volume $V_L$, we can perform a Taylor expansion,

$$\partial_\alpha\psi(\mathbf{q}) \simeq \partial_\alpha\psi(\mathbf{0}) + q^\beta\, \mathcal{T}_{\alpha\beta} \tag{5.6}$$

where we identified the tidal tensor $\mathcal{T}_{\alpha\beta} \equiv \partial_\alpha\partial_\beta\psi(\mathbf{0})$. Thus, we finally arrive to the expression

$$\boxed{L_\alpha(t) = a^2(t)\dot{D}(t)\, \varepsilon_{\alpha\beta}{}^{\gamma}\, \mathcal{T}^{\beta\sigma}\, \mathcal{I}_{\sigma\gamma}} \tag{5.7}$$

where we have identified the inertia tensor

$$\mathcal{I}_{\sigma\gamma} \equiv \int_{V_L} d\mathbf{q}\, \bar{\rho}(t) a^3(t)\, q_\sigma q_\gamma\,. \tag{5.8}$$

Note that we can write this expression as a function of the shear and shape tensors (which are the detraced part of the tidal and inertia tensors, respectively) as

$$\boxed{L_i(t) = a^2(t)\dot{D}(t)\, \varepsilon_{ijk}\, K^j{}_l\, I^{lk}} \tag{5.9}$$

We can also note that, for a spherical region, $\mathcal{I}_{\sigma\gamma} \propto q^5\delta_{\sigma\gamma}$, so that $L_\alpha \propto \varepsilon_{\alpha\beta}{}^{\gamma}\, \mathcal{T}^{\beta}_{\gamma} = 0$,

i.e., there is no angular momentum growth *at first order*.

In an EdS Universe, we have that $a^2\dot{D} \propto D^{3/2} \approx t$, the evolution of angular momentum in halos should then grow as $t \propto a^{3/2}$. They grow until approximately the turnaround time, when and influence of the matter surroundings ceases due to the distance to the collapsed halo. From that point, the angular momentum should remain constant. However, as we mentioned before, looking at numerical N-body simulations we know that this is not the case, as we can see e.g. in Figure 5.3.

Since spin bias is an effect which already manifests itself in high redshifts (see Chapter 7), we can in principle try to use TTT as a toy model to better understand spin bias, as it accurately predicts halo spin and only fails at recent times.

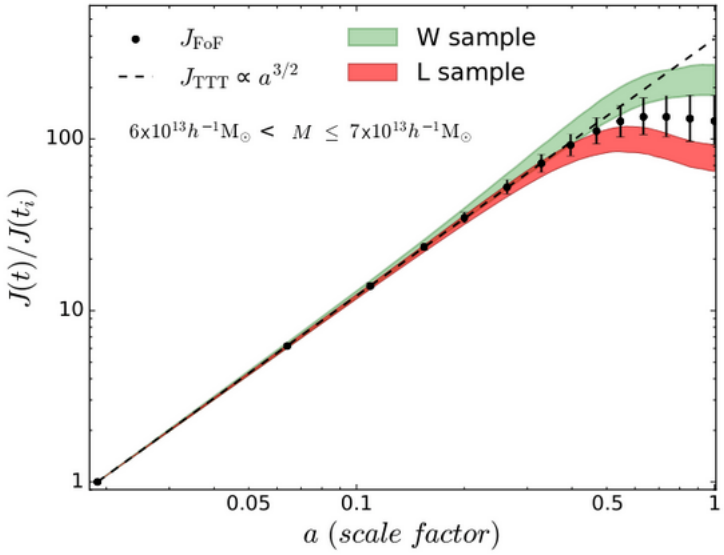


Figure 5.3: Evolution of halo angular momentum $J_{\rm FOF}$ in black circles compared with the TTT prediction $J_{\rm TTT}$ in dashed lines. The colored areas show the interquartile range of the W and L samples, which correspond to the terciles that have respectively most gained and lost angular momentum in each mass bin (López et al., 2019).

## 5.2 Tidal Torque Theory in the Peak Formalism

It has been shown in Catelan and Theuns (1996) that if we restrict that our halos are centred in density maxima peaks, we can write the ensemble average of angular momentum

$$\mathcal{L} \equiv \sqrt{\langle \mathbf{L}^2 \rangle_\psi} = \ell \mathcal{L}_* \,, \tag{5.10}$$

where

$$\mathcal{L}_* \equiv a^2 \dot{D} \eta_o \sigma_0 R_*^5 \tag{5.11}$$

and

$$\boxed{\ell \equiv \frac{96\pi}{\sqrt{15^3}}(1-\gamma^2)^{1/2}\left(\frac{\nu}{\gamma\, x}\right)^{5/2}\frac{\mathcal{A}(e,p)^{1/2}}{\mathcal{B}(e,p)^{3/2}}} \tag{5.12}$$

where the polynomials are defined as

$$\mathcal{A}(e,p) \equiv [p(p+1)]^2 + 3e^2(1-6p+2p^2+3e^2) \tag{5.13}$$

and

$$\mathcal{B}(e,p) \equiv (1-2p)[(1+p^2)-9e^2]\,. \tag{5.14}$$

The parameter $x$ is an indicator of the 'sharpness' of the peak, where sharper peaks have higher $x$; the parameters $e$ and $p$ characterize the asymmetry of the isodensity profile (see Section 2.3). The parameter $e(\geq 0)$ measures the ellipticity of the matter distribution in the plane $(q_1, q_3)$, while $p$ determines the oblateness $(0 \leq p \leq e)$ or the prolateness $(-e \leq p \leq 0)$ of the triaxial ellipsoid. If $e = 0$, then $p = 0$, and the ellipsoid is a sphere. For a sphere, $\mathcal{A}(0,0) = 0$ and $\mathcal{B}(e,p) = 1$.

The distribution of the peak shape parameters $\mathcal{P}(\nu, x, e, p)$ was obtained in Bardeen et al. (1986) and reads

$$\mathcal{P}(\nu,x,e,p) = \mathcal{P}_0\, x^8\, \mathcal{W}(e,p) \exp\left[-\frac{\nu^2}{2} - \frac{5}{2}x^2(3e^2+p^2) - \frac{(x-\gamma\nu)^2}{1(1-\gamma^2)}\right]. \tag{5.15}$$

inside the allowed domain of a triangle formed by $0 \leq e \leq 1/4$, $-e \leq p \leq e$ and $1/4 \leq e \leq 1/2$, $3e-1 \leq p \leq e$, where $\mathcal{P}_0 \equiv 3^2 5^{3/2}/(2\pi)^3 R_*^3 (1-\gamma^2)^{1/2}$ and $\mathcal{W}(e,p) \equiv e(e^2-p^2)\mathcal{B}(e,p)$.

## 5.3 Halo Spin from Major Mergers

As first presented in Vitvitska et al. (2002), it is possible to construct a random walk model in which halo spin is generated by the random accretion of satellites. Periods of enhanced mass accretion tend to increase the angular momentum of halos, while a long period of stacked accretion decrease angular momentum. This model is based on the evolution of the angular momentum of individual halo progenitors, instead of the entire Lagrangian volume as in TTT, and the evolution of the angular momentum of the progenitors is very irregular, instead of the constant increase predicted by TTT.

They do not assume correlation between the angular momenta of accreted objects, and their position is randomly distributed in a sphere around the protohalo. Here, tidal torques produce the random tangential velocities of merging satellites. They use the extended Press-Schechter approximation to calculate the growth of mass, angular momentum, and spin parameter for halos and compare with numerical simulations.

The random walk model naturally explain some features seen in N-body simulation, such as the bumps seen in the angular momentum of the progenitors when they pass through major mergers and the stead decline when only small satellites are accreted. It provides a good fit to the final lognormal distribution of the spin parameter, besides reproducing its weak dependence on mass and redshift.

Since halos which have no major mergers since $z = 3$ tend to have a lower spin, an interesting claim stated there is that halos which rotate faster should host elliptical galaxies and the ones which rotate slower should host spiral galaxies. This is because galaxy color correlate with assembly history and recent major mergers, where elliptical galaxies form from majors of spiral ones.

Note that, in principle, both pictures can be fitted together. An initial region corresponding to a halo of a final higher spin when compared to a low one might both have a higher initial misalignment between the protohalo shape and shear, as TTT predicts, and produce a higher fraction of final bigger satellites, which turn out to be accreted by the halo during its assemble history as predicted in the random walk model.

# PART II

# Secondary Halo Spin Bias

# CHAPTER 6

# Secondary Bias

It is well known that the bias of dark matter halos depends primarily on their masses (Kaiser, 1984). However it has been detected in numerical simulations that, at fixed halo mass, halo clustering also depends on secondary halo properties, such as age Sheth and Tormen (1999); Gao et al. (2004), concentration (Wechsler and Tinker, 2018; Sato-Polito et al., 2018), spin (Gao and White, 2007; Bett et al., 2007), substructure content (Villarreal et al., 2017), etc. Since it was first detected for age and concentration, which correlate with halo assembly history, the effect was denominated *halo assembly bias* at first. Nowadays the name secondary halo bias is preferred, since for other secondary properties the phenomenum seem to be caused by different physical mechanisms than assembly bias.

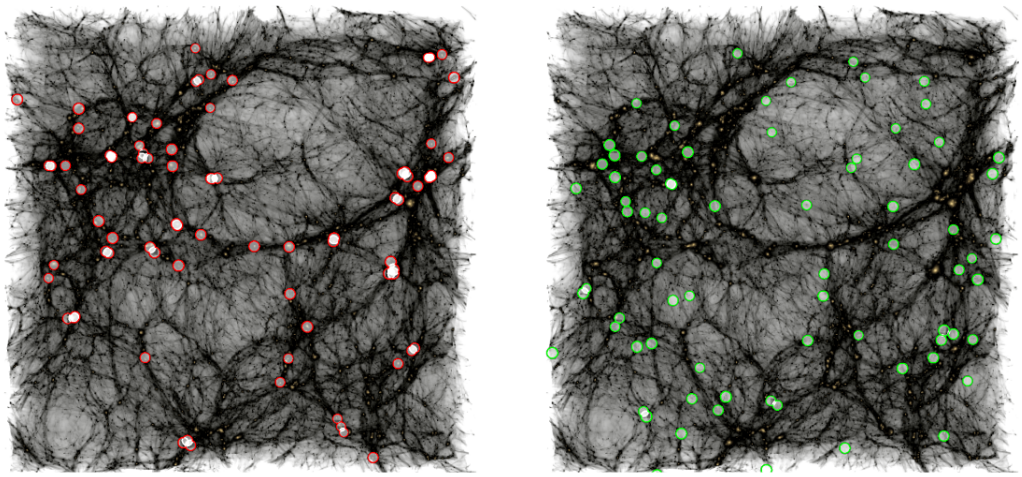

Figure 6.1: 90x90x30 Mpc slice of a cosmological simulation at z = 0. The open red circles indicate the 5% of halos at log M = 10.8 with the highest concentration. We can see that, for halos of this mass, more concentrated (red) halos are in more biased regions than less concentrated (green) halos. (Li et al., 2008; Wechsler and Tinker, 2018)

Understanding these trends is of interest not only for more realistic e.g. Halo Occupation Distribution (HOD) (Zentner, 2007; Hearin et al., 2016) and galaxy formation models

(Wechsler and Tinker, 2018), but also because to this day we do not have neither conclusive observational evidence neither a complete analytical framework to explain these phenomena. Also, separating the population of cosmological tracers into subgroups, e.g. according to their properties, can help in the extraction of cosmological parameters (Mergulhão et al., 2021).

## 6.1 Measuring Secondary Bias

We use the set of high-performance routines `Corrfunc` developed by Sinha (2016) to calculate the two-point correlation function in the aforementioned cosmological N-body simulations. Equipped with these tools, we can study secondary bias by estimating the relative bias of galaxies or halos of a given mass selected according to some secondary property, such as spin, concentration or age.

To quantify the dependence of halo clustering on a secondary property $S$, the relative bias between a subsample of halos selected according to $S$ and all halos in the same primary bias property ($B$) range can be measured by (Sato-Polito et al., 2018; Tucci et al., 2021)

$$b^2(r, B, S) = \frac{\xi(r, B, S)}{\xi(r, B)}\,, \tag{6.1}$$

where the primary bias paramater $B$ along this work will be $\mathrm{M_{vir}}$ and the secondary parameters $S$ can be $\lambda$, $a_{1/2}$ or $c_{200}$. This quantity therefore tell us whether a certain subsample of halos chosen according to a secondary property $S$ tend to be more clustered ($b^2 > 1$) or less clustered ($b^2 < 1$) than the entire sample in a given mass bin. In this work, the analysis is made for halos with the same virial mass, i.e., the same primary bias parameter $B$.

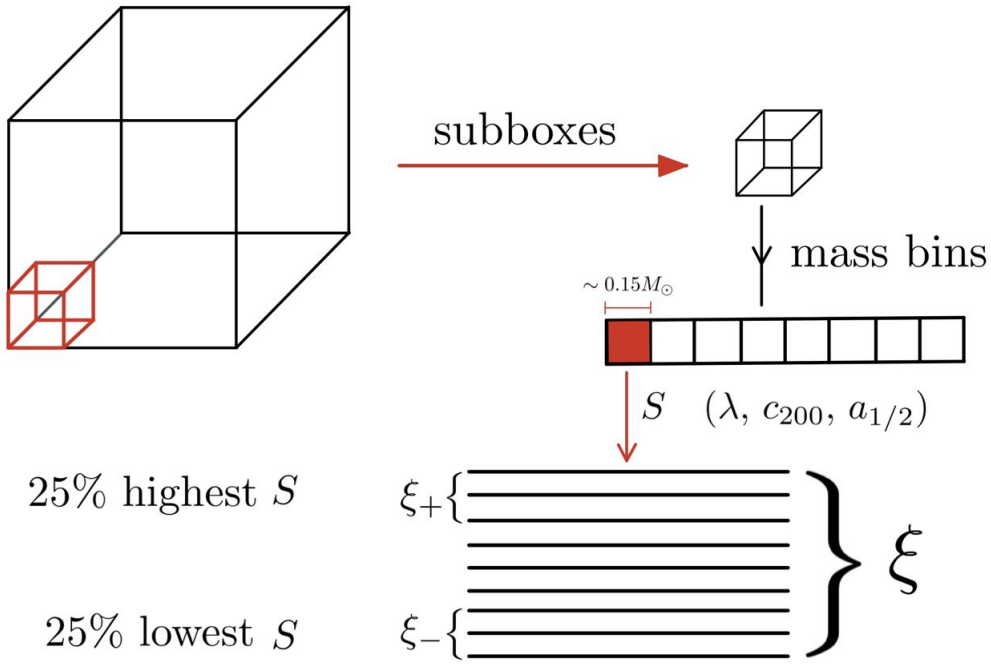


Figure 6.2: Methodology: divide each box into subboxes; divide the halos in each subbox into mass bins; for a fixes mass bin, calculate the correlation function for the population of 25% highest secondary parameter $S$ ($\xi_+$), 25% lowest ($\xi_-$) and entire population ($\xi$). Thanks to Nelson Schuback for the diagram.

