Abstract

During my thesis, I have worked on theoretical aspects of the large scale structure of the Universe. Clustering statistics of galaxies provide a wealth of information both on the initial conditions of cosmic structure formation and on its subsequent gravitational evolution. Extracting robust information from galaxy correlations presents several challenges, of which "galaxy bias" is one of the most significant: the relation between the observed galaxies and the underlying mass distribution generally is non-linear and scale-dependent. In this thesis, I have investigated this relation at the level of dark matter halos, which are the environments that harbor galaxies and clusters, using analytic and numerical approaches. Understanding the clustering properties of the late Universe allows also to glimpse the physics of the very early moments of history of the Universe. I have studied how to constrain signatures of primordial non-Gaussianity, namely if the random seeds of the initial perturbations are to be described with higher than 2-point statistics, from observations of the late Universe.
Theoretical Aspects of Large Scale Clustering of Dark Matter Halos

par

Matteo BIAGETTI

de

Forlì (Italie)

Thèse N° 4956

GENÈVE

Atelier de reproduction de la Section de Physique
2016
The fine understanding of the large scale structure of our Universe is a central theme in cosmology. Clustering statistics of the large scale structure provide a wealth of information both on the initial conditions of cosmic structure formation and on its subsequent gravitational evolution. Extracting robust information from galaxy correlations presents several challenges, of which galaxy bias is one of the most significant: the relation between the observed galaxies and the underlying mass distribution generally is non-linear and scale-dependent. In this thesis, we have investigated this relation at the level of dark matter halos, which are the environments that harbor galaxies and clusters, using analytic and numerical approaches. In particular, we have analyzed in detail predictions of the “excursion set peaks” model, a biasing scheme that starts from the assumption that peaks of the initial Gaussian dark matter field are the preferential environments for the formation of virialized haloes at later time. We have confirmed, through comparison of the model predictions against numerical N-body simulations, that indeed the bias is non-linear and scale-dependent. We have provided specific examples where this modeling can be applied, such as cosmologies with massive neutrinos and the study of the halo velocity bias.

Understanding the clustering properties of the late Universe allows also to glimpse the physics of the very early moments of history of the Universe. We have studied how to constrain signatures of primordial non-Gaussianity, namely if the random seeds of the initial perturbations are to be described with higher than 2-point statistics, from observations of the late Universe. Primordial non-Gaussianity is known to leave an imprint in the formation of structures, the strongest effect being a local modulation of the amplitude of matter fluctuations which generates a scale dependent bias whose signal is amplified at large scales. In this thesis, we have studied this effect in detail, providing new forecasts on the precision with which future galaxy surveys, such as the Euclid satellite, will be able to constrain such signatures. We have also analyzed the expected amplitude of this effect in numerical N-body simulations by measuring the power spectrum and bispectrum cross-correlations between haloes and dark matter. We show that models which assume a simple linear and scale independent bias do not always reach the theoretical precision which will be required for the analysis of the great amount of data coming from future galaxy redshift surveys.
REMERCIEMENTS

Tout d’abord, mes remerciements vont à Vincent, qui a supervisé ma thèse avec beaucoup de patience et de soins. Je lui suis reconnaissant d’être toujours disponible pour discuter des tous les problèmes que je rencontrais, à tout moment. Il m’a constamment encouragé à essayer de nouvelles choses et d’être le plus indépendant que possible.

Deuxièmement, à Toni, qui est la raison en premier lieu pour laquelle j’ai choisi de faire mon doctorat à Genève. Ma première rencontre avec lui à Padoue m’a impressionné au niveau que je savais d’accepter avant même de connaître les détails. Depuis lors, il a été un grand mentor et ami.

Un remerciement va certainement à mes collaborateurs: Hideki, Alex, Enrico, Matteo, Kwan, Aseem, Emanuela, Marco, Azadeh, Davide et Jorge. Il était toujours amusant de travailler avec eux et ce fut la meilleure motivation pour faire de la recherche.

Je dois un grand merci à tout le reste du groupe à Genève, en commençant par les professeurs, Ruth, Michele et Martin, qui m’ont donné un bel exemple de ce que cela signifie d’être accueilli dans un groupe de nouvelles personnes. Je suis également reconnaissant aux secrétaires, Francine et Cécile, qui étaient toujours prêts à me aider avec beaucoup de patience et qui ont toujours pris mes côtés à chaque fois que cela était nécessaire.

Une pensée particulière va aux personnes qui ont partagé leur amitié, autre que leur travail, avec moi: avec Dani, Kwan, Jorge, Azadeh, Julian et Veronica, Thomas, Enea et Alba, Francesco et Ipek j’ai partagé la “vie nocturne” à Genève, ainsi que de nombreuses pauses café. J’ai délibérément séparé Enrico et Davide de cette liste pour les remercier, en plus de leur grande compagnie en particulier au cours de l’année dernière, pour leur soutien dans les tâches d’enseignement, il était un vrai plaisir. Un grand merci également à Elisabetta, qui est toujours prêt à offrir son aide et dont l’invitation à dîner est toujours très bien accueillis et à ses (et Enrico bien sûr) deux belles filles, qui m’ont donné beaucoup de rires et de joie.

Je conclus avec les gens que je dois le plus pour leur soutien dans ces années: mes parents, Marco et Gilda, m’ont donné vraiment tout, le don le plus important étant la capacité de penser et de juger avec ma propre tête. Mon frère Giacomo et ma sœur Elena ont été un réel soutien: notre lien va toujours au-delà de toute barrière de la distance. Je tiens également à remercier la famille de Sara pour être si favorable et en particulier à Davide qui a eu le courage de dormir sur notre canapé deux fois (une fois avec Giacomo, ça c’est vraiment héroïque).

Mes remerciemets finaux et les plus importants sont à Sara. Pour être honnête, il n’y a vraiment pas de mots pour exprimer ma gratitude. Ce que je sais, et avec ça je veux conclure, c’est que toute autre chose de valeur que je vais jamais faire à l’avenir, elle sera ici, à la fin de la liste des remerciements.
ACKNOWLEDGEMENTS

First of all, my thanks are to Vincent, who has supervised my PhD with great patience and care. I am grateful to him for always being available for discussing any trouble I would be stuck with, at any time. He has constantly encouraged me to try new things and to be as much independent as I could.

Second of all, to Toni, who is the reason why I chose to do my PhD in Geneva in the first place. My first meeting with him in Padova impressed me at the level that I knew that I would accept to come even before to know the details. Since then, he has been a great mentor and friend.

A thank goes certainly to the people I have collaborated with closely: Hideki, Alex, Enrico, Matteo, Kwan, Aseem, Emanuela, Marco, Azadeh, Davide and Jorge. It was always fun to work with them and this was the best motivation for doing research.

I owe many thanks to all of the rest of the Geneva group, starting from the professors, Ruth, Michele and Martin, who gave me a nice example of what it means to be welcome in a group of new people. I am also grateful to the secretaries, Francine and Cecile, who were always ready to help me with great patience and who always took my side whenever it was needed.

A particular thought goes to the people who shared their friendship, other than their job, with me: with Dani, Kwan, Jorge, Azadeh, Julian and Veronica, Thomas, Enea and Alba, Francesco and Ipek I have shared some of the night-life in Geneva, as well as many coffee breaks. I purposely separated Enrico and Davide from this list to thank them, besides for their great companionship especially during the last year, for their support in the teaching duties, it was real fun. A special thank also to Elisabetta, who is always ready to offer her help and whose invitation to dinner is always very welcome and to her (and Enrico’s of course) beautiful two daughters, who gave me many laughs and joy.

I conclude with the people I owe the most for their support in these years: my parents, Marco and Gilda, have given me really everything, the most important gift being the ability to think and judge with my own head. My brother Giacomo and my sister Elena have been a real support: our bond goes always beyond any distance barrier. I want also to thank Sara’s family for being so supportive and in particular to Davide who was brave enough to sleep on our couch twice (once with Giacomo, that is pure courage).

My final and most important thanks are to Sara. To be honest, there really are no words to express my gratitude. What I know, and with this I would like to conclude, is that anything else good that I will ever do in the future, she will be here, at the end of the acknowledgments list.
I would like to thank them for having accepted to be part of this jury and for their careful reading of my thesis. A particular thank to Emiliano, who took the time to come to Geneva for a couple of days, he contributed to a very pleasant discussion.
LIST OF PUBLICATIONS

This thesis is based on the following papers, in order of appearance:


La cosmologie est la science qui étudie la formation et l'évolution de l'Univers dans son ensemble. Le domaine de la cosmologie a connu un énorme progrès au cours des deux dernières décennies, se transformant en une science de précision: les mesures des anisotropies de température du fond diffus cosmologique, de l'expansion accélérée de l'Univers à partir de l'observation de Supernovae de type Ia, et des structures à grande échelle ont conduit au développement d'un modèle cosmologique standard, appelé ΛCDM. A grande échelle, ce modèle décrit un univers homogène et isotrope, asymptotiquement plat et dont l'énergie est dominée aujourd'hui par une constante cosmologique, Λ, tandis que la formation de structures cosmiques est réalisée avec une composante de matière de nature inconnue, appelée “matière sombre”.

La mission Planck, qui a exploré en détail le fond diffus cosmologique avec une sensibilité en température de l'ordre de quelques micro Kelvin, est un des exemples les plus récents de la réussite de ce modèle, qui est en mesure d'expliquer avec une précision étonnante les statistiques de ces anisotropies avec seulement 6 paramètres. Toutefois, malgré ce succès, notre compréhension des principes fondamentaux régissant ce modèle ΛCDM est toujours sur un terrain instable.

Pour tester davantage notre compréhension des lois qui régissent l'Univers, une analyse approfondie de son évolution récente et, en particulier, de la distribution de matière à grande échelle, s'avère nécessaire. Les grands recensements de galaxies fournissent des cartes trois-dimensionnelles toujours plus grandes et complètes de la distribution des galaxies et des amas de galaxies en fonction de leur redshift et de leur position angulaire. Les analyses de cette énorme quantité d'information doivent être soutenues par une base théorique solide pour garantir une interprétation correcte des résultats.