Each box is divided into smaller subboxes (e.g., the box MDPL2 with box size 1.000 $h^{-1}$Mpc is divided into 64 subboxes with subbox size 250 $h^{-1}$Mpc - see Chapter 1), in order to have statistics for the halo correlation function.

Each subbox is divided into mass bins (e.g., pick up the halos with mass bewteen 10.6 and 11.8 and divide them in bins with a width of 0.15). The mass range is determined by the halo mass function of each box and by the mass cut of ∼500 particles. Since the primary bias parameter is mass, the analysis of halo bias dependence on its secondary properties has to be made for a fixed mass. However, in order to have a high enough number of halos to calculate the correlation function, halos with a mass in a small mass width are taken, small enough to don't affect the secondary bias analysis.

In each subbox, the halos are sorted with respect to one of the secondary property (e.g., spin) in such a way that they can be divided into 4 subpopulations, including the 25% of the halos which have the highest value of that parameter (e.g., spin) and the 25% of those with lowest values.

The relative bias for a given secondary parameter is calculated in Eq. 6.1 by the correlation function of the 25% highest (or lowest) subpopulation $\xi(r, B, S)$ divided by the correlation function obtained for the entire subpopulation $\xi(r, B)$, thus obtaining the secondary bias $b_+ = \xi_+/\xi$ for the highest and $b_- = \xi_-/\xi$ for the lowest spin quartile. A

further average over the scales is made to obtain the value of the relative bias of each quartile for a given mass bin. In Figure 6.3, we have averaged over the scales of 5 $h^{-1}$Mpc to 15 $h^{-1}$Mpc to calculate the relative bias.

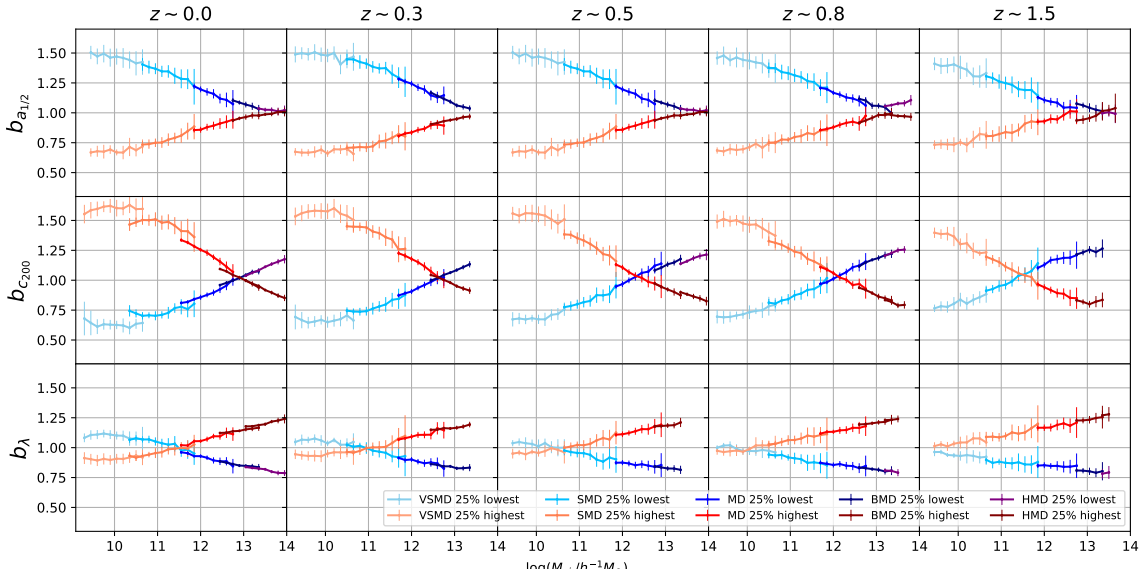


Figure 6.3: Secondary bias for halo age, concentration and spin measured in several MultiDark boxes and redshifts.

## 6.2 Interpretation of the Relative Bias

It is noteworthy that the relative bias measured here should be seen as an "effective" bias, as it does not correspond to any particular bias parameter of the bias expansion (see Section 2.5). Usually secondary bias are measured on highly non-linear scales, where the bias expansion is no longer valid. Unless the measurement is performed on quasi-linear or linear scales, secondary bias should be seen only as a difference of clustering of halo subpopulations (of a given mass) selected according to a secondary property.

The dependence of the large-scale bias parameters as defined in the bias expansion were measured with *separate universe* (SU) simulations (Wagner et al., 2015; Barreira and Schmidt, 2017), as done in e.g. Lazeyras et al. (2017); Ramakrishnan and Paranjape (2020). These are simulations are based on the "peak-background split" (PBS) argument (Kaiser, 1984; Bardeen et al., 1986), which states that for halo and galaxy formation processes a long-wavelength density perturbation acts like a local modification of the background density, such that the expectation value of the density in such a region is given by the average abundance of tracers in a fictitious FRW spacetime with modified physical background density – a "separate universe". Thus, each bias parameter can be seen as a response to the long wavelength perturbation of its corresponding operator, and can

be calculated by taking derivatives of the mass function in simulations with a modified background density. We have implemented SU with Gadget-4 (Springel et al., 2021) and we intend to use these results to investigate the dependence of $b_1$, $b_2$ and $b_{K^2}$ on halo spin (see Section 7.2).

## 6.3 Assembly Bias

Perhaps the most well studied case of secondary bias is assembly bias, i.e., the secondary dependence of halo clustering on halo assembly history. There are several quantities which encode the history of mass accretion of halos, which usually correlate one with another. We can define the half-mass scale $a_{1/2}$, which is the time at which the halo has accreted half of its peak mass (see Section 4.2), and say that old halos as those which accreted most of their mass early on, i.e., those with a low $a_{1/2}$. In Figure 6.3, we can see that old halos more tightly clustered than young halos over the entire mass range, since old halos display a higher relative bias. A crossover at the high-mass end has been discussed in the literature (Sato-Polito et al., 2018) (i.e., an inversion of the trend where very massive young halos turn to be more tightly clustered), but the limited quantity of massive halos in this range does not allow for a statistically significant detection. The crossover can be possibily detected by using different indicators of age with a fraction other than 1/2 used in the definition of $a_{1/2}$ (Chue et al., 2018).

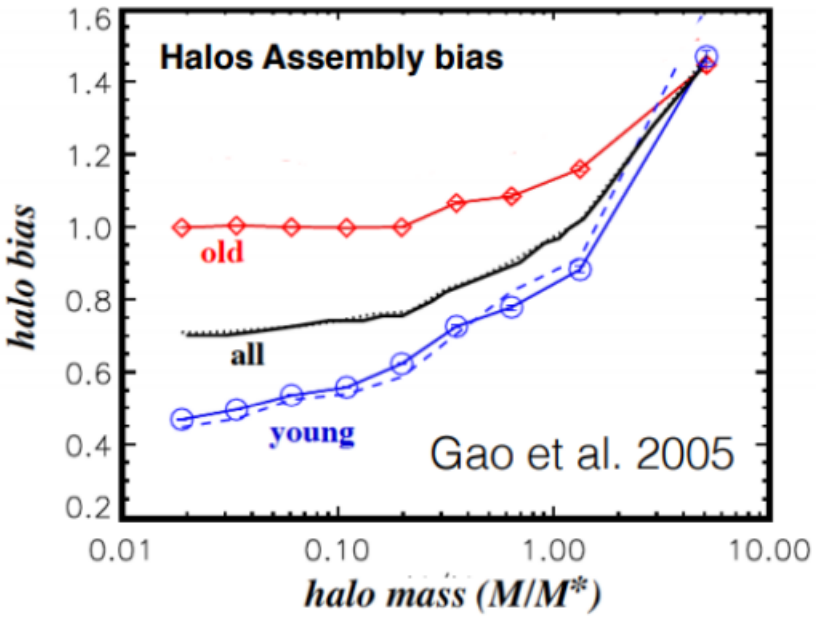


Figure 6.4: Bias at $z = 0$ as a function of halo mass and formation time. Gao et al. (2005)

Since the definition of $a_{1/2}$ is somehow vague and its determination in real data is

impossible, as we do not have access to the mass history of halos and thus neither the time at which they accreted half of their peak mass, a more convenient way to measure assembly bias is through halo concentration, which also encodes the assembly history of halos (see Section 4.2). We can see in Figure 6.3 that divide the concentration trend in two regimes, at high and low masses. At high masses, low-concentrated halos have a higher relative bias than high-concentrated halos, while the opposite is true below the crossover mass of $\log(M_{\rm vir}/h^{-1}M_\odot) \approx 13.0$ at $z = 0$. Note that this is the critical mass $M_*$ at which $\nu = \delta_{\rm cr}/\sigma(z = 0, M_*) = 1$. The concentration trend can be measured also as a function of $\nu$ for all redshifts, and its trend is quite similar (see Figure 6.5), as the peak height can be seen as a measure of halo mass independent of redshift.

The reason behind these trends will be clarified in the next two sections, where we denote "high-mass assembly bias" as the trend of the relative bias of concentration for masses higher than the crossover mass and "low-mass assembly bias" for masses lower than that.

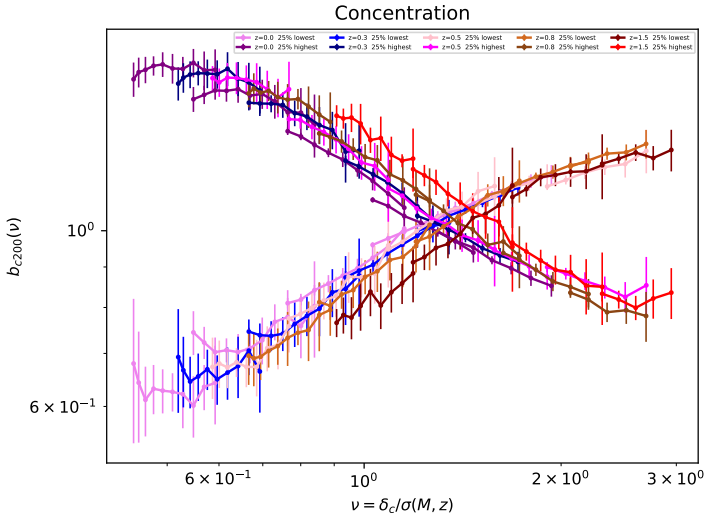


Figure 6.5: Relative bias for halo concentration as a function of $\nu$.

## 6.4 High-Mass Assembly Bias

### In the Peak Formalism

It was presented in Dalal et al. (2008) how high-mass assembly bias can be explained in the peak formalism. At fixed peak height, low concentrated halos are associated with peaks of a low curvature, and therefore of a higher bias. In the high-mass regime, the collapse of halos is nearly spherical (Bardeen et al., 1986), so that the collapse threshold can be well approximated by $\delta_c \approx 1.68$, which is derived from the spherical collapse model

(Gunn and Gott, 1972). Marginalizing over the peak parameters for halos of a given mass, we obtain that $b \sim \delta_c/\sigma^2$ (Press and Schechter, 1974; Cole and Kaiser, 1989). However, halos of different curvatures $s = d\langle\delta\rangle/d(\log M)$ also have different bias.

Knowing that we can estimate the environment density $\delta_{\rm env}$ where halos of a given mass live by the Taylor expansion $\delta_b = \delta + \Delta(\log M) \times d\delta/d(\log M) + \ldots$, we can pick two peaks with curvatures $s_1$ and $s_2$ and note that peak 1 of a shallower curvature (i.e., $s_1 > s_2$ and $|s_1| > |s_2|$) has a larger background density, and thus is associated to a larger bias. The dependence of halo bias on curvature can be estimated as in Bardeen et al. (1986),

$$b_L \approx \sigma^{-1}\frac{\nu - \gamma x}{1-\gamma^2}\,, \tag{6.2}$$

where $\nu = \delta/\sigma$, $x = \delta/\sigma_s$ and $\gamma = \langle \nu x\rangle$.

At high masses, there is also a direct relation between halo curvature and mass accretion history. In the spherical collapse picture, halos collapse when their density $\delta$ reaches $\delta_c$. Therefore, we can estimate the mass accretion history of halos by inspecting how the smoothed overdensity $\langle\delta\rangle$ varies with smoothing scale. We thus expect halo accretion rate to be related to halo curvature as $d(\log M)/d(\log a) = -[d(\log\delta)/d(\log a)]^{-1}$. Since the assembly history of halos and their concentration are correlated, we would expect that halos with a lower concentration to be formed by shallow peaks in the initial density field, which we just saw to be related to a higher bias. We now understand from peak arguments why massive, low-concentrated halos have a higher bias at fixed mass. The interpretation is that halos living in high density environments have an enhanced mass supply, such that they grow faster than halos residing in low density environments.

## In Excursion Sets

Since the peak curvature and the slope $\delta'$ strongly correlates [or completely in the case of a Gaussian filter (Desjacques et al., 2018)], the same conclusions can be obtained in the excursion set formalism (Zentner, 2007; Gao et al., 2005). However, since in assembly bias the small-scale assembly history of halos correlates with the large-scale background density, in this we have to account for correlated random walks instead of the usual Markovian random walks of the standard excursion-set approach (Maggiore and Riotto, 2010a). The steeper the slope $|\delta'|$, the lower $a_{1/2}$ is, since the collapse barrier is reached relatively earlier for halos of the same mass. Furthermore, these old, early-forming halos live in isolated environment (underdense regions with $\delta' < 0$) with less mass accretion at recent times, which leads to steeper halo profiles and higher concentration.

## 6.5 Low-Mass Assembly Bias

At low-masses, insights from the halo formation picture in the initial density field are not enough to understand the assembly bias trend, since non-linear effects turn to be more important. For example, in Figure 6.6 we see that the trend in concentration inverts at the characteristic mass scale (that is, the mass $M^*$ for which $\nu = \delta_c/\sigma(M^*, z) = 1$ at a given redshift $z$).

Mansfield and Kravtsov (2020) claims that there are three main causes of low-mass assembly bias: splashbacks, tidal fields and hot environments.

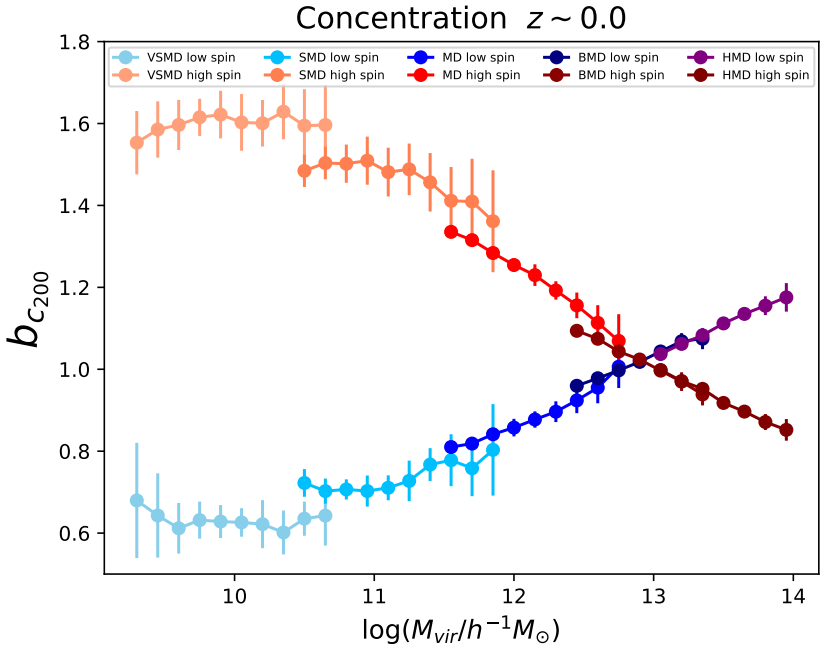


Figure 6.6: Relative bias for halo concentration calculated in several MultiDark boxes at $z = 0$.

It was shown also in Dalal et al. (2008) that *splashback halos* account for part of the assembly bias signal. This was later confirmed and discussed in several works (Wang et al., 2009; Sunayama et al., 2016; Mansfield and Kravtsov, 2020).

### Splashback Halos

Splashback halos are distinct halos that were subhalos at some previous time, i.e., passed through the virial radius of a larger halo.

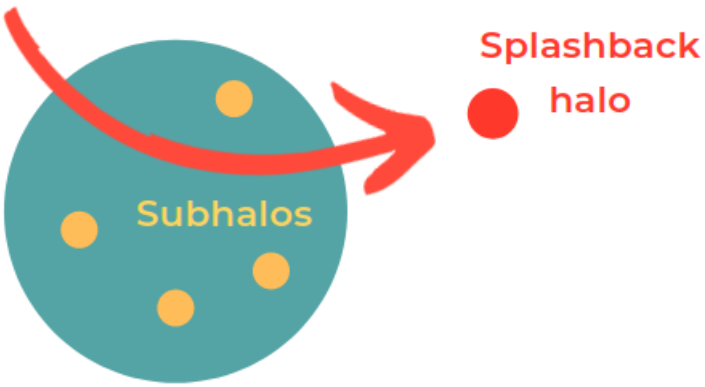


Figure 6.7: Splashback halos were previously subhalos that happen to be outside the virial radius of their hosts.

As we saw in Section 4.2, they live in the vicinity of massive halos in the so-called splashback radius (Wang et al., 2009; Adhikari et al., 2014). Since they still live near to their previous, massive hosts, they have a larger bias than other halos of same mass. Even on large scales these halos will track their previous hosts, such that they will be more strongly clustered. They are typically low-mass "ex-subhalos" which have their mass accretion ceased when inside their previous hosts. As a consequence, these halos are older and more concentrated than other halos of same mass. Their different mass assembly history combined with their higher bias leads to part of the low-mass assembly bias signal.

## Tidal Forces and Hot Environments

Splashback halos alone cannot fully explain low-mass assembly bias. There is still signal when the population is removed (Dalal et al., 2008; Sunayama et al., 2016; Mansfield and Kravtsov, 2020) or when the halo boundary definition is changed, such that splashbacks can lie inside the radius of larger halos (Villarreal et al., 2017).

Another effect which contributes for the low-mass assembly bias is the stalled mass accretion of halos due to tidal forces exerted by the surrounding distribution of matter. This can happen due to a gravitationally dominant massive halo (Hahn et al., 2009; Salcedo et al., 2018) or the structures of the cosmic web region in which the halo is located (Musso et al., 2018).

It could also be possible that small halos located in hot environments with deep potential wells have smaller accretion due to the high velocities of infalling matter (Dalal et al., 2008). Note that although tidal forces and gravitational heating usually are correlated

effects, their cause is different, as the first arises due to large second derivatives of the gravitational potential while the second is due to deep gravitational potentials.

# CHAPTER 7

# Spin Bias

*Halo spin bias* refers to the secondary dependence of halo clustering on *spin*, a dimensionless quantity defined in Section 4.2 which characterizes the angular momentum of halos. It was thought that high-spin halos were more tightly clustered than low-spin halos, independently of their mass (Gao and White, 2007; Bett et al., 2007; Faltenbacher and White, 2010; Lacerna and Padilla, 2012; Villarreal et al., 2017; Lazeyras et al., 2017; Salcedo et al., 2018). However, the research project carried out by Gabriela Sato-Polito, a former student of our group, together with Dr. Antonio Montero-Dorta and Prof. Raul Abramo (Sato-Polito et al., 2018), revealed that this is not the complete picture. They were the first in the literature to show the "low-mass spin bias inversion", i.e., the fact that the spin bias trend inverts at low masses.

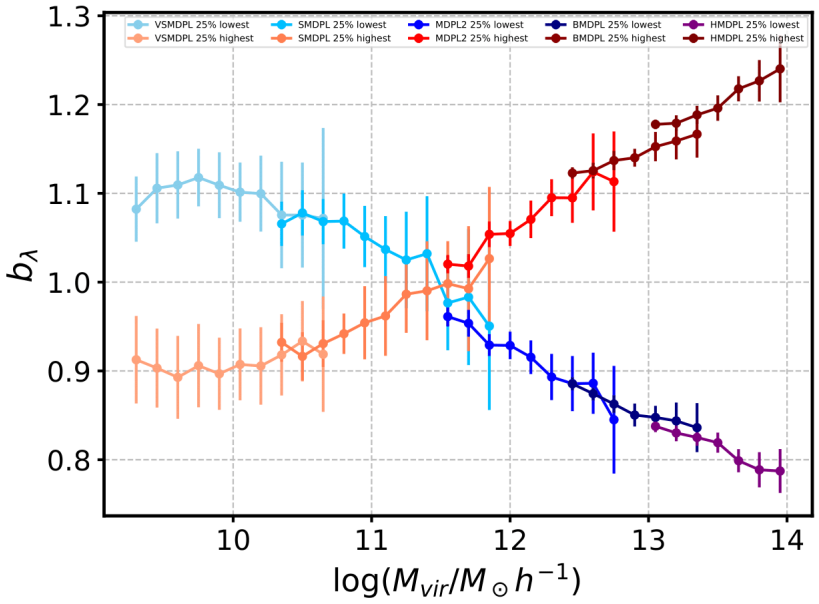


Figure 7.1: Relative spin bias measure with MultiDark boxes at z=0. Red (blue) tones correspond to 25% highest (lowest) spin halos.

In Figure 7.1, we can see this particular shape of the relative bias measured for halo spin in 5 different MultiDark simulations at $z = 0$ (see Section 6.1): at fixed halo mass, high-spin halos are more tightly clustered than low-spin halos, while the opposite is true at the very low-mass end. The inversion was not detected before Sato-Polito et al. (2018) in the literature as the simulations used there did not have enough resolution to probe spin bias at such low masses. This feature was only possible to be detected with the advent of the high-resolution simulation SMDPL (see Section 4.1), which allowed Sato-Polito et al. (2018) to investigate the clustering of halos below the "mass of crossover" $\log(\mathrm{M_c}/h^{-1}\mathrm{M}_\odot) \simeq 11.5$. In Figure 7.1, we further confirmed and extended this analysis with VSMDPL to even lower halo masses. A later work (Johnson et al., 2019) that was done comparing the MultiDark results with 2 other high-resolution simulations (Vishnu and Consuelo) showed the same "spin crossover" at low masses (also for $z = 0$).

Most of the attempts to provide a plausible explanation for the origins of spin bias have focused on the high-mass end (i.e., above the spin crossover). In this context, Salcedo et al. (2018) showed little correlation between spin bias and the proximity to a significantly more massive halo, which seems to discard the massive-neighbour theory. Lacerna and Padilla (2012), on the other hand, attributed the high bias of high-spin halos to their location in the cosmic web, as material from filaments accreted by massive halos (predominantly in high-density environments) can increase the halos' angular momenta. Johnson et al. (2019), introduces the notion of "twin" bias, a tweak to the massive-companion argument by which high-spin halos are slightly more likely to be found near other halos of comparable mass. Johnson et al. (2019), who actually address the entire mass range, further suggests that spin bias could in fact be described by a combination of twin bias and the contribution of other residual secondary dependencies.

The Lacerna and Padilla (2012) theory for the origin of spin bias at the high mass end is that highly biased halos are located in filaments, so that the accretion of matter in preferential directions spins up the halos. However, when Johnson et al. (2019) measure the local anisotropy of halos, defined upon the eigenvalues of the inertia tensor inside a neighbour radius, they find no correlation with halo spin. This seems to contradict Lacerna and Padilla (2012) theory, since halos located in filaments are known to have high anisotropy values when calculated in similar scales to the neighbour radius.

All the above works provide interesting information about spin bias, but a definite explanation for spin bias is yet to be established. The purposes of our work is precisely to provide a full and consistent picture for halo spin bias.

## 7.1 The origins of Low-Mass Spin Bias

Given that Sato-Polito et al. (2018) was the first work to show the low mass spin bias signal, confirmed later by Johnson et al. (2019), before the beginning of this research project there was no previous explanation in the literature for the origins of the low-mass spin bias signal. Sato-Polito et al. (2018) only characterizes the effect, while Johnson et al. (2019) proposes the twin bias mechanism to account for spin bias in the entire mass range, and it fails to account for the low-mass signal.

In Tucci et al. (2021), we explained the physical mechanism behind spin bias for low-mass halos. We showed that the low-mass spin bias inversion (the fact that, at fixed halo mass, low spin halos are more tightly clustered than high spin halos, while the opposite is true at the high-mass end) is completely caused by the population of typically low-mass *splashback halos*.

### Splashback Halos

As we already saw in Section 4.2, splashback halos are distinct halos that were subhalos at some previous time and live in the so-called splashback radius. Since they live in the vicinity of their previous, massive hosts, they have a larger bias than other halos of similar mass. As we discussed in Section 6.5, it is well known that splashback halos are responsible for part of the low mass secondary bias seen in concentration and age.

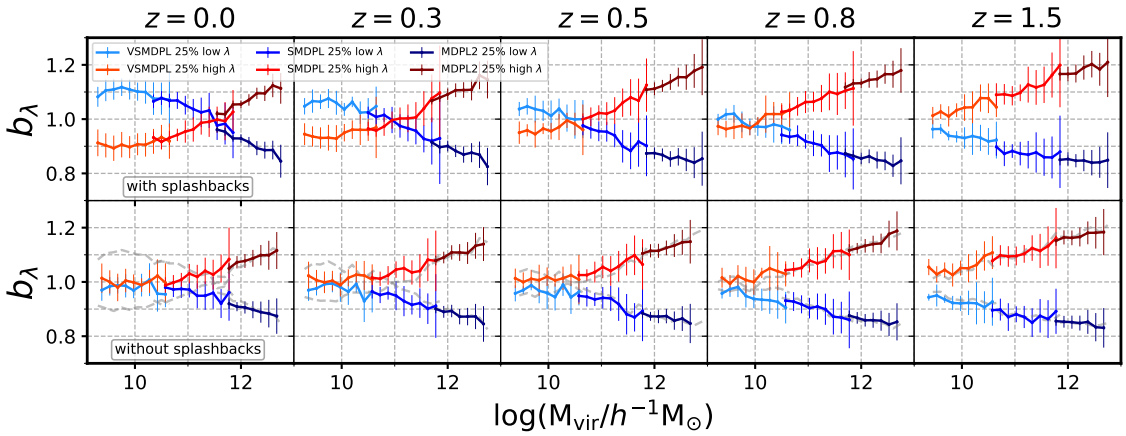


Figure 7.2: Relative spin bias measured in different MultiDark boxes and redshifts. The upper panel displays the results with all the halos, while the lower one shows the measurement after the removal of splashback halos. Dashed lines indicate the results with splashbacks.