La structure à grande échelle de notre Univers est un thème central de la cosmologie actuelle. L'analyse des statistiques de la distribution des galaxies ainsi que leur évolution fournit des informations à la fois sur les conditions initiales de notre Universe et sur son évolution gravitationnelle. Extraire des informations solides à partir des corrélations de galaxies présente plusieurs défis, dont le biais galactique est l'un des plus importants. Cette relation entre les galaxies observées et la distribution de matière sous-jacente est généralement non-linéaire et dépendante d'échelle.

Dans cette thèse, nous avons étudié cette relation au niveau des halos de matière sombre, qui sont les objets abritant les galaxies et les amas. Nous avons utilisé des approches analytiques et numériques. En particulier, nous avons analysé en détail les prévisions du modèle de biais connu sous le nom de “ excursion set peaks”. Ce modèle fait l'hypothèse que l'effondrement des pics du champ de densité initial conduit à la formation des halos virialisés. En comparant les prédictions de ce modèle avec des données
extraits de simulations numériques à N-corps, nous avons confirmé que le biais est non-linéaire et dépendant d’échelle. De plus, nous avons fourni des exemples précis dans lesquels cette modélisation peut être appliquée. Ceux-ci incluent des cosmologies avec des neutrinos massifs, ainsi qu’un biais (statistique) de la vitesse des halos par rapport à celle de la matière sombre.

Comprendre l’agrégation de la matière dans l’Univers permet également de sonder la physique des tout premiers moments de l’histoire de l’Univers. Nous avons étudié la façon de contraindre les signatures de non-gaussianité primordiale, qui indique si les perturbations initiales sont décrites avec des statistiques supérieures à 2 points. Une non-gaussianité initiale laisse une empreinte dans la formation de structures, en particulier sous la forme d’une modulation locale de l’amplitude des fluctuations de densité de matière. Ceci engendre un biais dépendant d’échelle dont le signal est amplifié à grande échelle. Dans cette thèse, nous avons étudié cet effet en détail, fournissant de nouvelles prévisions sur la précision avec laquelle les grands recensements futures, telles que le satellite Euclid, seront en mesure d’observer cette signature. Nous avons également analysé l’amplitude de cet effet dans des simulations numériques à N corps en mesurant les corrélations croisées du spectre de puissance et du bispectre entre les halos et la matière sombre. Nous avons démontré que les modèles qui supposent un simple biais linéaire et indépendent d’échelle n’atteignent pas la précision théorique nécessaire pour l’analyse de la grande quantité de données provenant des prochains recensements de galaxies.
CONTENTS

1 THE PEAKS APPROACH TO HALO CLUSTERING 25
1.1 A brief introduction to halo biasing 27
   1.1.1 The Press-Schechter formalism and correlation functions of high density regions 28
   1.1.2 The spherical collapse model 28
   1.1.3 The Press & Schechter formalism 30
   1.1.4 The two-point correlation function of high density regions 31
1.2 The peak-background split ansatz 32
1.3 The statistics of maxima of a gaussian field 34
1.4 The excursion set approach 38
1.5 Final remarks 40
2 THE EXCURSION SET PEAKS MODEL 41
   2.1 An excursion set for peaks 42
      2.1.1 Making the connection 42
      2.1.2 A more realistic picture: ESP with moving barriers 43
      2.1.3 Comparison with numerical simulations 44
   2.2 The ESP biasing scheme 46
      2.2.1 The effective bias expansion 48
   2.3 Measuring nonlocal Lagrangian peak bias 49
      2.3.1 Bias factors $b_{ijk}$: Hermite polynomials 49
      2.3.2 Bias factors $\chi_i$ and $\omega_{j0}$: Laguerre polynomials 51
      2.3.3 Test with numerical simulations 53
   2.4 Conclusions 60
3 THE HALO BOLTZMANN EQUATION 63
   3.1 From the Klimontovich-Dupree equation to the Boltzmann equation 66
   3.2 Halos as extrema of the smoothed density field 69
   3.3 The halo Boltzmann equation: the left-hand side 70
      3.3.1 Renormalization at the linear order 72
      3.3.2 Renormalization at higher order with the path integral technique 74
      3.3.3 Non-renormalization of the DM Boltzmann equation 77
   3.4 The halo Boltzmann equation: the right-hand side 77
   3.5 Discussion and conclusion 80
4 NON LOCAL HALO BIAS WITH AND WITHOUT MASSIVE NEUTRINOS 83
   4.1 Non-local bias up to third-order 84
Contents

4.1.1 Non local Lagrangian bias and velocity bias 85
4.2 Perturbative approach to cosmologies with massive neutrinos 88
4.3 Bias Parameters 90
4.4 Results 92

II CONSTRaining PRIMordial NON-GAUsSIANITY THROUGH THE LARGE SCALE STRUCTURE OF THE UNIVERSE 97
5 SIGNATURES OF PRIMORDIAL NON-GAUSSIANITY IN THE UNIVERSE 99
5.1 The Suyama - Yamaguchi inequality 101
5.2 The running of primordial non-Gaussianity 102
5.3 Scale dependent bias from the inflationary bispectrum 103

6 FORECASTING PRIMORDIAL NON-GAUSSIANITY THROUGH LSS OBSERVABLES 109
6.1 Testing the Suyama-Yamaguchi inequality 109
   6.1.1 The Fisher matrix formalism applied to galaxy surveys 110
   6.1.2 Halo model predictions 113
   6.1.3 Results and conclusions 114
6.2 Testing the running of primordial non-Gaussianity 117
   6.2.1 Conclusion 122

7 FACING THE CHALLENGE OF PRECISION MEASUREMENTS OF THE AMPLITUDE OF THE NON-GAUSSIAN BIAS 123
7.1 Non Gaussian bias in the framework of the ESP model 124
   7.1.1 A closer look at the peak prediction 127
   7.1.2 The squeezed limit of the galaxy bispectrum 129
   7.1.3 Final remarks 130
7.2 Non Gaussian bias from N-body simulations 132
   7.2.1 The N-body simulations 133
   7.2.2 Halo mass function 134
   7.2.3 Linear bias 135
   7.2.4 Final results and discussion 142
7.3 Squeezing the halo bispectrum 142
   7.3.1 Halo bispectra in the presence of primordial non-Gaussianity 143
   7.3.2 Comparison to numerical simulations 152
   7.3.3 Conclusion 160

III APPENDIX 187
8 THE ESP MODEL: USEFUL FUNCTIONS 189
   8.1 The curvature function of density peaks 189
   8.2 Bivariate $\chi^2$ distributions 189

9 LAGRANGIAN ESP PERTURBATIVE EXPANSION AND BIAS PARAMETERS 193
10 THE HALO BOLTZMANN EQUATION: DETAILS AND CALCULATIONS 197
Contents

10.1 Conditional probabilities 197
10.2 Cumulants at second-order 200
10.3 The peak force 203

11 NON LOCAL BIAS WITH AND WITHOUT MASSIVE NEUTRINOS: DETAILED CALCULATIONS 205
11.1 Non local bias up to third order 205
11.1.1 First-order 205
11.1.2 Second-order 206
11.1.3 Third-order 207
11.2 The halo-mass correlator 212

12 NON GAUSSIAN BIAS FROM N-BODY SIMULATIONS 215
12.1 Convergence of the logarithmic derivative 215
12.2 The stochasticity matrix 216

13 GENERAL EXPRESSIONS FOR THE BISPECTRA 219

14 THE S121 MODEL: THEORETICAL CONSIDERATIONS 221
14.1 Excursion set bias from a peak-background split 221
14.2 Analytic prediction for the halo-matter bispectrum 223
Cosmology as a branch of modern physics can be considered to be born at the beginning of the 20th century. Einstein’s theory of general relativity in 1915 [9] raised from the very beginning the question of how to apply the theory to the entire Universe. It was clear that the field equations of general relativity could not allow a static Universe, but instead a constantly evolving and expanding (or collapsing) environment. This idea was particularly difficult to accept due to the common philosophical, religious and cultural background of an unchanging and static Universe which had its roots in many centuries of history; indeed, in a paper in 1917, Einstein forced the general relativity equations of motion to predict a static Universe by introducing a cosmological constant by hand with negative pressure, to counteract the effect of gravity [10]. On the same year, De Sitter found an alternative scenario describing a Universe without matter as a solution of Einstein’s equations [11]. In 1922, Friedmann was the first to introduce an expanding homogenous and isotropic Universe as a solution to Einstein’s equations [12], which was then confirmed by Lemaître in 1927 [13] and by Hubble in 1929 [14].

Observations in the early 1920s started to raise some serious questions about the distribution of stars in our galaxy: at the time, the concept of the existence of galaxies other than ours was not yet fully conceived, so that, at first, observations of distant galaxies were interpreted as “faint spiral nebulae”. The so-called “Great Debate” [15] dealt precisely with this fundamental question and it was solved by Edwin Hubble’s observations of the Cepheid variable stars in the Andromeda M31 galaxy, with which he proved that these stars were so distant from our galaxy that they must belong to another one [16].

Along with the discovery of the existence of many galaxies in the Universe came the question of homogeneity. Einstein, together with De Sitter, was the first to speculate that while the distribution of stars was clearly inhomogeneous, the systems in which they were included, called “island Universes” at the time, could be uniformly distributed [17].

The Cosmological Principle1, stating that the Universe is homogeneous and isotropic on large scales, had therefore already been introduced as a starting point for cosmology a few years after Einstein’s general relativity article. Its observational confirmation, however, came much later and its precise definition of “large scale” is still under debate. One of the clearest indication of the validity of this principle is the discovery of the Cosmic Microwave Background (CMB) by Penzias and Wilson in 1964 [19] and the subsequent higher precision measurements (see results from the Planck satellite mission

---

1 The first to name it “cosmological principle” was Milne in 1933 [18].
[20] for the most recent measurements) proved this background to be a very isotropic picture of the very early Universe. Isotropy, however, does not imply homogeneity without assuming that our point of observation is not privileged in the Universe, which is a stronger version of the Copernican Principle.