In MultiDark, we identified splashback halos with the *redshift of last accretion* $z_{\rm firstacc}$ given by `ROCKSTAR`, which is higher than the analyzed redshift if the halo has ever passed

through the virial radius of a larger halo ($z_{\mathrm{firstacc}} > z$). The halos presenting this feature and which are also distinct halos (`pid`=1), i.e., those which do not lie inside the virial radius of a larger halo in the analyzed redshift, are thus identified as splashback halos. Even composing only a small percentage of the total population (see Figure 7.3), when this subpopulation is removed and the spin bias is analyzed, the crossover at low masses disappears and there is no significant spin bias signal (Figure 7.2).

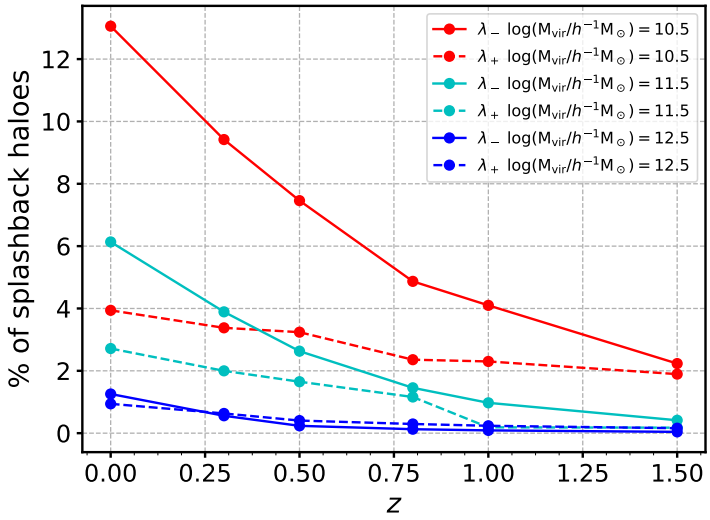


Figure 7.3: At $z = 0.0$, percentage of splashback halos in each mass bin (all halos); inside each mass bin, percentage of splashback halos in the low spin quartile ($\lambda_-$: 25% lower spin halos) and in the high spin quartile ($\lambda_+$: 25% higher spin halos). (Tucci et al., 2021)

Indeed, the spin bias at low masses without the subpopulation of splashback halos is statistically null for the entire low mass range. In order to prove that they are responsible for the spin bias signal, we have to understand how they are connected with halo spin and the large scale environment. As a first step, we characterized the subpopulation of splashback halos. The following facts contribute for the idea that splashback halos are indeed responsible for the low mass spin bias.

As Figure 7.2 displays, spin bias at low masses is a recent phenomenon which becomes important only at low masses, low redshifts and specific of low spin halos. Figure 7.3 shows that splashback halos are only a considerable part of the population of halos at low masses, low redshifts and they are significantly more present in the population of low spin halos.

Since splashbacks naturally have a higher bias than other halos of same mass, it is the unbalance between the fraction of splashbacks present in the spin quartiles that causes the low-mass inversion. The larger fraction of splashbacks in the low-spin quartile causes it to have a larger bias than the high-spin one, and when this specific population of halos

is removed, there is no more difference of clustering between the spin quartiles. The physical mechanism responsible for the low spin of splashbacks will be discussed in the next subsections.

Our results unveil a specific link between spin bias and assembly bias, since splashback halos have been shown to account for some part of the assembly bias signal at the low-mass end. However, their impact on spin bias is far more severe, since the signal completely vanishes after their removal (see Figure 7.4).

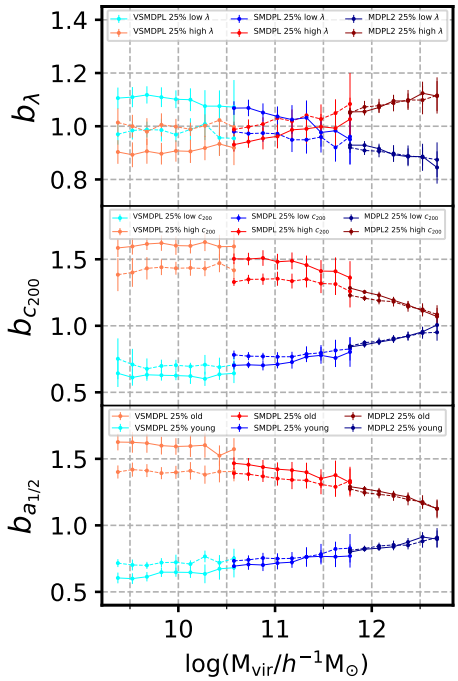


Figure 7.4: The effect of splashback halos on spin bias and assembly bias (i.e., the secondary bias for halo concentration and age) at $z = 0$ in VSMDPL, SMDPL, and MDPL2. Solid (dashed) lines display the relative bias measured by including (removing) splashback halos. (Tucci et al., 2021)

## Consistency checks

We can see in Figure 7.5 that the fact that the spin bias signal disappears without splashback halos is not an artifact related to the amount of halos which are taken out of the sample. If we randomly remove the same amount of halos in each mass bin, the trend remains roughly the same as the spin bias measure with the entire population.

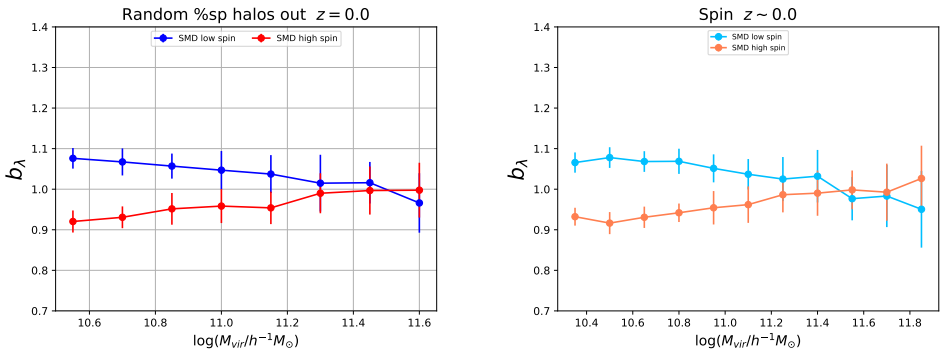


Figure 7.5: Relative spin bias measured in SMDPL at $z = 0$. *Left*: The same amount of splashback halos in each mass bin is randomly removed. *Right*: Reference plot with all the halos.

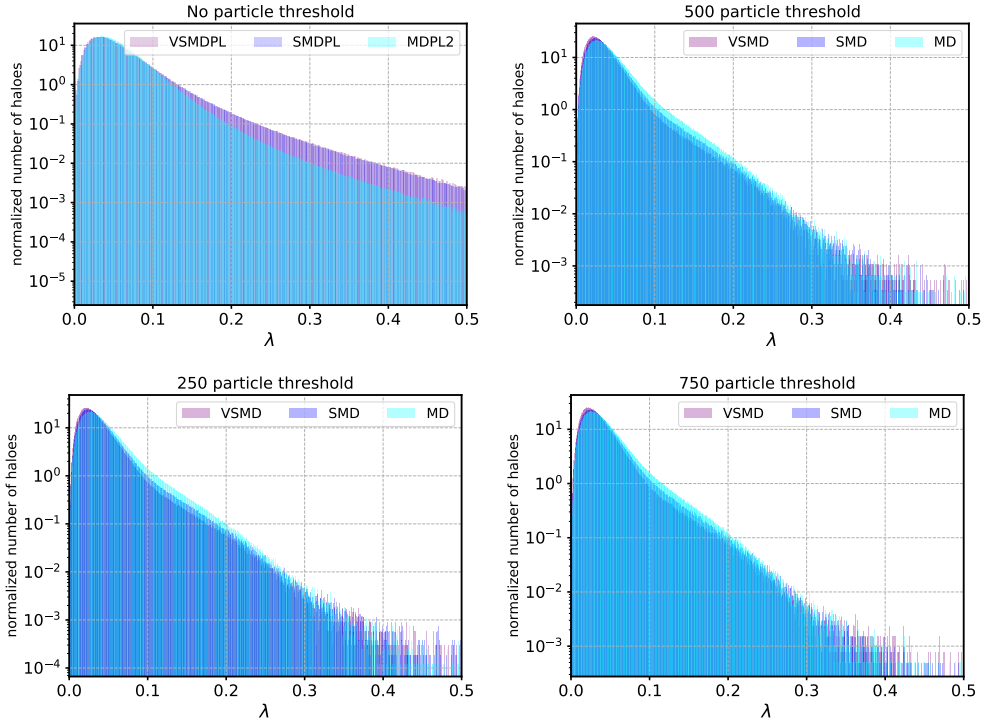


Figure 7.6: Spin distributions over the MultiDark boxes used. *Upper left:* For halos of all masses (no threshold). *Upper right:* 500 particle threshold used in the paper. *Bottom left:* 250 particle threshold for comparison with the 500 particle threshold employed in the analysis. *Bottom right:* 750 particle threshold also for comparison.

We have tested the consistency of spin distributions at a given mass across different boxes/particle resolutions. The normalized histograms for VSMDPL, SMDPL, and MDPL2 are shown in Figure 7.6 for different mass cut resolutions. In this work, we use the 500 particle threshold, for which the spin distribution is consistent across the different boxes. Also, the distributions are consistent for different particle resolutions (250 and 750 particle

thresholds). We show in Figure 7.7 the mean and standard deviation of spin as a function of mass. The 500 particle threshold guarantee a consistent spin distribution across the mass range analysed in this work.

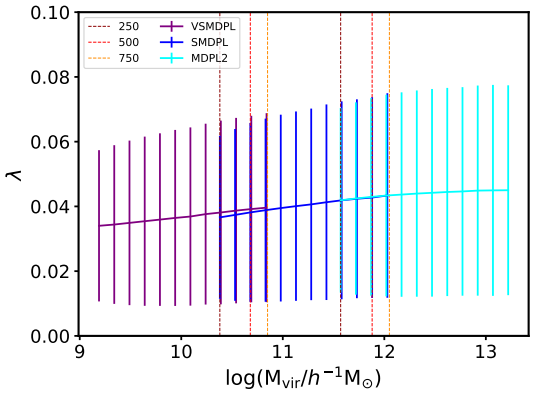


Figure 7.7: Mean and standard deviation of spin in the MultiDark boxes used. Dashed lines show different particle resolution thresholds.

We have tested defining the spin quartiles and then removing splashback halos, instead of removing the splashbacks before separating the population in quartiles. The results are displayed in Figure 7.8 and we notice that they are very similar when defining the quartiles before or after the removal of splashback halos. The high spin quartile only "absorbs" $\sim 5\%$ of the lower spin halos.

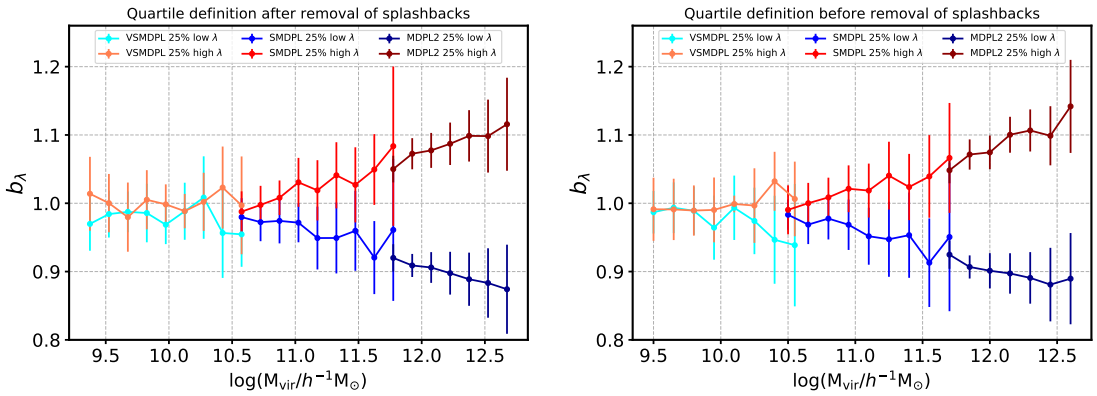


Figure 7.8: Spin bias at $z = 0$. *Left:* The low and high spin quartiles are defined *after* all splashback halos being removed from the sample. *Right:* Spin quartiles are defined *before* the removal of all splashback halos.

We have further investigated the effect of splashback halos in Uchuu (see Section 4.1). Since the halo catalogs are also produced with ROCKSTAR, the splashbacks are also identified as distinct halos with $z_{\rm firstacc} > z$. As Figure 7.9 shows, in Uchuu the inversion at low

masses also vanishes after the removal of splashbacks. We can also see the splashback effect as a function of scale. As in the case of assembly bias (Sunayama et al., 2016), their effect is larger on small scales.

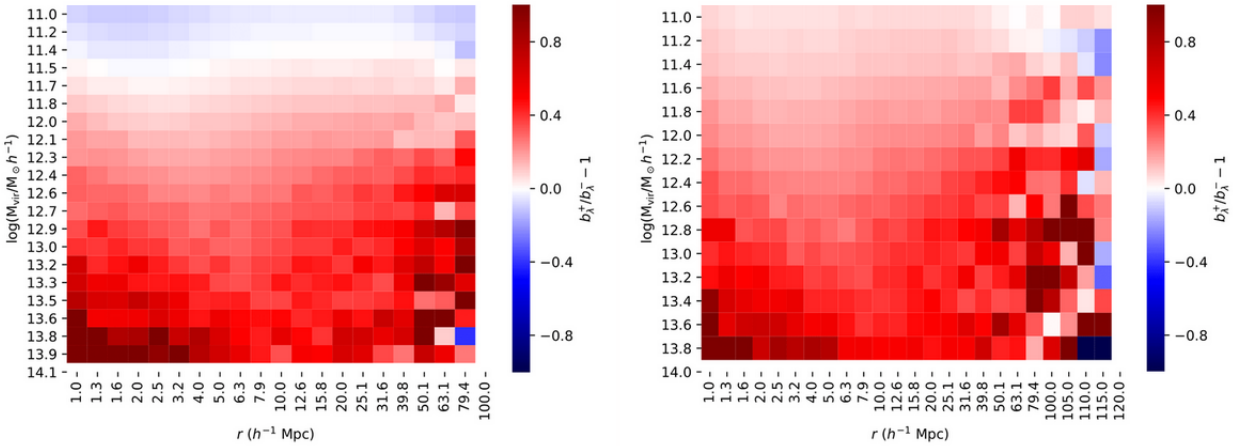


Figure 7.9: Ratio between the relative bias of the 25% highest and lowest spin subpopulations minus one in each mass and radial bin calculated in Uchuu at $z = 0$. *Left:* with splashback halos, the inversion is present at low masses. *Right:* without splashback halos, the inversion disappears.