Here we have to make an important remark: the field of cosmology is a peculiar one among the many branches of modern physics. We have only access to one observable Universe, the one we live in, that is to say, in a statistical sense, a single realization. Hence, the point is not only to have to assume the Copernican Principle, that is, that we are not privileged observers, but also that it is impossible in cosmology to “repeat the experiment” many times and validate our predictions based on a statistical analysis. Here the Cosmological Principle has a crucial importance: if we assume that on sufficiently large scales the Universe is homogeneous and isotropic, then when we are averaging over many small patches of the sky it is as if we were looking at many realizations of the Universe. This corresponds to assuming that each path is a fair sample of the statistical ensemble. This concept is known under the name of the ergodic theorem applied to cosmology. It requires that spatial correlations decay sufficiently fast at large separation, so as to consider each patch independent and therefore replace the ensemble average with the spatial average.

If the Universe can be considered homogeneous, on first approximation, at large scales, the next step is to be able to explain how, on general grounds, the (relatively) small scale structures, such as galaxies and clusters, are formed. Since the early 1930s, Lemaître [21] and Jeans [22] argued that gravitational instabilities may cause an initially uniform distribution of matter to slowly fragment into structures. The problem was to characterize this instability in the framework of an expanding Universe: in this case, the exponential growth is suppressed by the expansion into a power law in time. The slowing down of the growth has as the important consequence that the evolving distribution of matter retains memory of the initial conditions. Hence, the theories of structure formation must embed the treatment of the initial state and of the primordial perturbations to be predictive. The accurate study of the CMB have indeed revealed the presence of very small Gaussian and adiabatic fluctuations of the temperature, of the order of $10^{-5}$, first measured by COBE [23] in 1992. These fluctuations tell us about precisely about the primordial seed that, through gravitational evolution, evolved in the later distribution of galaxies and matter we see today in the Universe.

The leading scenario that provides such perturbations is inflation [24]. The inflationary scenario was first introduced by Guth in 1981 as a solution to well-known problems mostly related to CMB observations. For example, in the “horizon problem”, regions of the sky with a causally disconnected past appear to us with very similar characteristics, in particular with a CMB temperature extremely isotropic; the flatness problem dealt instead with the fact that, according to the dynamic equations obtained using the Einstein equation, the Universe would have been flat with a very high precision in the past (e.g. at Plank times $t_P \sim 10^{-42}$, one part in $10^{-60}$) in order to give the observed
The curvature of the space-time metric today (which is consistent with zero). The idea of
cosmic inflation is that at early times, the expansion of the Universe began to accelerate,
as a consequence of the fact that the energy density of the Universe was dominated by
the vacuum energy associated to the potential of a scalar field, called the “inflaton field”. In
this scenario, primordial perturbations are generated by quantum fluctuations of the
scalar fields (one or more) dominating the energy density at that time. These fluctua-
tions are redshifted outside the cosmological horizon on very large scales, where their
amplitude remains “frozen” on superhorizon scales. They subsequently reenter the hori-
zon as curvature perturbations at the end of inflation and are afterwards observable as
temperature anisotropies in the CMB.

Significant progress in the development of a theory of structure formation was achieved
in 1970s and 1980s. At first, it seemed reasonable to deal with two component models,
baryonic matter and radiation, considering neutrinos subdominant as long as they were
massless. In such two component models, one could distinguish two main classes of
structure formation: the “top-down” scenario [25] and the “bottom-up” scenario [26, 27].
In the first scenario, structures form first on large scales, at masses $M \approx 10^{12} - 10^{14} M_\odot$, into clusters and superclusters, undergoing subsequent fragmentation to form smaller
and smaller structures. This model is dominated by adiabatic perturbations, where matter
and radiation fluctuations are coupled together. In the second case, the opposite
scenario is favored: structures form first on small scales, at masses $M \approx 10^5 - 10^6 M_\odot$, subsequently accreting to form larger and larger structures in what is called “hierarchi-
cal clustering”; perturbations in this model are isothermal, that is, radiation fluctuations
are negligible. Both scenarios are tightly constrained by the measurements of the CMB
anisotropies, since they predict a very large amplitude for the fluctuations inconsistent
with CMB observations.

These problems gave motivation to study models with a third, non-baryonic dark mat-
ter component. In such models, perturbations can be classified into curvature and isocur-
vature type. In both cases, all the three components are perturbed, with the difference
that curvature perturbations cause fluctuations in the energy-density and consequently
in the curvature of space-time, while isocurvature perturbations are characterized by
surfaces of constant total energy density, so that the total spatial curvature is unaffected.
The idea of the existence of an exotic component of invisible matter interacting gravita-
tionally was present ever since 1933, when Zwicky observed that visible matter could
not explain the measured galaxy velocities in the Coma cluster [28]. Since its discovery,
beside the velocity dispersion of galaxies inside clusters, there have been many indirect
evidences of the existence of dark matter, such as, for example, weak and strong gravita-
tional lensing, X-ray emission by hot gas in clusters, bullet clusters and of course angular
CMB fluctuations. In spite of the amount of cosmological probes favoring the existence
of a dark matter component in the Universe, its characteristics still remain elusive today.
Beside being subject to gravitational interaction, we know that electromagnetic interac-
tions with other baryonic particles and itself must be negligible and that most probably
it is a particle, rather than a finite size compact object, even though alternative scenarios like, for example, the possibility of being constituted by primordial black holes have been studied [29].

In the context of a structure formation model dominated by dark matter, two classes can be distinguished and are related to the “top-down” and “bottom-up” picture that were known already from the earlier models we introduced above. In the hot dark matter (HDM) scenario, the energy density is dominated by collisionless particles which decoupled from the other components when they were still in the relativistic regime, so that their velocity dispersion is generally very large and their mass small. A typical example of HDM are massive neutrinos with mass $M_\nu \sim 10$eV. Structures in this scenario have a “top-down” type of formation, being the accretion at small scales particularly slow due to the particles large velocity dispersion. In the cold dark matter CDM scenario, the opposite scenario of a “bottom-up” picture is considered, where particles are also collisionless but with small velocity dispersion, having decoupled from other components deep in the non-relativistic regime. The mass of CDM particles is therefore generally very large as compared to HDM ones. Note however that light candidates exist as well: this is possible if the particle has never been in thermal equilibrium with the other components, such as for axions [30]).

Observations of the CMB temperature anisotropies since their discovery in 1992 by the experiment COBE [23] and galaxy clustering data have strongly favored curvature perturbations over isocurvature perturbations in the last decades (see the WMAP [31] and Planck [32] experiments for the most recent results) and a CDM scenario. Moreover, measurements of the distance-redshift relation for type Ia supernovae (SNIa) in 1998 [33, 34, 35] revealed the accelerated expansion of the late Universe, so that a CDM with a cosmological constant $\Lambda$, the so-called $\Lambda$CDM model, is the standard cosmological model that best fits the data as of today.

The accurate study of the formation and evolution of structures over time reveals therefore a number of very compelling challenges: what is the nature of dark matter? Can we explain the measurements of type Ia supernovae with a simple cosmological constant? Is general relativity still a good description of the Universe at very large scales? Answering these questions is the goal of current cosmological probes that aim at the statistical analysis of galaxy clustering through galaxy redshift surveys, such as VIMOS [36], WiggleZ [37], DES [38], etc. These surveys provided independent constraints on cosmological parameters, complementary to the CMB ones. Forthcoming surveys such as Euclid [42], SKA [43] and LSST [44] will improve even more such constraints and look for new physics beyond the standard $\Lambda$CDM model.

What emerges from this brief overview of the historical process that brought to the $\Lambda$CDM scenario as the most predictive model as of today is that everything that drives the structure formation on large and small scales is dark, that is, invisible to our direct

---

2 The observation of the universe on its largest scales dates back to the late eighties ( CfA [39] ) and subsequently evolved into larger and larger surveys such as 6dFGS [40], SDSS [41], etc.
observations. Dark matter (DM hereafter) and dark energy (DE hereafter) are the main components of the energy budget of the Universe and drive the formation of structures as we see them today, yet we have only indirect evidence of their existence and we know very little of their nature. Every consideration we can make on DM and DE has therefore to go through our observation of the visible Universe, that is, of luminous tracers of the DM such as galaxies, quasars, clusters of galaxies, etc., made of baryonic matter. The fundamental issue to face is therefore of how to relate the distribution of luminous tracers with the underlying DM field. The study of this relation goes under the name of theory of bias and it is the main subject of the first part of this thesis.

To investigate the properties of the statistical distribution of luminous objects on large scales such as galaxies and cluster of galaxies, the typical quantity of interest is the N-point correlation function of such objects. The existence of an ergodic theorem applied to the entire Universe allows to take ensemble averages of correlation functions by integrating over volumes. This is indeed how Kaiser in 1984 [45] pioneered three decades of galaxy biasing studies: he was able to explain the unexpectedly high correlations at large scales of Abell clusters by calculating the two-point correlation function of regions of DM lying above a high density threshold. He showed that these statistics are different than the ones obtained by computing the correlation of two randomly chosen DM points and need to be rescaled, on first approximation, by a factor which depends on the mass and redshift of the region considered.

Since 1984, several approaches have been introduced to tackle this issue. The final goal of a theoretical study of galaxy biasing is being able to calculate N-point correlation functions of tracers of the DM field. This is precisely what galaxy and cluster surveys can provide given their constantly increasing catalog of observed galaxies classified by their redshift and angular position in the sky. In the first part of this thesis, we discuss a few theoretical approaches to the study of the formation of the large scale structures of the Universe. In Chapter §1, we overview how we can relate the statistics of overdense regions of the DM, the so-called DM halos, to the distribution of the underlying DM field. 3 We focus on analytic models of collapse and halo clustering statistics, which try to characterize this process starting from first principles, such as the Press & Schechter formalism [46], the peak-background split ansatz [45], the peaks model [47] and the excursion set approach [48]. In the following Chapter §2, we describe in details the excursion set peaks model, a combination of the peaks model in the framework of an excursion set approach. We then compare predictions for the halo mass function and for non-local bias parameters against numerical N-body simulations. Chapter §3 investigates another prediction of the peaks model, namely the existence of a statistical velocity bias between the halos and the DM field. We formulate this problem in the context of

---

3 This constitutes only a first step to the issue of galaxy biasing, as galaxies populate DM halos in a complicated, non linear way. In this thesis, we choose to restrict to this simpler problem of studying the only dark matter clustering. The study of how dark matter halos are populated with galaxies can be carried on separately and, on a theoretical ground, the modeling of the halo biasing has already a number of complications still to be solved without including the further complication of their relation with galaxies.
Boltzmann equations and we compute the effect to second order in perturbation theory. Concluding the first part, in Chapter §4 we show how to compute bias parameters up to third order in Eulerian perturbation theory and we apply this results to a Universe with massive neutrinos, showing that predictions where we include non-local effects in the bias up to third order agree well with numerical N-body simulations.

In the second part of the thesis, we utilize the excursion set peaks model to make theoretical predictions for large scale structure observables with a particular interest in constraining the physics of the early Universe, i.e. the nature of the mechanism of production of the primordial perturbations. Despite the great success of the inflationary scenario, it is also true that there is a plethora of different models of inflation that provide the properties constrained by CMB observations. In fact, producing Gaussian and adiabatic curvature perturbations is rather simple, so that we are not able to distinguish any model of inflation as long as it satisfies these conditions. One of the features that could help to discriminate viable models is the statistical nature of the primordial perturbations: they are constrained to be nearly gaussian, but even a small amount of non-Gaussianity would rule out the entire class of single field models of inflation. The second part of this thesis investigates about how we can observe such primordial non-Gaussianity features in the large scale structures of the Universe today. For instance, it is well known that the two point correlation function of biased tracer contains information about this primordial non-Gaussianity (PNG) through a unique scale dependent bias feature which peaks at very large scales [49, 50, 51]. We review in detail this effect in Chapter §5, where we also introduce some features of inflationary models, such as the Suyama-Yamaguchi inequality [52] and the running of PNG, which could help in discriminating the mechanism for the production of the primordial perturbations. In Chapter §6, we make forecasts on the precision with which future experiments such as Euclid will be able to constrain such signatures of PNG using the Fisher matrix formalism. In Chapter §7 we discuss the accurate measurement and prediction of the amplitude of the scale dependent bias signature in the presence of PNG: we use numerical N-body simulations to measure the effect of primordial non-Gaussian initial conditions in the two- and three-point function of halos and dark matter and we provide theoretical predictions of these quantities from the excursion set peaks model, comparing the results. We conclude with our final remarks and summary of the results.
Part I

THE PEAKS APPROACH TO HALO CLUSTERING
A BRIEF INTRODUCTION TO HALO BIASING

Over the last decade, we have accumulated a good deal of observational evidence that the large scale structure of the Universe originated from seed fluctuations in the very early Universe, which grew via gravitational instability at later times. In this picture, galaxies form within dark matter halos, which are therefore to some extent the building blocks of what we observe. Knowing how dark matter halos form and evolve under the action of gravity is crucial to understand the environment that harbors galaxies and clusters. Understanding the clustering properties of halos represents, therefore, a key ingredient in comprehending galaxy statistical properties. The important goal of such a study is to relate observable properties of tracers such as the density contrast of galaxies to the underlying matter distribution and ultimately to the initial conditions.

Heuristic arguments like the peak-background split [45], and approximations like local bias [53] have been very helpful for modeling the clustering of dark matter halos and extensively used to predict and compare with both observations from galaxy and cluster surveys and measurements from N-body simulations. A number of analytic approaches like the peak model [47], the excursion set framework [48] or perturbation theory (see e.g. for a review [54]) have been developed as well to better understand the physics of the problem and to improve predictions, as the advent of large scale galaxy surveys have considerably increased the need for an accurate description of halo clustering. These approaches have the advantage of being built on basic and well motivated physical principles, but the drawback of lacking predictability on non-linear scales, where the clustering becomes important.

PLAN OF THE CHAPTER This chapter is organized as follows. In Sec. §1.1 we summarize the basic aspects of the spherical collapse model and the Press-Schechter formalism and overview how the biasing problem was first raised in [45]. In Sec. §1.2 we review the peak-background split ansatz. In Secs. §1.3 and §1.4 we summarize two analytic approaches that have been developed to study the halo biasing, namely the peak model and the excursion set theory. In Sec. §1.5 we make a few final remarks.
1.1 THE PRESS-SCHECHTER FORMALISM AND CORRELATION FUNCTIONS OF HIGH DENSITY REGIONS

As we discussed in the introduction, in the standard CDM scenario the growth of structures follows an hierarchical clustering, that is, virialized objects form first on small scales and subsequently grow on larger and larger scales. These bounded objects appear when regions of the dark matter field have undergone sufficient nonlinear evolution such that the primordial density perturbation have grown enough for the region to decouple from the expansion and collapse. Visible galaxies and clusters tend to arise from the baryonic material that is mostly dense in these overdense regions of the dark matter field.

1.1.1 The spherical collapse model

Given that the collapse process is highly non-linear, most of its characteristics have to be studied with the help of numerical simulations. An analytic approach can be nevertheless possible if a number of assumptions are made, as in the spherical collapse model, first introduced in [55]. In this framework, the central assumption is that, at some early time $t_i \ll 1$, there is a spherically symmetric density perturbation at the center of our coordinate system within a radius $R_i$ and with density $\bar{\rho}(t_i)(1 + \delta_{R_i}(x,t_i))$, being $|\delta(R_i)| \ll 1$ still linear at this early times. The mass inside this region is therefore

$$M_i = \frac{4\pi}{3}R_i^3\bar{\rho}(t_i)(1 + \delta_{R_i}(x,t_i)).$$  \hspace{1cm} (1)

Here we consider a flat background Universe in a matter-dominated stage, with a density $\bar{\rho}(t) = \rho_c(t)$, where the critical solution for such a background is the Einstein de-Sitter (EdS) Universe with scale factor $a \propto t^{2/3}$ and density $\rho_c(t) = 1/(6\pi G t^2)$. Because of the spherical symmetry (Birkhoff’s theorem), this region evolves independently from the outer mass distribution and each concentric shell remains concentric and evolves like a closed ($\delta > 0$) or open ($\delta < 0$) Friedmann Universe. Friedmann equations for a closed Universe have a parametric solution that we can use to study the time evolution of an overdense shell,

$$\frac{R(\theta)}{R_i} = A(1 - \cos \theta),$$  \hspace{1cm} (2)

$$\frac{t(\theta)}{t_i} = B(\theta - \sin \theta),$$  \hspace{1cm} (3)

where $R$ is the physical radius. $A$ and $B$ can be constrained to be related by $B^2/A^3 = 9/2$ by imposing that we recover the background solution at early times, which corresponds
1.1 THE PRESS-SCHECHTER FORMALISM AND CORRELATION FUNCTIONS OF HIGH DENSITY REGIONS

to $\theta \ll 1$. The ratio of the average density within the shell to that of the background at a generic time $t$ is therefore

$$1 + \Delta = \frac{M}{V(t) \rho_c(t)} \approx \frac{R^3_i \rho_c(t_i)}{R^3(t) \rho_c(t)} \approx \frac{9}{2} \frac{(\theta - \sin \theta)^2}{(1 - \cos \theta)^3}. \quad (4)$$

At $\theta = \pi$ the Hubble expansion of the shell is exactly balanced by its self-gravity and reaches the maximum radius, $R_{\text{max}}$. Subsequently, the region starts collapsing until, at $\theta = 2\pi$, the nonlinear solution formally collapses to a point, $R \to 0$ and the density tends to infinity, $\Delta_{\text{max}} \to +\infty$. However, shell-crossings make the assumptions of the spherical collapse invalid for small radii. Due mainly to angular momentum conservation the region remains finite and virializes, forming a bound object. The ultimate state of collapse is a system in equilibrium whose structure is governed by collisionless dynamics. For collisionless systems, the energy is conserved, so that during the collapse the gravitational potential energy is converted to kinetic energy of the particles. The sphere eventually relaxes to a quasi-static structure supported by random motions of the particles. To compute the final radius of this bounded region, we can therefore use the virialization theorem which states that at equilibrium one has to impose that

$$2K_{\text{vir}} + W_{\text{vir}} = 0, \quad (5)$$

where $K$ is the kinetic energy and $W$ is the potential energy, and at the same time that the energy is conserved,

$$E = K_{\text{vir}} + W_{\text{vir}}, \quad (6)$$

so that we get

$$E = \frac{W_{\text{vir}}}{2} \approx -\frac{GM}{2R_{\text{vir}}} = -\frac{GM}{R_{\text{max}}}, \quad (7)$$

which shows that $R_{\text{vir}} = R_{\text{max}}/2$ and, assuming that $t_{\text{vir}}$ occurs at $t = 2\pi$, $t_{\text{vir}} = 2t_{\text{max}}$. Using Eq. (4) we finally get

$$1 + \Delta_{\text{vir}} \approx \frac{R^3_{\text{vir}} \rho_c(t_{\text{max}})}{R^3(t) \rho_c(t)} \approx 178. \quad (8)$$

Notice that this value does not depend on the radius (and therefore the mass) of the region we considered, therefore this suggests that all virialized objects will have the same density relative to the background within this model.

It is interesting to compare this result to the one obtained by applying linear theory only. A simpler calculation is performed for the linear density field, $\delta_L$, which in a EdS Universe scales with time as

$$\delta_L = \frac{3}{5} \delta_i \left( \frac{t}{t_i} \right)^{2/3} = \frac{3}{5} \left( \frac{3}{4} \right)^{2/3} (\theta - \sin \theta)^{2/3}, \quad (9)$$
and therefore at $\theta = 2\pi$ we get

$$\delta_{sc} = \frac{3}{5} \left( \frac{3\pi}{2} \right)^{2/3} \approx 1.687,$$

so that, as expected, the non-perturbative result is much greater than the linear one, $\Delta_{\text{vir}} \gg \delta_{sc}$.