## Major Mergers

What remains to be understood is the reason why splashbacks have a lower spin than other halos of the same mass. In principle, this fact could be explained by the picture presented in Vitvitska et al. (2002) (see Section 5.3), where the low spin of halos is caused by periods of slow mass accretion. Therefore, in principle, splashback halos could have a lower spin due to a lack of recent major mergers. However, we see in Figure 7.10 that they do not differ from the mean population in the fraction of recent major mergers, suggesting that this may not be the origin of their lower spin when compared to other halos of same mass.

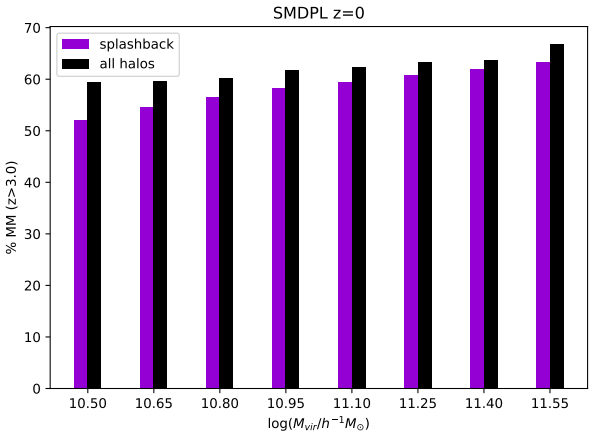


Figure 7.10: Percentage of non-splashback and splashback halos which have undergone a recent major merger ($z < 3$) in each mass bin in SMDPL at $z = 0$.

## Tidal Stripping

A very important feature of splashback halos is that have a very particular mass assembly history, the reason why they are one of the causes of low-mass assembly bias (see Section 6.5). *Tidal stripping* is a dynamical process which takes place due to tidal interactions with a more massive halo and results in mass and spin loss (Lee et al., 2018). Of course, subhalos and splashback halos are expected to suffer from intense tidal stripping with their (previous) host.

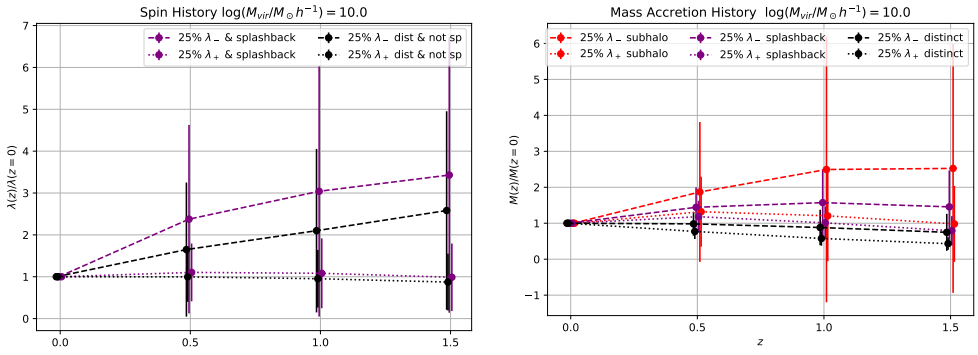


Figure 7.11: *Left:* Spin history for splashback and not splashback halos of $\log(M_{\rm vir}/h^{-1}M_\odot) = 10.0$ at VSMDPL and $z = 0$. *Right:* Mass history for splashback, not splashback halos and subhalos of $\log(M_{\rm vir}/h^{-1}M_\odot) = 10.0$ at VSMDPL and $z = 0$.

We see in Figure 7.11 that low mass, low spin halos at $z = 0$ had their spin and mass decreased with time, what suggests that there is a dynamical process responsible for it. Also, splashback halos of a lower spin at $z = 0$ lost more spin, while the spin of high spin halos didn't evolve with time, indicating that this dynamical mechanism lowers their spin.

We also see in this figure that splashback halos (and subhalos) loose mass with time, the effect behind more pronounced for the low spin ones. All these facts contribute to the tidal stripping mechanism, which is a dynamical processes that has as consequences mass and spin loss. It is more intense in high dense regions and affects at most low-mass halos. Therefore splashback halos are be more subjected by it than other halos of similar mass due to intense tidal forces from the previous host halo, not only during the subhalo epoch but also when still orbiting the previous host.

In MultiDark we have access to the tidal force felt by each halo, defined as the ratio $R_{vir}/R_{Hill}$, where $R_{Hill}$ is the "Hill radius", which can be expressed as

$$R_{Hill} \simeq d \left(\frac{m}{3M}\right)^{1/3} \tag{7.1}$$

and represents the *sphere of influence* of a halo. In essence, in the restricted three-body problem, consisting of a body of mass $M$ at a distance $d$ from a smaller body of mass $m \ll M$, a third body of negligible mass can have stable circular orbits around the smaller mass $m$ only within the Hill radius. However, halos in general do not follow circular orbits around each other and it only considers the strongest tidal force from any nearby halo. For splashbacks and subhalos their (previous) host halo is in general the most tidally influential nearby halo. A halo is prone to be subject to significant tidal stripping when the tidal force is typically greater than one (Hahn et al., 2009).

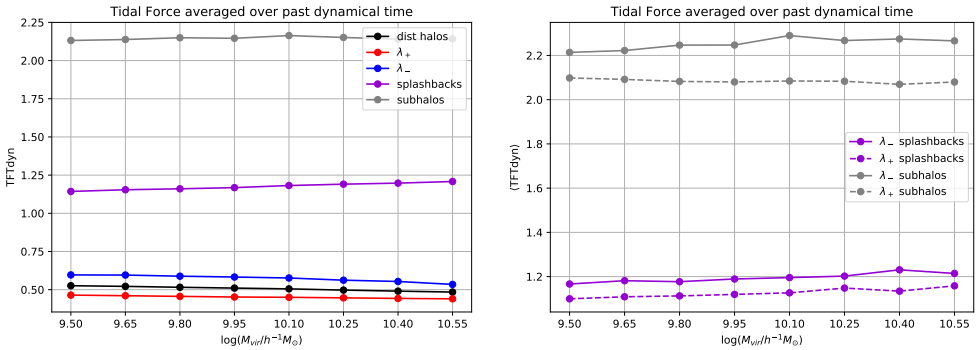


Figure 7.12: *Left:* Tidal force averaged over past dynamical time at VSMDPL and $z = 0$. *Right:* Same but dividing splashback halos and subhalos into spin quartiles.

We see in Figure 7.12 that splashback halos (and subhalos) have suffered more tidal forces during their past dynamical time than not splashback halos, and also that the low spin subpopulations felt even higher ones, what also contributes to the tidal stripping picture. Subhalos feel the most intense tidal forces due to the distribution of dark matter particles from the host halo around it. Splashbacks then were subject to these very intense tidal forces during their subhalo epoch, and most of them are still subject to tidal forces

from their previous host when living near it (see the diagram in Figure 7.13).

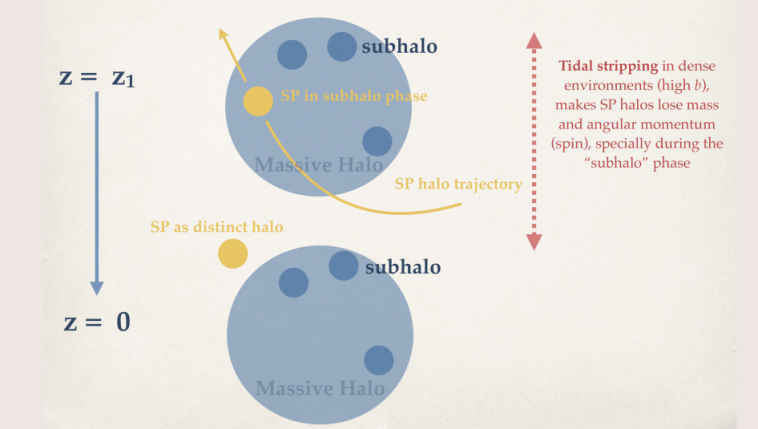


Figure 7.13: Diagram of how splashback halos lose spin due to tidal stripping during the subhalo epoch. Made by Antonio Montero-Dorta.

We conclude here then that splashback halos account for the low-mass spin bias inversion. They have a higher bias due to the proximity to a massive neighbour and a lower spin due to the tidal stripping mechanism (Tucci et al., 2021). After elucidating the inversion and clarifying what seemed to be a complex scenario, our task is now to develop an explanation for the origins of the *intrinsic spin bias*, i.e., the trend of a higher bias for high-spin halos that remains after the removal of splashbacks at high-masses.

## 7.2 What are the origins of High-Mass Spin Bias?

### Some clues

Our measurements displayed here so far, such as in Figure 7.1, do not concern the dependence of the strict bias parameters on spin. However, if we are going to develop a model for spin bias, we should be able to reproduce the dependence of the strict bias parameters on mass and spin, e.g., $b_1(m, \lambda)$, and then test it against simulations. In Lazeyras et al. (2021), it has been measured with forward model techniques how the bias parameters $b_1$, $b_2$ and $b_{K_2}$ vary with respect to the spin parameter, as we can see in Figure 7.14. For a fixed mass, $b_1$ seems to have a linear dependence on spin, while $b_2$ has a quadratic one. It is very hard to interpret the $b_{K_2}$ dependence, but we will try to give a simple reasoning here.

It is well known that the relation $b_{K_2}(b_1)$ is monotonically decreasing and has negative values for massive halos (Lazeyras and Schmidt, 2018), as the more massive the halo (those with higher $b_1$), the more unlikely it is to be found in a high-$K_2$ region. This is because a higher shear field tends to slow down the formation of halos.

According to TTT, halos with a higher spin are more likely to be formed in regions with a higher initial shear (and hence $K_2$). Actually, as we know halo spin is proportional to the misalignment between the shape and tidal tensors, but it is also proportional to the shear magnitude. For the highest mass bin it is much more unlikely for a halo of low-spin to be found in a high-$K_2$ region, which seems to be consistent with the TTT picture.

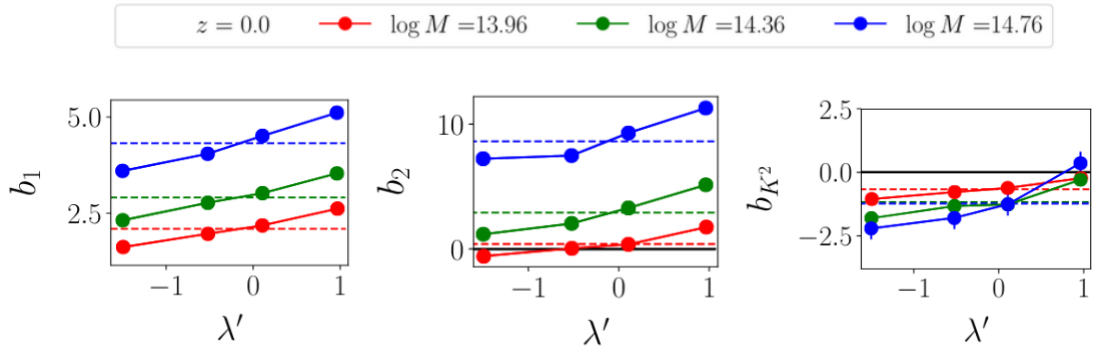


Figure 7.14: Large-scale bias parameters $b_1$, $b_2$ and $b_{K^2}$ as a function of halo spin for three different mass bins (Lazeyras et al., 2021).

## Our Hypothesis

As we saw in Chapter 3, in the ESP scenario we have the variables the density $\delta$ and the curvature $x$, where $\delta$ is constrained by the barrier $B$. A first approach would be to relate spin to curvature, which is known to be related to the large-scale density and thus to bias. This will be our "curvature" hypothesis. Another possibility is to relate the barrier to spin, and then derive a dependence of bias on spin. We will denote this hypothesis as the "barrier" one.

Can we use TTT to better understand spin bias? In Figure 7.2, we can note that high-mass spin bias is an effect which already manifests itself at $z = 1.5$ ($a = 0.4$), where TTT is still a good prediction for the evolution of the angular momentum of massive halos, as we can see in Figure 5.3. Therefore, in both hypothesis we will use the TTT framework to predict halo angular momentum.

We will be trying to investigate whether the bias dependence on spin at the high-mass end come from the relation between curvature and spin or from the relation between the barrier and spin during the next sections (and months, as we have not reached a conclusion yet).

## Curvature

Inspired by Dalal et al. (2008), where high-mass assembly bias was explained by the peak curvature of halos (see Section 6.4), which is related to the mass accretion rate and thus provides a connection between halo concentration and bias. Here, we can try to use the same line of reasoning, since TTT in the peak formalism (see Section 5.2) gives us the relation between halo angular momentum and peak curvature,

$$\ell \equiv \frac{96\pi}{\sqrt{15^3}}(1-\gamma^2)^{1/2}\left(\frac{\nu}{\gamma\, x}\right)^{5/2}\frac{\mathcal{A}(e,p)^{1/2}}{\mathcal{B}(e,p)^{3/2}}\,, \tag{7.2}$$

which we display here again for simplicity. As we can see, angular momentum is inversely proportional to the curvature $x$, which reflects the "sharpness" of the peak: the higher is $x$, the sharper is the peak (see Section 2.3).

The idea is then similar to assembly bias, where halos of a higher spin are formed by shallower peaks, and thus inherit a larger bias than other halos of same mass. In this sense, both trends (at least at high-masses) would arise due to the same mechanism, which is the relation to different peak curvatures of the initial density field.