The value of the spherical collapse threshold is frequently used to define collapsed objects in numerical simulations, while its linear counterpart is employed in models that describe the clustering of virialized object from the initial density fluctuation field. This result has its central weakness in the basic assumption that the initial perturbation has spherical symmetry, we will deal with this choice in more detail in the next chapters.

### 1.1.2 The Press & Schechter formalism

Utilizing the spherical collapse model one can predict the abundance of virialized objects of mass $M$ at a certain redshift $z$. This calculation was first performed in [46] and goes under the name of Press-Schechter formalism.

Primordial density perturbations that grow via gravitational interaction to form the non-linear structure we see today have been thoroughly tested through observations of the CMB to be very close to a Gaussian distribution [31, 56]. It is therefore reasonable to consider the initial density fluctuation as a Gaussian random field.

Even though the main goal is to calculate the abundance at late times, when gravitational clustering has modified the initial Gaussian distribution into a non-Gaussian one, we can work with the initial density field linearly extrapolated at the redshift of interest, and treat the linear density field as a “clock” for the non-linear collapse.

Let us define the smoothed linear perturbations on scale $R$ as

$$\delta_R(x, z) = \int \frac{d^3k}{(2\pi)^3} \delta_L(k, z) W_T(kR) e^{ik \cdot r},$$

where we smooth with a top-hat radius $W_T(x) = 3(\sin x - x \cos x) / x^3$. Fluctuations $\delta_R(x, z)$ above the threshold $\delta_{sc}$ will correspond to virialized objects of mass $M$ at redshift $z$, where the mass-to-radius relation is given by Eq. (1). Since the density field $\delta_R(x, z)$ is Gaussian, its probability distribution is simply

$$P_G(\delta_R) d\delta_R = \frac{1}{\sqrt{2\pi \sigma^2_R}} e^{-\frac{\delta_R^2}{2\sigma^2_R}} d\delta_R$$

with zero mean and variance $\sigma^2_R = \langle \delta_R^2 \rangle$. Hence, we can compute the fraction of Lagrangian volume in halos of mass greater than $M$ to be

$$F_{>M}(\delta_{sc}) = \frac{1}{\sqrt{2\pi \sigma^2_R}} \int_{\delta_{sc}}^{\infty} e^{-\frac{\delta_R^2}{2\sigma^2_R}} d\delta_R = \frac{1}{2} \text{erfc}(\frac{\nu_R}{\sqrt{2}})$$
where we have defined the peak height \( v_R = \delta_{sc}/\sigma_R(z) \). This quantity is not properly normalized, as in the limit of \( M \to 0 \), and therefore \( \sigma_R \to \infty \), the collapse fraction tends to one half of the total mass. This is the well-known cloud-in-cloud problem, that is, the fact that we are not including the possibility that an overdense region could be contained in a bigger one. In other terms, we are considering the only possibility that \( \delta > 0 \), thus excluding underdense regions from the total mass count. We will come back to this point later in Sec. §1.4. One may be interested in the number density of objects that have a mass in the interval \([M, M + dM]\), that is the differential number density per unit volume and unit mass, commonly known as the halo mass function, which reads

\[
\frac{dn}{dM} \equiv \bar{n}(M) = -\frac{\bar{\rho}}{M} \frac{dF_{>M}(\delta_{sc})}{dM} = \sqrt{\frac{2}{\pi M^2}} \bar{\rho} \nu e^{-\nu^2} \frac{d\ln \nu}{d\ln M} \tag{14}
\]

where we corrected for the factor of 2 missing in Eq. (13). This mass function is of the general form

\[
\frac{dn}{dM} = \frac{\bar{\rho}}{M^2} f(\nu) \frac{d\ln \nu}{d\ln M}, \tag{15}
\]

where \( f(\nu) \) is called the multiplicity function and for universal mass functions, such as the Press & Schechter one, it is an only function of \( \nu \). Notice that with this treatment we are assuming that all the mass is inside halos. From this abundance one can identify a characteristic mass \( M_\ast(z) \) that corresponds to the mass at which \( \nu = 1 \), that is when \( \sigma_R = \delta_{sc} \). We can therefore identify two regimes: for halos of mass lower than \( M_\ast \), the scaling of the halo mass function is \( \bar{n}(M) \propto M^{-2} \), while for higher masses it is exponentially suppressed by \( e^{-\nu^2/2} \).

### 1.1.3 The two-point correlation function of high density regions

Having a prediction for the abundance of halos, the natural step forward is to calculate their N-point correlation functions to characterize their statistical distribution. Ever since the pioneering work of [45], it is known that the statistical distributions of galaxies and clusters are biased with respect to the one of the dark matter field, as they act as its tracers, forming where it is most overdense. In [45], it was first argued that the correlation function of very massive objects, such as the Abell clusters may be different from the one of a randomly selected dark matter field point. This is due to the fact that they form in regions where the primordial density perturbations lay above a certain threshold.

Let us consider for example the correlation function \( \xi_{>\nu}(r) \), that is, the correlation between two regions lying above a certain threshold. This is the fractional excess probability that the density in one point lay above a threshold \( \delta_2(x_2) > \nu \sigma \) given that it also does for another point \( \delta_1(x_1) > \nu \sigma \), being \( |x_1 - x_2| = r \).
If we stick to the assumption that the initial linear density $\delta$ is a Gaussian distributed variable, the joint probability distribution of $\delta_1$ and $\delta_2$ is therefore

$$P_{12}^G = \frac{1}{2\pi(\sigma^4 - \sigma_{12}^4)} e^{-\frac{\sigma^2(\delta_1^2 + \delta_2^2) - 2\sigma_{12}\delta_1\delta_2}{2(\sigma^4 - \sigma_{12}^4)}}$$

(16)

where we define $\sigma_{12} = \langle \delta_1 \delta_2 \rangle$. Therefore, the probability that $\delta_1$ and $\delta_2$ lay above the threshold is

$$\mathcal{P}(\delta_1 > \delta_{sc}, \delta_2 > \delta_{sc}) = \int_{\delta_{sc}}^{\infty} \int_{\delta_{sc}}^{\infty} d\delta_1 d\delta_2 P_{12}^G.$$  

(17)

We can thus compute the fractional excess probability in the limit $\nu \gg 1$

$$\xi_\nu(r) = \frac{\mathcal{P}(\delta_1 > \delta_{sc}, \delta_2 > \delta_{sc})}{P_{2G}^2} - 1 \nu^{\gg 1} \frac{\nu^2}{\sigma^2} \xi(r).$$  

(18)

We get therefore that the higher is the threshold, the higher is the bias between $\xi_\nu(r)$ and $\xi(r)$ and that this bias generally depends on the smoothing scale $R$.

### 1.2 The Peak-Background Split Ansatz

It is useful at this point to try to understand more in detail the physical meaning of the biasing effect we introduced in the previous section.

Let us suppose that we can split the density field $\delta$ in short- and long-wavelength modes

$$\delta = \delta_L + \delta_S.$$  

(19)

While the long-wavelength $\delta_L$ is a coherent perturbation and can be though of as a smooth background field acting as a constant shift in density in regions much smaller than its characteristic scale, the short-wavelength $\delta_S$ is noisy and acting on the typical scales at which density peaks are formed, as in Fig. (1). Let $n_h(x, M, \delta_{sc})$ be the number density of halos with mass $M$ inside a region of Lagrangian volume $V$ centered in $x$. The short modes $\delta_S$ inside this region will determine the formation of a halo of mass $M$ if they lay above the threshold $\delta_{sc} - \delta_L(x)$, where $\delta_L(x)$ is the constant value of the long mode inside this region. Therefore we can write the overdensity of halos as

---

1 Here we momentarily drop the subscript $R$, although it is understood that $\delta$ is smoothed.
\[ \delta_h(x) \equiv \frac{n_h(x, M, \delta_{sc})}{\bar{n}_h(M, \delta_{sc})} - 1 \]
\[ \approx \frac{\bar{n}_h(M, \delta_{sc} - \delta_L(x))}{\bar{n}_h(M, \delta_{sc})} - 1 \]
\[ \approx -\frac{1}{\bar{n}_h \delta_{\delta_{sc}}} \delta_L(x) + \frac{1}{2} \left( \frac{d^2 \bar{n}_h}{d \delta_{sc}^2} \right) \delta_L^2(x) + ... \] (20)
\[ = b_1 \delta_L(x) + \frac{1}{2} b_2 \delta_L^2(x) + ... \] (21)

being \( \bar{n}_h(M, \delta_{sc}) \) the mean number density of halos of mass \( M \) and wherein the third line we Taylor expanded the expression in the second line.

This calculation shows that the peak-background split ansatz allows to relate the halo overdensity with the long-wavelength density perturbation \( \delta_L(x) \) through a series expansion with deterministic bias parameters \([58]\)

\[ b_N = \frac{(-1)^N}{\bar{n}_h} \frac{d^N \bar{n}_h}{d \delta_{sc}^N}. \] (22)

The physical meaning of this expansion resides in the fact that regions with a higher background density \( \delta_L(x) \) tends to form halos more easily, that is, it has to overcome a lower barrier for collapse, than regions with lower \( \delta_L(x) \). The values of the bias param-
A BRIEF INTRODUCTION TO HALO BIASING

...ters can be computed explicitly for the Press-Schechter mass function in Eq. (14), e.g. for $b_1$ we have\(^2\)

$$b_1 = -\frac{1}{\bar{n}_h} \frac{d\bar{n}_h}{d\delta_{sc}} = -\frac{1}{\bar{n}_h} \frac{d\bar{n}_h}{d\nu} = \frac{1}{\sigma} \left( \nu - \frac{1}{\nu} \right).$$

(23)

The expansion in Eq. (21) is deterministic and local: bias parameters are the response to the local change in the barrier on the number density of halos in that region. Furthermore, it is a Lagrangian bias expansion: we have not dealt with the gravitational evolution yet and we are assuming that the separation of scales in Eq. (19) is possible. Long and short modes are non-linearly coupled by the gravitational interaction, so that at later times the density has a non Gaussian distribution and the scale separation is not possible anymore.