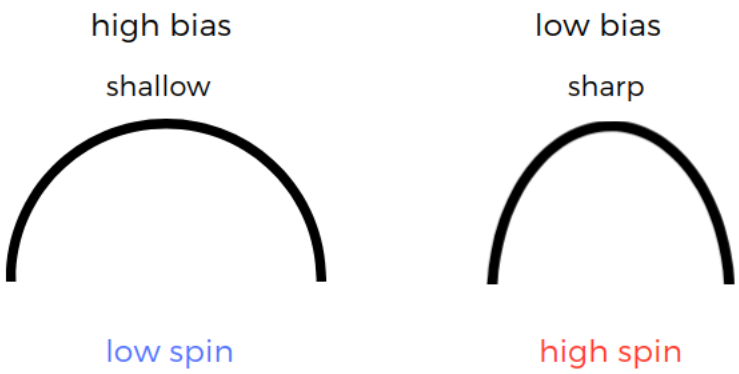


Figure 7.15: Halos of same height formed by shallower peaks (thus inheriting a higher bias) are associated with a high angular momentum.

## Barrier

In the ESP$\tau$ model proposed by Castorina et al. (2016), the barrier for gravitational collapse of halos is modelled to depend on $\tau$, which in turn depends on the misalignment between the inertia and shear tensors (see Section 3.6). Halo spin is also known to depend on such misalignment, as predicted by TTT (see Section 5.1). Therefore, in principle, we could relate halo spin to $\tau$. In Castorina et al. (2016), they integrate the halo mass function obtained by the barrier over $\tau$ in order to provide a prediction for the number of

halos of a given mass $m$. Here, we could refrain from performing the integral and derive a mass function which depends both on halo mass and spin, and with that provide a prediction for spin bias in the form of $b(m, \lambda)$.

The interpretation here would be that halos of a higher spin form in regions where the shear is higher. As shear tends to slow down the formation of halos, these have to struggle more to form and thus to be in a higher density region, what in turn is the reason why they have a higher bias.

### Tests in simulations

We can test our two hypothesis by first dividing the halos into high-spin and low-spin quartiles at $z = 0$. Then, we take the ID of the dark-matter particles contained in each halo and trace them back to the initial conditions by performing a matching. The value of a certain field $f$ associated to a halo with $N_k$ particles of position $\boldsymbol{q}_i \in L_K$ (Lagrangian space) is defined as

$$\langle f \rangle = \frac{1}{N_k} \sum_{\boldsymbol{q}_i \in L_K} f(\boldsymbol{q}) \,, \tag{7.3}$$

where then we can calculate interesting quantities such as the initial smoothed density and peak curvature for a given protohalo. The value of $f(\boldsymbol{q})$ can be determined by making a density grid, calculating the field in Fourier space (where derivatives are more easily calculated) and Fourier transforming back to the grid (see Section 2.3).

Some interesting, useful tests we are doing are the relation between the initial peak curvature and the misalignment between the initial shape and shear tensors (i.e., the prediction given by TTT for halo angular momentum, whose validity is also going to be tested), the relation between $\tau$ and angular momentum, and how the scatter of the initial smoothed density at fixed halo mass correlates with $\tau$ and the final spin.

## 7.3 Observing Spin Bias

There is a great debate on the literature about finding conclusive evidences for secondary bias in real data (Montero-Dorta et al., 2017). While most the attempts have focused on observational signatures of assembly bias, in a recent publication Montero-Dorta et al. (2021) we have proposed a new way of observing halo spin bias. We have analyzed in IllustrisTNG300 how the kinetic Sunyaev-Zel'dovich effect (kSZ) can be used as a proxy for halo spin, while the thermal Sunyaev-Zel'dovich effect (tSZ) traces halo mass.

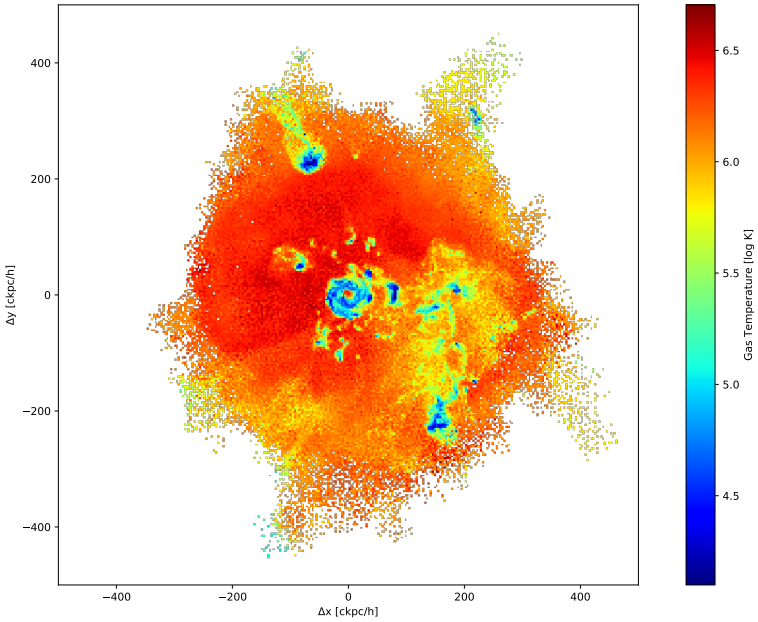


Figure 7.16: Gas temperature of a single halo of $\log(M_{\rm vir}/h^{-1}M_{\odot}) = 12.83$ in Illustris TNG100. The coordinates are with respect to the halo center.

The Sunyaev-Zel'dovich effects (Sunyaev and Zeldovich, 1970, 1980b,a) happen due to the inverse Tompson scattering of photons when interact with the electrons of galaxy clusters, causing anisotropies in the CMB fluctuations. In the tSZ, the scattering takes place due to the thermal, random motion of the electrons inside the cluster, such that the temperature of the photons will depend on the number density of electrons along the line of sight, which is proportional to the total gas mass in the halo and as a consequence to halo mass as well.

The kSZ in turn reflects the rotation of the gas, since it is sensitive to the dot product between the velocity of electrons and the line of sight. As a consequence, the effect is maximum when the rotational axis of the halo is perpendicular to the line of sight, where in one hemisphere will have most of the electrons "entering the plane" and in the other they will be "exiting the plane". This produces a dipole pattern in the gas temperature due to its coherent rotation inside the cluster (see Figure 7.16), which is known to correlate with the angular momentum of the host halo and thus represents a proxy for halo spin. Of course, there is also a huge amount of turbulent motion of the electrons, and the rotation axis of the halo will not always be perpendicular to the line of sight. However, the gas rotation will usually be expected to generate a dipole signal in its temperature.

Given that halo spin bias is the relative clustering of halos selected according to spin at fixed halo mass, we compare this signal with the relative clustering of halos selected according to kSZ at fixed tSZ, as shown in Figure 7.17. Since we recover a very similar trend, this represents an alternative route of measuring secondary bias in real data.

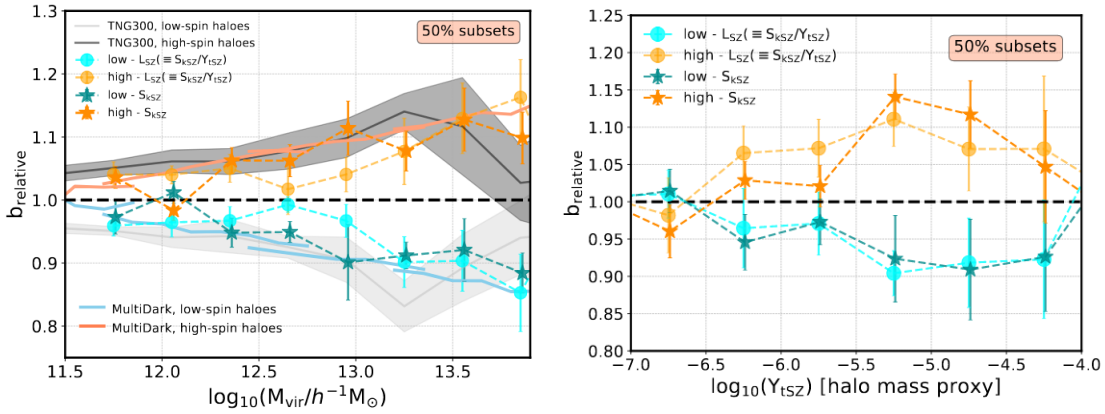


Figure 7.17: (Left): Relative bias of subsets of kSZ (tracing halo spin) as a function of halo mass. We see that the signal is very similar to the halo spin bias measured in IllustrisTNG300 and MultiDark. (Right): Same but as a function of tSZ (tracing halo mass). Montero-Dorta et al. (2021)

# Bibliography

I. Achitouv, Y. Rasera, R. K. Sheth, and P. S. Corasaniti. Testing the Self-Consistency of the Excursion Set Approach to Predicting the Dark Matter Halo Mass Function. *PhRvL*, 111(23):231303, December 2013. doi: 10.1103/PhysRevLett.111.231303.

Susmita Adhikari, Neal Dalal, and Robert T. Chamberlain. Splashback in accreting dark matter halos. *JCAP*, 2014(11):019, November 2014. doi: 10.1088/1475-7516/2014/11/019.

Susmita Adhikari, Jeremy Sakstein, Bhuvnesh Jain, Neal Dalal, and Baojiu Li. Splashback in galaxy clusters as a probe of cosmic expansion and gravity. *JCAP*, 2018(11):033, November 2018. doi: 10.1088/1475-7516/2018/11/033.

Luca Amendola, Stephen Appleby, David Bacon, Tessa Baker, Marco Baldi, Nicola Bartolo, Alain Blanchard, Camille Bonvin, Stefano Borgani, Enzo Branchini, Clare Burrage, Stefano Camera, Carmelita Carbone, Luciano Casarini, Mark Cropper, Claudia de Rham, Cinzia Di Porto, Anne Ealet, Pedro G. Ferreira, Fabio Finelli, Juan García-Bellido, Tommaso Giannantonio, Luigi Guzzo, Alan Heavens, Lavinia Heisenberg, Catherine Heymans, Henk Hoekstra, Lukas Hollenstein, Rory Holmes, Ole Horst, Knud Jahnke, Thomas D. Kitching, Tomi Koivisto, Martin Kunz, Giuseppe La Vacca, Marisa March, Elisabetta Majerotto, Katarina Markovic, David Marsh, Federico Marulli, Richard Massey, Yannick Mellier, David F. Mota, Nelson J. Nunes, Will Percival, Valeria Pettorino, Cristiano Porciani, Claudia Quercellini, Justin Read, Massimiliano Rinaldi, Domenico Sapone, Roberto Scaramella, Constantinos Skordis, Fergus Simpson, Andy Taylor, Shaun Thomas, Roberto Trotta, Licia Verde, Filippo Vernizzi, Adrian Vollmer, Yun Wang, Jochen Weller, and Tom Zlosnik. Cosmology and Fundamental Physics with the Euclid Satellite. *Living Reviews in Relativity*, 16(1):6, September 2013. doi: 10.12942/lrr-2013-6.

Valentin Assassi, Daniel Baumann, Daniel Green, and Matias Zaldarriaga. Renormalized halo bias. *JCAP*, 2014(8):056, August 2014. doi: 10.1088/1475-7516/2014/08/056.

J. M. Bardeen, J. R. Bond, N. Kaiser, and A. S. Szalay. The Statistics of Peaks of Gaussian Random Fields. *ApJ*, 304:15, May 1986. doi: 10.1086/164143.

Alexandre Barreira and Fabian Schmidt. Responses in large-scale structure. *JCAP*, 2017 (6):053, June 2017. doi: 10.1088/1475-7516/2017/06/053.

Daniel Baumann. Inflation. In *Theoretical Advanced Study Institute in Elementary Particle Physics: Physics of the Large and the Small*, 7 2009. doi: 10.1142/9789814327183_0010.

Peter S. Behroozi, Risa H. Wechsler, and Hao-Yi Wu. THE ROCKSTAR PHASE-SPACE TEMPORAL HALO FINDER AND THE VELOCITY OFFSETS OF CLUSTER CORES. *ApJ*, 762(2):109, December 2012. doi: 10.1088/0004-637x/762/2/109. URL https://doi.org/10.1088/0004-637x/762/2/109.

F. Bernardeau, S. Colombi, E. Gaztañaga, and R. Scoccimarro. Large-scale structure of the Universe and cosmological perturbation theory. *PhysRep*, 367(1-3):1–248, September 2002. doi: 10.1016/S0370-1573(02)00135-7.

Philip Bett, Vincent Eke, Carlos S. Frenk, Adrian Jenkins, John Helly, and Julio Navarro. The spin and shape of dark matter haloes in the Millennium simulation of a Λ cold dark matter universe. *MNRAS*, 376(1):215–232, March 2007. doi: 10.1111/j.1365-2966.2007.11432.x.

J. R. Bond, S. Cole, G. Efstathiou, and N. Kaiser. Excursion Set Mass Functions for Hierarchical Gaussian Fluctuations. *ApJ*, 379:440, October 1991. doi: 10.1086/170520.

Michael Boylan-Kolchin, Volker Springel, Simon D. M. White, Adrian Jenkins, and Gerard Lemson. Resolving cosmic structure formation with the Millennium-II Simulation. *Monthly Notices of the Royal Astronomical Society*, 398(3):1150–1164, 09 2009. ISSN 0035-8711. doi: 10.1111/j.1365-2966.2009.15191.x. URL https://doi.org/10.1111/j.1365-2966.2009.15191.x.

Greg L. Bryan and Michael L. Norman. Statistical Properties of X-Ray Clusters: Analytic and Numerical Comparisons. *ApJ*, 495:80–99, Mar 1998. doi: 10.1086/305262.

J. S. Bullock, T. S. Kolatt, Y. Sigad, R. S. Somerville, A. V. Kravtsov, A. A. Klypin, J. R. Primack, and A. Dekel. Profiles of dark haloes: evolution, scatter and environment. *MNRAS*, 321:559–575, Mar 2001. doi: 10.1046/j.1365-8711.2001.04068.x.