Another point to remark is that, while this relation is deterministic, we saw in the previous Sec. §1.1, that the biasing relation arises from the study of two-point correlation functions, that is from statistical quantities such as $\xi(\nu r)$ and $\xi(r)$, where we averaged over the whole space and therefore lost every information about the local relation between halos and matter. Indeed what we are usually able to extract from observations are correlation functions between objects averaged over the volume of the surveys. The two treatments can be connected by calculating correlation functions using the relation in Eq. (21) For instance, for the two-point function at first order, plugging the Press-Schechter bias computed in Eq. (23), we have

$$\langle \delta_h(x_1)\delta_h(x_2) \rangle = b_1^2 \langle \delta(x_1)\delta(x_2) \rangle + ...$$

$$= \frac{1}{\sigma^2} \left( \nu - \frac{1}{\nu} \right)^2 \langle \delta(x_1)\delta(x_2) \rangle + ...$$

$$\nu \gg 1 \Rightarrow \nu \nu^2 \langle \delta(x_1)\delta(x_2) \rangle + ...$$

(24)

where we took the limit of high peaks $\nu \gg 1$ on the last line, thus recovering what found in the previous section in Eq. (18).

1.3 THE STATISTICS OF MAXIMA OF A GAUSSIAN FIELD

It is rather natural to think that, to some degree of approximation, virialized halos at late time may have formed out of density maxima of the initial density field [47]. This amounts to saying that, if we trace back halos of a certain mass $M$ to the initial conditions, they sit on a peak of the initial density field. The corresponding Lagrangian patch is called “proto-halo” and it defines the halo Lagrangian scale $R \propto (M/\bar{\rho})^{1/3}$. This

\(^2\) Note that in [59] it was first argued that the only knowledge of the unconditional mass function is sufficient to predict the large-scale bias with sufficient accuracy, without having to know the whole merger histories of dark matter halos.
assumption has been extensively analyzed in numerical N-body simulations and shown to work rather well for a wide range of masses \cite{60, 61, 62}.

In Sec. §1.1, we argued that it is well motivated by current observations to assume that the DM field at very early times is described by a Gaussian random field. We will therefore assume that our initial density field is Gaussian distributed and study the maxima of this field.

We can generally define the number density of points that are local maxima of the Gaussian field $\delta(x)$ as a point process,

$$n_{pk}(x) = \sum_{x_{pk}} \delta_D(x - x_{pk}),$$

where $\delta_D$ is the 3-dimensional delta function. The continuous field $\delta(x)$ can be Taylor expanded around a peak position $x_{pk}$,

$$\delta(x) \approx \delta(x_{pk}) + \frac{1}{2} \sum_{i,j} [\nabla_i \nabla_j \delta(x_{pk})](x - x_{pk})_i (x - x_{pk})_j, \quad (26)$$

where we imposed the peak condition on the first derivative, $\nabla_i \delta(x_{pk}) = 0$. In a similar manner, we can Taylor expand $\nabla_i \delta(x)$ around the peak as

$$\nabla_i \delta(x) \approx [\nabla_i \nabla_j \delta(x_{pk})](x - x_{pk})_j. \quad (27)$$

Using Eq. (26) and Eq. (27), we can express the delta functions of Eq. (25) as

$$\delta_D(x - x_{pk}) \approx |\det [\nabla_i \nabla_j \delta(x)]| \delta_D(\nabla \delta(x)), \quad (28)$$

provided that the hessian $\nabla_i \nabla_j \delta$ is definite negative, to ensure to be looking at maxima and not at minima of the field $\delta$, and non singular in $x_{pk}$.

Up to now, we have not imposed any additional condition, other than being a maximum of the DM field, to this density of discrete peaks. As we said, the identification of proto-halos inserts a scale into the problem, as we want to model the clustering of halo centers, and not their internal substructure. We therefore need to smooth the DM field on the Lagrangian size $R_s$. Moreover, we may impose a high enough threshold on the value of the density at the peak, that is, a peak height, to make sure that we are dealing with regions sufficiently dense as to favor virialization, in the spirit of \cite{45}.

As a result, the number density of peaks of height $\nu'$ at position $x$ can be expressed in terms of the smoothed field $\delta_s$ and its derivatives as \cite{47}

$$n_{pk}(\nu, R_s, x) = \frac{3^{3/2}}{R_s^3} |\det \zeta(x)| \delta_D(\eta(x)) \theta_H(\lambda_3(x)) \delta_D(\nu(x) - \nu'), \quad (29)$$
where we have conveniently defined the normalized peak density and its derivatives as

\[ \nu(x) = \frac{1}{\sigma_0 s} \delta_s(x) \]  
\[ \eta_i(x) = \frac{1}{\sigma_1 s} \nabla_i \delta_s(x) \]  
\[ \zeta_{ij}(x) = \frac{1}{\sigma_2 s} \nabla_i \nabla_j \delta_s(x) \]  

(30) (31) (32)

and we will adopt the following notation for the variance of the smoothed density field, (linearly extrapolated to present-day) and its derivatives,

\[ \sigma_j^2 = \frac{1}{2\pi^2} \int_0^\infty dk P(k)k^{2(j+1)}W_s^2(kR), \]  

(33)

where \( P(k) \) is the power spectrum of the mass density field and \( W(kR) \) is the filtering kernel. Moreover, \( R_s \) is the Lagrangian smoothing scale (which may depend on the choice of kernel) and \( R_* = \sqrt{3\sigma_1 s/\sigma_2 s} \) is the characteristic radius of the peak. We have imposed explicitly that the stress tensor \(-\zeta_{ij}\) is definite negative through the positivity of \( \lambda_3 \), its lowest eigenvalue.

The number density \( n_{pk} \) can be used, in principle, to compute any \( N \)-point correlation functions among the peaks of the density field by ensemble averaging products of \( n_{pk} \),

\[ \rho_{pk}^{(N)}(v, R_s, x_1, ..., x_N) = \langle n_{pk}(v, R_s, x_1) \times ... \times n_{pk}(v, R_s, x_N) \rangle, \]  

(34)

and in the Gaussian initial conditions considered here, multivariate normal distribution are assumed to perform the ensemble average. The simplest example is the case of \( N = 1 \), that is, the averaged peak number density

\[ \bar{n}_{pk} = \langle n_{pk}(v, R_s, x) \rangle = \int d\nu d^3\eta d^6\zeta n_{pk}(v, R_s, x)P_1(v, \eta, \zeta_A) \]  

(35)

where the joint probability distribution \( P_1 \) is given by

\[ P_1(y)dy^{10} = \frac{1}{(2\pi)^5|\det M|^{1/2}} e^{-Q_1(y)}dy^{10}, \]  

(36)

and we define \( y = (\nu, \eta, \zeta_A) \), with \( A = 1, ..., 6 \) since \( \zeta \) is symmetric and \( M \) is the corresponding covariance matrix. Owing to rotational invariance, we can factorize the exponential into the following form

\[ Q_1 = \frac{\nu^2 + J_1^2 - 2 \gamma_1 v J_1}{2(1 - \gamma_1)} + \frac{3}{2} \eta^2 + \frac{5}{2} J_2^2 \]  

(37)

where \( J_1 = -\text{tr}(\zeta_{ij}) \) is the peak curvature and \( J_2 = \frac{3}{2}\text{tr}(\zeta_{ij}^2) \), \( \bar{\zeta}_{ij} = \zeta_{ij} + \delta_{ij} J_1/3 \) are the components of the traceless part of the Hessian and we define the correlation

\[ \gamma_1^2 = \frac{\langle \delta J_1 \rangle^2}{\langle \delta^2 \rangle} = \frac{\sigma_4^4}{\sigma_2^2 \sigma_0^2}. \]  

(38)
This decomposition allows to write Eq. (36) in a more compact way as a product of a bivariate Gaussian in the variables \( v \) and \( J_1 \) and 3- and 5- degrees of freedom \( \chi \)-squared distributions in \( 3\eta^2 \) and \( 5J_2 \) respectively,

\[
P_1(y) \, d^{10} y = \mathcal{N}(v, J_1) \, dv \, dJ_1 \, \chi_3^2(3\eta^2) \, d(3\eta^2) \, \chi_5^2(5J_2) \, d(5J_2) \, P(\Omega),
\]

and \( P(\Omega) \) represents the probability distribution of the five remaining d.o.f. Since they are all angular variables, they do not generate bias factors because the peak (and halo) overabundance can only depend on scalar quantities (see e.g. [63, 64]). Without entering into the details of the full calculation (see [47]), the integral Eq. (35) can be simplified into the only integration over \( J_1 \)

\[
\hat{n}_{\text{pk}}(v, R_*, \nu) = \int_0^{+\infty} dJ_1 \, N_{\text{pk}}(J_1, \nu)
\]

where we define

\[
N_{\text{pk}}(J_1, \nu) = \frac{e^{-v^2/2}}{\sqrt{2\pi}} \frac{1}{(2\pi R_*^2)^{3/2}} \, F_1(J_1) \, P_C(J_1 - \gamma_1 v; 1 - \gamma_1^2)
\]

and

\[
F_1(x) = \frac{1}{2} (x^3 - 3x) \left[ \text{erf} \left( x\sqrt{\frac{5}{2}} \right) + \text{erf} \left( x\sqrt{\frac{5}{8}} \right) \right] + \sqrt{\frac{2}{5\pi}} \left( \frac{31x^2}{4} + \frac{8}{5} \right) e^{-5x^2/8} + \left( \frac{x^2}{2} - \frac{8}{5} \right) e^{-5x^2/2}
\]

and we used Bayes’ theorem to write \( \mathcal{N}(v, J_1) = \mathcal{N}(v) \, \mathcal{N}(J_1 | v) \) and consequently written the conditional gaussian distribution \( P_C(x - \mu; s) \) with shifted mean \( \mu = \gamma_1 v \) and variance \( s = 1 - \gamma_1^2 \). The integral can be computed analytically and gives the final result

\[
\hat{n}_{\text{pk}}(v) = \frac{e^{-v^2/2}}{\sqrt{2\pi}} \frac{1}{V_*} \mathcal{G}_0^{(1)}(\gamma_1, \gamma_1 v),
\]

where we define the function \( \mathcal{G}_0^{(1)} \) in the appendix §8.1 and \( V_* = (2\pi R_*^2)^{3/2} \) is the characteristic volume size of the peak.