Matteo Carrera and Domenico Giulini. Influence of global cosmological expansion on local dynamics and kinematics. *Reviews of Modern Physics*, 82(1):169–208, January 2010. doi: 10.1103/RevModPhys.82.169.

Emanuele Castorina, Emiliano Sefusatti, Ravi K. Sheth, Francisco Villaescusa-Navarro, and Matteo Viel. Cosmology with massive neutrinos II: on the universality of the halo mass function and bias. *JCAP*, 2014(2):049, February 2014. doi: 10.1088/1475-7516/2014/02/049.

Emanuele Castorina, Aseem Paranjape, Oliver Hahn, and Ravi K. Sheth. Excursion set peaks: the role of shear. *arXiv e-prints*, art. arXiv:1611.03619, November 2016.

Paolo Catelan and Tom Theuns. Evolution of the angular momentum of protogalaxies from tidal torques: Zel'dovich approximation. *MNRAS*, 282(2):436–454, September 1996. doi: 10.1093/mnras/282.2.436.

Chun Yin Ricky Chue, Neal Dalal, and Martin White. Some assembly required: assembly bias in massive dark matter halos. *JCAP*, 2018(10):012, October 2018. doi: 10.1088/1475-7516/2018/10/012.

Alison L. Coil. *The Large-Scale Structure of the Universe*, volume 6, page 387. 2013. doi: 10.1007/978-94-007-5609-0_8.

Shaun Cole and Nick Kaiser. Biased clustering in the cold dark matter cosmogony. *MNRAS*, 237:1127–1146, April 1989. doi: 10.1093/mnras/237.4.1127.

P. Coles. Large-scale Structure, Theory and Statistics. In Hector J. de Vega, Isaak M. Khalatnikov, and Norma G. Sanchez, editors, *Phase Transitions in the Early Universe: Theory and Observations*, volume 40, page 217, January 2001.

Asantha Cooray and Ravi Sheth. Halo models of large scale structure. *PhysRep*, 372(1): 1–129, December 2002. doi: 10.1016/S0370-1573(02)00276-4.

N. Dalal, O. Doré, D. Huterer, and A. Shirokov. Imprints of primordial non-Gaussianities on large-scale structure: Scale-dependent bias and abundance of virialized objects. *Physical Review D*, 77(12):123514, June 2008. doi: 10.1103/PhysRevD.77.123514.

Neal Dalal, Martin White, J. Richard Bond, and Alexander Shirokov. Halo assembly bias in hierarchical structure formation. *The Astrophysical Journal*, 687(1):12–21, 2008. ISSN 1538-4357. doi: 10.1086/591512. URL http://dx.doi.org/10.1086/591512.

Neal Dalal, Martin White, J. Richard Bond, and Alexander Shirokov. Halo Assembly Bias in Hierarchical Structure Formation. *ApJ*, 687(1):12–21, November 2008. doi: 10.1086/591512.

M. Davis and P. J. E. Peebles. A survey of galaxy redshifts. V - The two-point position and velocity correlations. *ApJ*, 267:465–482, April 1983. doi: 10.1086/160884.

A. Dekel and J. Silk. The Origin of Dwarf Galaxies, Cold Dark Matter, and Biased Galaxy Formation. *ApJ*, 303:39, April 1986. doi: 10.1086/164050.

A. Del Popolo and M. Gambera. Tidal torques and the clusters of galaxies evolution. *A&A*, 337:96–104, September 1998.

DESI Collaboration et al. The DESI Experiment Part I: Science,Targeting, and Survey Design. *arXiv e-prints*, art. arXiv:1611.00036, October 2016.

Vincent Desjacques, Donghui Jeong, and Fabian Schmidt. Large-scale galaxy bias. *PhysRep*, 733:1–193, February 2018. doi: 10.1016/j.physrep.2017.12.002.

Benedikt Diemer, Philip Mansfield, Andrey V. Kravtsov, and Surhud More. The Splashback Radius of Halos from Particle Dynamics. II. Dependence on Mass, Accretion Rate, Redshift, and Cosmology. *ApJ*, 843(2):140, July 2017. doi: 10.3847/1538-4357/aa79ab.

S. Dodelson and F. Schmidt. *Modern Cosmology*. Elsevier Science, 2020. ISBN 9780128159484. URL https://books.google.com.br/books?id=GGjfywEACAAJ.

Scott Dodelson. *Modern Cosmology*. Academic Press, 2003.

A.G. Doroshkevich. Spatial structure of perturbations and origin of galactic rotation in fluctuation theory. *Astrophysics*, 6(4):320–330, October 1970. doi: 10.1007/BF01001625.

G. Efstathiou, M. Davis, S. D. M. White, and C. S. Frenk. Numerical techniques for large cosmological N-body simulations. *ApJS*, 57:241–260, February 1985. doi: 10.1086/191003.

Albert Einstein and Ernst G. Straus. The Influence of the Expansion of Space on the Gravitation Fields Surrounding the Individual Stars. *Reviews of Modern Physics*, 17 (2-3):120–124, April 1945. doi: 10.1103/RevModPhys.17.120.

Andreas Faltenbacher and Simon D. M. White. Assembly Bias and the Dynamical Structure of Dark Matter Halos. *ApJ*, 708(1):469–473, January 2010. doi: 10.1088/0004-637X/708/1/469.

Carlos S. Frenk, Simon D. M. White, Marc Davis, and George Efstathiou. The Formation of Dark Halos in a Universe Dominated by Cold Dark Matter. *ApJ*, 327:507, April 1988. doi: 10.1086/166213.

L. Gao and S. D. M. White. Assembly bias in the clustering of dark matter haloes. *MNRAS*, 377:L5–L9, April 2007. doi: 10.1111/j.1745-3933.2007.00292.x.

L. Gao, S. D. M. White, A. Jenkins, F. Stoehr, and V. Springel. The subhalo populations of ΛCDM dark haloes. *MNRAS*, 355:819–834, December 2004. doi: 10.1111/j.1365-2966.2004.08360.x.

L. Gao, V. Springel, and S. D. M. White. The age dependence of halo clustering. *MNRAS: Letters*, 363(1):L66–L70, Oct 2005. ISSN 1745-3933. doi: 10.1111/j.1745-3933.2005.00084.x. URL http://dx.doi.org/10.1111/j.1745-3933.2005.00084.x.

James E. Gunn and III Gott, J. Richard. On the Infall of Matter Into Clusters of Galaxies and Some Effects on Their Evolution. *ApJ*, 176:1, August 1972. doi: 10.1086/151605.

Oliver Hahn and Aseem Paranjape. The locations of halo formation and the peaks formalism. *MNRAS*, 438(1):878–899, February 2014. doi: 10.1093/mnras/stt2256.

Oliver Hahn, Cristiano Porciani, Avishai Dekel, and C. Marcella Carollo. Tidal effects and the environment dependence of halo assembly. *MNRAS*, 398(4):1742–1756, October 2009. doi: 10.1111/j.1365-2966.2009.15271.x.

Andrew P. Hearin, Andrew R. Zentner, Frank C. van den Bosch, Duncan Campbell, and Erik Tollerud. Introducing decorated HODs: modelling assembly bias in the galaxy-halo connection. *MNRAS*, 460(3):2552–2570, August 2016. doi: 10.1093/mnras/stw840.

Alan Heavens and John Peacock. Tidal torques and local density maxima. *MNRAS*, 232: 339–360, May 1988. doi: 10.1093/mnras/232.2.339.

Y. Hoffman. The Dynamics of Superclusters: The Effect of Shear. *ApJ*, 308:493, September 1986. doi: 10.1086/164520.

Tomoaki Ishiyama, Francisco Prada, Anatoly A. Klypin, Manodeep Sinha, R. Benton Metcalf, Eric Jullo, Bruno Altieri, Sofía A. Cora, Darren Croton, Sylvain de la Torre, David E. Millán-Calero, Taira Oogi, José Ruedas, and Cristian A. Vega-Martínez. The Uchuu simulations: Data Release 1 and dark matter halo concentrations. *MNRAS*, 506 (3):4210–4231, September 2021. doi: 10.1093/mnras/stab1755.

James W. Johnson, Ariyeh H. Maller, Andreas A. Berlind, Manodeep Sinha, and J. Kelly Holley-Bockelmann. The secondary spin bias of dark matter haloes. *MNRAS*, 486(1): 1156–1166, Jun 2019. doi: 10.1093/mnras/stz942.

N. Kaiser. On the spatial correlations of Abell clusters. *ApJ*, 284:L9–L12, September 1984. doi: 10.1086/184341.

Anatoly Klypin, Gustavo Yepes, Stefan Gottlöber, Francisco Prada, and Steffen Heß. MultiDark simulations: the story of dark matter halo concentrations and density profiles. *MNRAS*, 457:4340–4359, Apr 2016. doi: 10.1093/mnras/stw248.

Anatoly A. Klypin, Sebastian Trujillo-Gomez, and Joel Primack. Dark Matter Halos in the Standard Cosmological Model: Results from the Bolshoi Simulation. *ApJ*, 740:102, Oct 2011. doi: 10.1088/0004-637X/740/2/102.

A. et al. Knebe. Haloes gone MAD: The Halo-Finder Comparison Project. *MNRAS*, 415: 2293–2318, August 2011. doi: 10.1111/j.1365-2966.2011.18858.x.

I. Lacerna and N. Padilla. The nature of assembly bias - II. Halo spin. *MNRAS*, 426: L26–L30, October 2012. doi: 10.1111/j.1745-3933.2012.01316.x.

S. D. Landy and A. S. Szalay. Bias and variance of angular correlation functions. *ApJ*, 412:64–71, July 1993.

Titouan Lazeyras and Fabian Schmidt. Beyond LIMD bias: a measurement of the complete set of third-order halo bias parameters. *JCAP*, 2018(9):008, September 2018. doi: 10.1088/1475-7516/2018/09/008.

Titouan Lazeyras and Fabian Schmidt. A robust measurement of the first higher-derivative bias of dark matter halos. *Journal of Cosmology and Astroparticle Physics*, 2019(11): 041–041, Nov 2019. ISSN 1475-7516. doi: 10.1088/1475-7516/2019/11/041. URL http://dx.doi.org/10.1088/1475-7516/2019/11/041.

Titouan Lazeyras, Marcello Musso, and Fabian Schmidt. Large-scale assembly bias of dark matter halos. *JCAP*, 2017(3):059, Mar 2017. doi: 10.1088/1475-7516/2017/03/059.

Titouan Lazeyras, Alexandre Barreira, and Fabian Schmidt. Assembly bias in quadratic bias parameters of dark matter halos from forward modeling. *arXiv e-prints*, art. arXiv:2106.14713, June 2021.

Christoph T. Lee, Joel R. Primack, Peter Behroozi, Aldo Rodríguez-Puebla, Doug Hellinger, and Avishai Dekel. Tidal stripping and post-merger relaxation of dark matter haloes: causes and consequences of mass-loss. *MNRAS*, 481(3):4038–4057, December 2018. doi: 10.1093/mnras/sty2538.

Yun Li, H. J. Mo, and L. Gao. On halo formation times and assembly bias. *MNRAS*, 389 (3):1419–1426, September 2008. doi: 10.1111/j.1365-2966.2008.13667.x.

Pablo López, Manuel E. Merchán, and Dante J. Paz. Deviations from tidal torque theory: environment dependences on halo angular momentum growth. *MNRAS*, 485(4): 5244–5255, June 2019. doi: 10.1093/mnras/stz762.

LSST Science Collaboration et al. LSST Science Book, Version 2.0. *arXiv e-prints*, art. arXiv:0912.0201, December 2009.

Shinji Tsujikawa Luca Amendola. *Dark Energy: Theory and Observations*. Cambridge University Press, 2010.

Michele Maggiore and Antonio Riotto. The Halo Mass Function from Excursion Set Theory. I. Gaussian Fluctuations with Non-Markovian Dependence on the Smoothing Scale. *ApJ*, 711(2):907–927, March 2010a. doi: 10.1088/0004-637X/711/2/907.

Michele Maggiore and Antonio Riotto. The Halo Mass Function from Excursion Set Theory. III. Non-Gaussian Fluctuations. *ApJ*, 717(1):526–541, July 2010b. doi: 10.1088/0004-637X/717/1/526.

Philip Mansfield and Andrey V. Kravtsov. The Three Causes of Low-Mass Assembly Bias. *MNRAS*, February 2020. doi: 10.1093/mnras/staa430.

Ian G. McCarthy, Joop Schaye, Simeon Bird, and Amandine M. C. Le Brun. The BAHAMAS project: calibrated hydrodynamical simulations for large-scale structure cosmology. *MNRAS*, 465(3):2936–2965, March 2017. doi: 10.1093/mnras/stw2792.

Thiago Mergulhão, Henrique Rubira, Rodrigo Voivodic, and L. Raul Abramo. The Effective Field Theory of Large-Scale Structure and Multi-tracer. *arXiv e-prints*, art. arXiv:2108.11363, August 2021.

A. D. Montero-Dorta, E. Pérez, F. Prada, S. Rodríguez-Torres, G. Favole, A. Klypin, R. Cid Fernandes, R. M. González Delgado, A. Domínguez, A. S. Bolton, R. García-Benito, E. Jullo, and A. Niemiec. The Dependence of Galaxy Clustering on Stellar-mass Assembly History for LRGs. *ApJ, Letters*, 848:L2, October 2017. doi: 10.3847/2041-8213/aa8cc5.

Antonio D. Montero-Dorta, M. Celeste Artale, L. Raul Abramo, Beatriz Tucci, Nelson Padilla, Gabriela Sato-Polito, Ivan Lacerna, Facundo Rodriguez, and Raul E. Angulo. The manifestation of secondary bias on the galaxy population from IllustrisTNG300. *MNRAS*, 496(2):1182–1196, June 2020. doi: 10.1093/mnras/staa1624.

Antonio D. Montero-Dorta, M. Celeste Artale, L. Raul Abramo, and Beatriz Tucci. On the kinetic Sunyaev-Zel'dovich effect as an observational probe for halo spin bias. *MNRAS*, 504(3):4568–4582, July 2021. doi: 10.1093/mnras/stab1026.

Azadeh Moradinezhad Dizgah, Kwan Chuen Chan, Jorge Noreña, Matteo Biagetti, and Vincent Desjacques. Squeezing the halo bispectrum: a test of bias models. *JCAP*, 2016 (9):030, September 2016. doi: 10.1088/1475-7516/2016/09/030.

Surhud More, Benedikt Diemer, and Andrey V. Kravtsov. The Splashback Radius as a Physical Halo Boundary and the Growth of Halo Mass. *ApJ*, 810(1):36, September 2015. doi: 10.1088/0004-637X/810/1/36.

Surhud More, Hironao Miyatake, Masahiro Takada, Benedikt Diemer, Andrey V. Kravtsov, Neal K. Dalal, Anupreeta More, Ryoma Murata, Rachel Mandelbaum, Eduardo Rozo, Eli S. Rykoff, Masamune Oguri, and David N. Spergel. Detection of the Splashback Radius and Halo Assembly Bias of Massive Galaxy Clusters. *ApJ*, 825(1):39, July 2016. doi: 10.3847/0004-637X/825/1/39.

V. Mukhanov. *Physical Foundations of Cosmology*. Cambridge University Press, 2005. ISBN 9781139447119. URL https://books.google.com.br/books?id=OqIgAwAAQBAJ.

M. Musso, C. Cadiou, C. Pichon, S. Codis, K. Kraljic, and Y. Dubois. How does the cosmic web impact assembly bias? *MNRAS*, 476(4):4877–4906, June 2018. doi: 10.1093/mnras/sty191.

Marcello Musso and Ravi K. Sheth. One step beyond: the excursion set approach with correlated steps. *MNRAS*, 423(1):L102–L106, June 2012. doi: 10.1111/j.1745-3933.2012.01266.x.

Julio F. Navarro, Carlos S. Frenk, and Simon D. M. White. The Structure of Cold Dark Matter Halos. *ApJ*, 462:563, May 1996. doi: 10.1086/177173.

Aseem Paranjape and Ravi K. Sheth. Peaks theory and the excursion set approach. *MNRAS*, 426(4):2789–2796, November 2012. doi: 10.1111/j.1365-2966.2012.21911.x.

Aseem Paranjape, Tsz Yan Lam, and Ravi K. Sheth. Halo abundances and counts-in-cells: the excursion set approach with correlated steps. *MNRAS*, 420(2):1429–1441, February 2012. doi: 10.1111/j.1365-2966.2011.20128.x.

Aseem Paranjape, Oliver Hahn, and Ravi K. Sheth. Halo assembly bias and the tidal anisotropy of the local halo environment. *MNRAS*, 476(3):3631–3647, May 2018. doi: 10.1093/mnras/sty496.

Kyle Parfrey, Lam Hui, and Ravi K. Sheth. Scale-dependent halo bias from scale-dependent growth. *PhRvD*, 83(6):063511, March 2011. doi: 10.1103/PhysRevD.83.063511.

J. A. Peacock and A. F. Heavens. Alternatives to the Press-Schechter cosmological mass function. *MNRAS*, 243:133–143, March 1990. doi: 10.1093/mnras/243.1.133.

P.J.E. Peebles. Origin of the Angular Momentum of Galaxies. *ApJ*, 155:393, February 1969. doi: 10.1086/149876.

Planck Collaboration et al. Planck 2018 results. VI. Cosmological parameters. *arXiv e-prints*, art. arXiv:1807.06209, Jul 2018.

Cristiano Porciani, Avishai Dekel, and Yehuda Hoffman. Testing tidal-torque theory - I. Spin amplitude and direction. *MNRAS*, 332(2):325–338, May 2002. doi: 10.1046/j.1365-8711.2002.05305.x.

W. H. Press and P. Schechter. Formation of Galaxies and Clusters of Galaxies by Self-Similar Gravitational Condensation. *ApJ*, 187:425–438, February 1974. doi: 10.1086/152650.

Sujatha Ramakrishnan and Aseem Paranjape. Separate Universe calibration of the dependence of halo bias on cosmic web anisotropy. *arXiv e-prints*, art. arXiv:2007.03711, July 2020.

Sujatha Ramakrishnan, Aseem Paranjape, Oliver Hahn, and Ravi K. Sheth. Cosmic web anisotropy is the primary indicator of halo assembly bias. *MNRAS*, 489(3):2977–2996, November 2019. doi: 10.1093/mnras/stz2344.

Brant E. Robertson, Andrey V. Kravtsov, Jeremy Tinker, and Andrew R. Zentner. Collapse Barriers and Halo Abundance: Testing the Excursion Set Ansatz. *ApJ*, 696(1):636–652, May 2009. doi: 10.1088/0004-637X/696/1/636.

Andrés N Salcedo, Ariyeh H Maller, Andreas A Berlind, Manodeep Sinha, Cameron K McBride, Peter S Behroozi, Risa H Wechsler, and David H Weinberg. Spatial clustering of dark matter haloes: secondary bias, neighbour bias, and the influence of massive neighbours on halo properties. *MNRAS*, 475(4):4411–4423, jan 2018. doi: 10.1093/mnras/sty109. URL https://doi.org/10.1093/mnras/sty109.

Gabriela Sato-Polito, Antonio D. Montero-Dorta, L. Raul Abramo, Francisco Prada, and Anatoly Klypin. The dependence of halo bias on age, concentration and spin. *arXiv e-prints*, art. arXiv:1810.02375, Oct 2018.

Joop Schaye, Robert A. Crain, Richard G. Bower, Michelle Furlong, Matthieu Schaller, Tom Theuns, Claudio Dalla Vecchia, Carlos S. Frenk, I. G. McCarthy, John C. Helly, Adrian Jenkins, Y. M. Rosas-Guevara, Simon D. M. White, Maarten Baes, C. M. Booth, Peter Camps, Julio F. Navarro, Yan Qu, Alireza Rahmati, Till Sawala, Peter A. Thomas, and James Trayford. The EAGLE project: simulating the evolution and assembly of galaxies and their environments. *MNRAS*, 446(1):521–554, January 2015. doi: 10.1093/mnras/stu2058.

R. K. Sheth and G. Tormen. Large-scale bias and the peak background split. *MNRAS*, 308:119–126, September 1999. doi: 10.1046/j.1365-8711.1999.02692.x.

Ravi K. Sheth and Giuseppe Tormen. An excursion set model of hierarchical clustering: ellipsoidal collapse and the moving barrier. *MNRAS*, 329(1):61–75, January 2002. doi: 10.1046/j.1365-8711.2002.04950.x.

Ravi K. Sheth, H. J. Mo, and Giuseppe Tormen. Ellipsoidal collapse and an improved model for the number and spatial distribution of dark matter haloes. *MNRAS*, 323(1): 1–12, May 2001. doi: 10.1046/j.1365-8711.2001.04006.x.

Ravi K. Sheth, Kwan Chuen Chan, and Román Scoccimarro. Nonlocal lagrangian bias. *Physical Review D*, 87(8), Apr 2013. ISSN 1550-2368. doi: 10.1103/physrevd.87.083002. URL http://dx.doi.org/10.1103/PhysRevD.87.083002.

Manodeep Sinha. Corrfunc: Corrfunc-1.1.0. http://dx.doi.org/10.5281/zenodo.55161, June 2016.

V. Springel, S. D. M. White, A. Jenkins, C. S. Frenk, N. Yoshida, L. Gao, J. Navarro, R. Thacker, D. Croton, J. Helly, J. A. Peacock, S. Cole, P. Thomas, H. Couchman, A. Evrard, J. Colberg, and F. Pearce. Simulations of the formation, evolution and clustering of galaxies and quasars. *Nature*, 435:629–636, June 2005. doi: 10.1038/nature03597.

Volker Springel, Rüdiger Pakmor, Oliver Zier, and Martin Reinecke. Simulating cosmic structure formation with the GADGET-4 code. *MNRAS*, 506(2):2871–2949, September 2021. doi: 10.1093/mnras/stab1855.

Joachim Gerhard Stadel. *Cosmological N-body simulations and their analysis*. PhD thesis, UNIVERSITY OF WASHINGTON, January 2001.

Tomomi Sunayama, Andrew P. Hearin, Nikhil Padmanabhan, and Alexie Leauthaud. The scale-dependence of halo assembly bias. *MNRAS*, 458(2):1510–1516, May 2016. doi: 10.1093/mnras/stw332.

R. A. Sunyaev and Ia. B. Zeldovich. Microwave background radiation as a probe of the contemporary structure and history of the universe. *ARA&A*, 18:537–560, January 1980a. doi: 10.1146/annurev.aa.18.090180.002541.

R. A. Sunyaev and Ya. B. Zeldovich. The interaction of matter and radiation in the hot model of the Universe, II. *APSS*, 7(1):20–30, April 1970. doi: 10.1007/BF00653472.

R. A. Sunyaev and Ya. B. Zeldovich. The velocity of clusters of galaxies relative to the microwave background - The possibility of its measurement. *MNRAS*, 190:413–420, February 1980b. doi: 10.1093/mnras/190.3.413.

R. Teyssier. Cosmological hydrodynamics with adaptive mesh refinement. A new high resolution code called RAMSES. *A&A*, 385:337–364, April 2002. doi: 10.1051/0004-6361:20011817.

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. *ApJ*, 724(2):878–886, December 2010. doi: 10.1088/0004-637X/724/2/878.

Beatriz Tucci, Antonio D. Montero-Dorta, L. Raul Abramo, Gabriela Sato-Polito, and M. Celeste Artale. The physical origins of low-mass spin bias. *MNRAS*, 500(3):2777–2785, January 2021. doi: 10.1093/mnras/staa3319.

Francisco Villaescusa-Navarro, ChangHoon Hahn, Elena Massara, Arka Banerjee, Ana Maria Delgado, Doogesh Kodi Ramanah, Tom Charnock, Elena Giusarma, Yin Li, Erwan Allys, Antoine Brochard, Cora Uhlemann, Chi-Ting Chiang, Siyu He, Alice Pisani, Andrej Obuljen, Yu Feng, Emanuele Castorina, Gabriella Contardo, Christina D. Kreisch, Andrina Nicola, Justin Alsing, Roman Scoccimarro, Licia Verde, Matteo Viel, Shirley Ho, Stephane Mallat, Benjamin Wandelt, and David N. Spergel. The Quijote Simulations. *ApJS*, 250(1):2, September 2020. doi: 10.3847/1538-4365/ab9d82.

Antonio S. Villarreal, Andrew R. Zentner, Yao-Yuan Mao, Chris W. Purcell, Frank C. van den Bosch, Benedikt Diemer, Johannes U. Lange, Kuan Wang, and Duncan Campbell. The immitigable nature of assembly bias: the impact of halo definition on assembly bias. *MNRAS*, 472(1):1088–1105, November 2017. doi: 10.1093/mnras/stx2045.

Maya Vitvitska, Anatoly A. Klypin, Andrey V. Kravtsov, Risa H. Wechsler, Joel R. Primack, and James S. Bullock. The Origin of Angular Momentum in Dark Matter Halos. *ApJ*, 581(2):799–809, December 2002. doi: 10.1086/344361.

Rodrigo Voivodic and Alexandre Barreira. Responses of Halo Occupation Distributions: a new ingredient in the halo model & the impact on galaxy bias. *arXiv e-prints*, art. arXiv:2012.04637, December 2020.

C. Wagner, F. Schmidt, C. T. Chiang, and E. Komatsu. Separate universe simulations. *MNRAS*, 448:L11–L15, March 2015. doi: 10.1093/mnrasl/slu187.

Huiyuan Wang, H. J. Mo, and Y. P. Jing. The distribution of ejected subhaloes and its implication for halo assembly bias. *MNRAS*, 396(4):2249–2256, July 2009. doi: 10.1111/j.1365-2966.2009.14884.x.

R. H. Wechsler and J. L. Tinker. The Connection between Galaxies and their Dark Matter Halos. *ArXiv e-prints*, April 2018.

Risa H. Wechsler, Andrew R. Zentner, James S. Bullock, Andrey V. Kravtsov, and Brandon Allgood. The Dependence of Halo Clustering on Halo Formation History, Concentration, and Occupation. *ApJ*, 652(1):71–84, November 2006. doi: 10.1086/507120.

S. D. M. White. Angular momentum growth in protogalaxies. *ApJ*, 286:38–41, November 1984. doi: 10.1086/162573.

Simon D. M. White and Carlos S. Frenk. Galaxy Formation through Hierarchical Clustering. *ApJ*, 379:52, September 1991. doi: 10.1086/170483.

Jesus Zavala and Carlos Frenk. Dark matter haloes and subhaloes. *Galaxies*, 7:81, 09 2019. doi: 10.3390/galaxies7040081.

Y. B. Zel'Dovich. Reprint of 1970A&A.....5...84Z. Gravitational instability: an approximate theory for large density perturbations. *A&A*, 500:13–18, March 1970.

Andrew R. Zentner. The Excursion Set Theory of Halo Mass Functions, Halo Clustering, and Halo Growth. *International Journal of Modern Physics D*, 16(5):763–815, January 2007. doi: 10.1142/S0218271807010511.

Zheng Zheng, Andreas A. Berlind, David H. Weinberg, Andrew J. Benson, Carlton M. Baugh, Shaun Cole, Romeel Davé, Carlos S. Frenk, Neal Katz, and Cedric G. Lacey. Theoretical Models of the Halo Occupation Distribution: Separating Central and Satellite Galaxies. *ApJ*, 633(2):791–809, November 2005. doi: 10.1086/466510.

Zheng Zheng, Idit Zehavi, Daniel J. Eisenstein, David H. Weinberg, and Y. P. Jing. Halo Occupation Distribution Modeling of Clustering of Luminous Red Galaxies. *ApJ*, 707 (1):554–572, December 2009. doi: 10.1088/0004-637X/707/1/554.