The explicit calculation of higher order correlations function is rather complicated in this framework, owing to the large number of variables in the problem. The 2-point correlation function was computed in [65] and involves a total of 20 variables and a 20 dimensional covariance matrix. To simplify these calculations, an effective local bias expansion was developed in [66] and subsequently nicely formalized in [67]. We discuss this in more detail in the next chapter.
ON THE LAGRANGIAN RADIUS OF PEAKS

It is useful at this point to remark a very important point about the framework of the peak model. We have stressed that, by imposing that halos at late time are connected to proto-halos centered on peaks of the initial DM field, we introduce a scale in the problem, that is, the Lagrangian peak patch radius. This scale is not arbitrary, on the contrary it is a physical scale connected to the mass of the halo considered and it is indeed observed also in numerical N-body simulations, as found in the analysis of [68].

1.4 THE EXCURSION SET APPROACH

In Sec. §1.1, we introduced the Press & Schechter model to estimate cluster abundances in models characterized by an hierarchical clustering, as is the case for the standard CDM scenario.

A problem that arises from this framework, as we already mentioned, is that it does not appropriately account for the total collapse fraction, the cloud-in-cloud problem: overdense regions may be contained into bigger clouds.

The natural framework for not double-counting smaller overdense regions that are embedded in larger ones is known as the excursion set approach [69, 48, 70, 71, 72]. The goal of this approach is precisely to find regions that are sufficiently overdense on a given smoothing scale, but not on larger ones. To perform this check, one needs to consider the density field $\delta$ at any given point as a function of the smoothing scale. This function looks similar to a random walk, the starting point being the limit of infinitely large smoothing scales, where the overdensity is zero. In this case, the critical density for collapse defines another curve, typically independent of the point of space, namely a “barrier” for the random walk. In the spherical collapse model of Sec. §1.1, this barrier is constant and flat (in R), $B = \delta_{sc}$, but one may consider more involved, and more realistic, examples, as we do in the next chapter.

The cloud-in-cloud problem is solved by identifying the largest smoothing scale on which the walk first crosses the barrier. The excursion set ansatz therefore relates the abundance of halos of mass $M$ to the fraction of random walks that first cross a barrier $B(R)$ on the scale $R$, the halo mass being the one contained inside $R$,

$$M = \frac{4\pi}{3} R^3 \bar{\rho}.$$  \hspace{1cm} (44)

We relate to Eq. (15) by defining the first crossing fraction $f(s)ds$ to be

$$\frac{M dn(M)}{\bar{\rho} dM} dM = f(s) ds,$$  \hspace{1cm} (45)

where it is customary in this framework to use the variance $s = \sigma^2_0(R)$ as the reference variable for the random walk.

The simplest scenario is when steps are uncorrelated, that is the walk is completely stochastic. In this case, one has just to impose that a randomly chosen walk has not
crossed the barrier \( B(s) \) prior to some \( S \). This corresponds to the ‘survival probability’, \( P_u \), that for a constant barrier, \( B(s) = \delta_{sc} \), is given by \( P_u(s) = \text{erf}(\delta_{sc}/\sqrt{2s}) \). The first crossing distribution is therefore [48, 73]

\[
f_u(s)ds = -\frac{\partial P_u(s)}{\partial s} = \frac{\delta_{sc}}{\sqrt{2\pi s}} e^{-\frac{\delta_{sc}^2}{2s}} ds.
\]  

(46)

Note that this is precisely 2 times what found in Eq. (13) as expected, since we are explicitly solving the cloud-in-cloud issue. Analytical solutions in this case exist also for linear barriers [71] and can be approximated for more generic barriers [74, 75].

The assumption of the steps being uncorrelated is, however, not well motivated: even though different k-modes in the initial Gaussian field are indeed uncorrelated, this is generally not the case in real space [73, 76, 77]. For example, one may think of a fully deterministic walk, in which the density field is directly related to the smoothing, \( \delta \sim \sqrt{s} \). In this case, the walk is completely correlated and the knowledge of one single step provides information about the entire walk, since \( \delta_1 \sim \sqrt{s_1} \delta_0 / \sqrt{s_0} \). Despite this being an apparently unrealistic picture, in [77] it was showed that it actually provides a very good description of the first crossing distribution at small \( s \), that is, in the regime where the smoothing scales are large enough, which is the case of interest for cosmological studies.

The reason for this success is rather intuitive: a walk that retains a certain degree of correlation can be viewed as a stochastic succession of steps deviating from a smooth, deterministic path. The deviation at small \( s \) can not be very large, since it had no time to depart significantly from it yet. In what follows, we overview the treatment of [78] about correlated walks, as it will be preparatory to the discussion about the excursion set peaks model, in the next chapter.

As we said, the excursion set ansatz consists in requiring that \( \delta < B(s) \) for all \( s < S \). This is an infinite set of conditions and, in a generic case, calculating the first crossing distribution can be very complicated. Let us assume that, when the walk is strongly correlated, we can instead impose the condition on the one preceding step, that is, that \( \delta < B(S - \Delta s) \) for \( \Delta s \to 0 \). We can then Taylor expand in \( \delta \) and \( B \) and get the following condition:

\[
B(S) \leq \delta \leq B(S) + \Delta s \left( \frac{d\delta}{ds} - \frac{dB}{ds} \right)
\]  

(47)

with \( d\delta/ds \geq dB/ds \) at \( s = S \), which is a condition on the ‘velocity’ of \( \delta \) being greater than the tilt of the barrier. This implies that now we have to deal with a bivariate distribution on \( \delta \) and its velocity \( \delta' \equiv d\delta/ds \), \( P(\delta, \delta') \), but with the simplification of having only one condition to impose, rather than an infinite number of them. The first crossing distribution reads

\[
f(s)ds \simeq \int_{B'}^{+\infty} d\delta' \int_{B(s)}^{B(s)+\Delta s(\delta'-B')} d\delta P(\delta, \delta'),
\]  

(48)

\[
\simeq \Delta s \int_{B'}^{+\infty} d\delta' p(\delta', B)(\delta' - B').
\]  

(49)
If we now take the limit $\Delta s \to 0$ and change the integration variable $\delta' \to v = \delta' - B'$ we get the general formula

$$f(s) = \int_0^{+\infty} dv \, v \, p(v + B', B). \quad (50)$$

As an example, we can compute Eq. (50) for Gaussian initial conditions and a constant barrier, $B(s) = \delta_{sc}$. In this case, it is convenient to work with the following change of variables

$$\delta_{sc} \to v \equiv \delta_{sc} / \sqrt{s} \quad v \to x \equiv 2\gamma \sqrt{s}v \quad (51)$$

and we have defined

$$\gamma^2 \equiv \frac{\langle \delta \delta' \rangle^2}{\langle \delta^2 \rangle \langle \delta' \rangle}. \quad (52)$$

We obtain therefore the following first crossing distribution, as a function of $v$,

$$f(v) = \frac{e^{-v^2/2}}{\sqrt{2\pi}} \frac{1}{\gamma v} \int_0^{+\infty} dx \, x \, p_G(x - \gamma v; 1 - \gamma^2) \quad (53)$$

where the subscript $G$ indicates the Gaussian distribution and, as for Eq. (40), we have used the conditional probability $p_G(x, \delta_{sc}) = p_G(\delta_{sc}) p_G(x|\delta_{sc})$.

### 1.5 Final Remarks

Understanding the clustering of dark matter halos has been a topic of active research for many years. Astrophysical and cosmological observations tell us that galaxy properties depend strongly on their environment, yet on a theoretical ground it is often assumed that their statistics can be described by a continuous galaxy density field, with the DM halo mass as the only relevant parameter [79]. Although they provide a reasonably good description of the weakly nonlinear clustering of dark matter halos and galaxies, it is yet unclear whether they can meet the recent improvements in computational power and numerical algorithms as well as the accuracy of forthcoming large scale galaxy surveys, which will provide unprecedented descriptions of the clustering of galaxies and clusters.

Moreover, in this chapter we have never taken into account the coherent motions of halos and their evolution through time. Including the gravitational interaction among halos and the DM field implies studying a non-linear and nonlocal process, so that the local Lagrangian bias relation is transformed into a nonlocal Eulerian one. In the next two chapters, we propose some methods to improve our understanding both of the Lagrangian and the Eulerian approaches to halo biasing.
In the previous chapter, we have introduced the issue of halo biasing, with a particular attention to analytic approaches such as the peaks model Sec. 81.3 and the excursion set ansatz Sec. 81.4. The fundamental difference between these two frameworks is that while the peak model describes the statistics of the positions of the maxima in the initial conditions around which virialized halos of a given mass, and therefore scale, will form, the excursion set aims at calculating the fraction of collapsed objects considering all points in space by analyzing the density field at different smoothing scales.

Working out the connection between them has been the subject of several recent papers. In [66], building on earlier work by [65], it was showed that correlation functions of discrete density peaks can be computed using an effective (i.e. which does not involve measurable counts-in-cells quantities) generalized bias expansion in which all the bias parameters, including those of the nonlocal terms\(^1\), can be computed from a peak-background split. In parallel, [80] demonstrated how the peak formalism can be combined with excursion set theory into one model. The excursion set peaks (ESP hereafter) model therefore arises by imposing that peaks of the density field at a given smoothing scale satisfy a first crossing condition, in such a way as to have a first crossing distribution for peaks.

Several details have been worked out in this model, which we present in this chapter. One important step is the study of a realistic barrier for collapse in this framework: in [80] an ellipsoidal barrier in the style of the one derived in [81] was generalized to an excursion set with peaks. Subsequently, [82] introduced a moving barrier with a stochastic component lognormally distributed with mean and variance fitted to simulations. This constituted a good improvement in the predictivity of the ESP halo mass function, as we show in Figure 2.

The ESP biasing scheme as an effective bias expansion was fully developed in a number of papers, starting from the early work of [65] for peaks and going through subsequent refinements in [66, 83, 67]. We review this scheme in this chapter summarizing results from this works.

Moreover, in [84], following a suggestion from [85], first- and second-order local bias factors were first measured from N-body simulations using one-point statistics, rather

---

1 To facilitate the comparison with other studies, we will call nonlocal terms all contributions to Lagrangian clustering that are not of the form $\delta^n(x)$, where $\delta(x)$ is the linear mass density field.
than the more conventional technique of studying the cross-correlation between halo-matter over matter-matter power spectra. We look at this technique in detail, showing a generalization of the method to measure non-local second-order bias parameters from N-body simulations.

**Plan of the chapter** In this chapter, we focus on this ESP framework. In Sec. §2.1 we review recent work analyzing how to build this model and its general features, providing a comparison with numerical N-body simulation of the ESP halo mass function. In Sec. §2.2, we introduce the ESP biasing scheme summarizing a series of recent results and in Sec. §2.3 we show a novel method for measuring non-local bias parameters in N-body simulations with one-point statistics and compare it to ESP predictions. In Sec. §2.4 we conclude.

### 2.1 An excursion set for peaks

#### 2.1.1 Making the connection

The goal is to apply the excursion set ansatz to the statistics of peaks of the density field. This amounts to impose to select only those peaks which are found at a certain scale \( R \) which have a smaller height on the next larger smoothing scale, that is,

\[
\frac{B(R_s)}{\sigma_{0s}} < \nu \leq \frac{B(R_s)}{\sigma_{0s}} + \frac{\Delta R_s}{\sigma_{0s}} \left( \frac{dB}{dR_s} \frac{d\delta}{dR_s} \right),
\]

where we just adapted Eq. (47) to the variable \( R_s \), with the condition that \( \delta' - B' \leq 0 \), where now the prime indicates derivative with respect to \( R_s \). Combining Eq. (35) and Eq. (48) we get [83]

\[
\bar{n}_{\text{ESP}}(\nu, R_s) \Delta R_s = \frac{3^{3/2}}{R_s} \int d^6 \zeta \int d^3 \eta \int_{-\infty}^{B'} d\delta' \int_{B}^{B-\Delta R_s(\delta'-B')} \frac{d\delta}{\sigma_{0s}} |\det \zeta| \delta_D(\eta) \theta_H(\lambda_3) P_1(\omega) \]

\[
= \int d^6 \zeta \int d^3 \eta \int_0^{\infty} d\xi \frac{\xi}{\sigma_{0s}} n_{pk}(y) P_1(\omega)
\]

where we have defined the variable \( \xi = B' - \delta' \) and the 11-dimensional vector of variables \( \omega = (y, \xi) \), being \( y \) as in Eq. (36). Using a similar approach to Eq. (40), we can calculate the ESP multiplicity function to be

\[
f_{\text{ESP}}(\nu) = \frac{M}{\rho} \bar{n}_{\text{ESP}} \frac{dR_s}{d\nu}
\]

\[
= -\frac{e^{-\nu^2/2}}{\sqrt{2\pi}} \frac{1}{\sigma_{0s}^2} \frac{V}{V_s} G_1^{(1)}(\gamma_1, \gamma_1 \nu)
\]
where \( V = m/\rho \), \( \sigma'_{0s} = d\sigma_{0s}/dR_s \) and \( \varphi^{(1)}_1(x) \) is defined in appendix 8.1. Hence, the corresponding ESP discrete number density can be written as
\[
n_{\text{ESP}}(\omega) = -\frac{\zeta}{\sigma'_{0s}v} \theta_H(\xi) n_{pk}(y). \tag{58}
\]

### 2.1.2 A more realistic picture: ESP with moving barriers

Up to now, we have limited our examples to the case of a constant, flat barrier \( B(s) = \delta_{sc} \), which is related to the spherical collapse model we discussed in Sec. §1.1. However, it has been long known that the collapse is not spherical in general, especially when the region forming the virialized object is not very massive. This was first argued in \([81]\), where they predicted that the critical density for collapse may be distributed around a mean value which increases with decreasing halo mass. Analyses of N-body simulations have confirmed this prediction and showed the scatter around the mean barrier is always significant \([86, 87, 61]\).

Several implementations of moving and stochastic barriers exist in the literature, ranging from direct implementations of triaxial collapse \([88, 81]\), multidimensional excursion set approaches \([74, 89, 90, 91]\) to the diffusive drifting barrier approach of \([92, 93, 94]\). Since the stochasticity induced by triaxial collapse is somewhat cumbersome to implement in analytic models of halo collapse (see e.g. for a tentative implementation with the peak constraint \([95]\)), we consider a simple approximation calibrated with numerical simulations. Namely, the square-root stochastic barrier
\[
B(R_s) = \delta_{sc} + \beta \sigma_{0s}, \tag{59}
\]
wherein the stochastic variable \( \beta \) closely follows a lognormal distribution, furnishes a good description of the critical collapse threshold as a function of halo mass \([87]\). In \([82]\), this result was interpreted as follows: each halo “sees” a moving barrier \( B = \delta_{sc} + \beta \sigma_{0s} \) with a value of \( \beta \) drawn from a lognormal distribution. Here however, we will assume that each halo “sees” a constant (flat) barrier, whose height varies from halo to halo. Therefore, we will implement the first-crossing condition simply as
\[
B(R_s) < \delta_s < B(R_s) + \mu \Delta R_s, \tag{60}
\]
where \( \mu = -d\delta_s/dR_s \). Consequently, the variable \( \mu \) will satisfy the constraint \( \mu > 0 \) rather than \( \mu > B' \).

With the aforementioned modifications, the excursion set peak multiplicity function reads
\[
f_{\text{ESP}}(\nu_c) = \left( \frac{V}{V_\ast} \right) \frac{1}{\gamma_{\nu\mu} \nu_c} \int_0^\infty d\beta \, p(\beta) \times \int_0^\infty d\mu \, \mu \int_0^\infty dJ_1 \, F_1(J_1) \mathcal{N}(\nu_c + \beta, J_1, \mu), \tag{61}
\]
where \( p(\beta) \) is a log-normal distribution, for which we take \( \langle \beta \rangle = 0.5 \) and \( \text{Var}(\beta) = 0.25 \) as in \cite{82}, and we define the cross correlations \( \gamma_{\nu\mu} = \langle \nu \mu \rangle \) and \( \gamma_{1\mu} = \langle J_1 \mu \rangle \). We can now apply Bayes’ theorem and write \( \mathcal{N}(\nu, J_1, \mu) = \mathcal{N}(\nu, J_1) \mathcal{N}(\mu | \nu, J_1) \). The integral over \( \mu \),

\[
\int_0^\infty d\mu \, \mu \mathcal{N}(\mu | \nu, u),
\]

is the same as in \cite{78} and, therefore, is equal to

\[
\bar{\mu} \left[ \frac{1 + \text{erf}(\bar{\mu} / \sqrt{2\Sigma})}{2} + \frac{\Sigma}{\sqrt{2\pi}} e^{-\bar{\mu}^2 / 2\Sigma^2} \right],
\]

where

\[
\bar{\mu} = J_1 \left( \frac{\gamma_{1\mu} - \gamma_{\nu\mu}}{1 - \gamma_1^2} \right) + (\nu + \beta) \left( \frac{\gamma_{\nu\mu} - \gamma_{1\nu} \gamma_{1\mu}}{1 - \gamma_1^2} \right)
\]

\[
\Sigma^2 = \Delta_0^2 - \frac{\gamma_{\nu\mu}^2 - 2\gamma_{\nu\nu} \gamma_{1\mu} + \gamma_{1\mu}^2}{1 - \gamma_1^2}.
\]

Substituting this expression into Eq. \( (61) \) and performing numerically the integrals over \( J_1 \) and \( \beta \), we obtain an analytic prediction for the halo mass function without any free parameter. Our ESP mass function differs at most by 2 - 3\% over the mass range \( 10^{11} - 10^{15} \, M_\odot / h \) from that obtained with the prescription of \cite{82}. Likewise, the linear and quadratic local bias parameters are barely affected by our modifications.

### 2.1.3 Comparison with numerical simulations

To test the validity of this approach, we show a comparison of the ESP mass function with that of halos extracted from N-body simulations. For this purpose, we ran a series of N-body simulations evolving \( 1024^3 \) particles in periodic cubic boxes of size 1500 and 250 \( h^{-1} \text{Mpc} \). The particle mass thus is \( 2.37 \times 10^{11} \) and \( 1.10 \times 10^9 \, M_\odot / h \), respectively. The transfer function was computed with CAMB [97] assuming parameter values consistent with those inferred by WMAP7 [31]: a flat \( \Lambda \text{CDM} \) cosmology with \( h = 0.704, \Omega_m = 0.272, \Omega_{\Lambda} = 0.0455, n_s = 0.967 \) and a normalisation amplitude \( \sigma_8 = 0.81 \). Initial conditions were laid down at redshift \( z = 99 \) with an initial particle displacement computed at 2nd order in Lagrangian perturbation theory with 2LPTic [98]. The simulations were run using the N-body code GADGET-2 [99] while the halos were identified with the spherical overdensity (SO) halo finder AHF [100] assuming an overdensity threshold \( \Delta = 200 \) constant throughout redshift.
2.1 An Excursion Set for Peaks

Figure 2: Halo mass function measured from N-body simulation at redshift $z = 0$ (left panels) and $z = 1$ (right panels) with different box sizes as indicated in the figures. The error bars are Poisson. The data is compared to the theoretical prediction Eq. (61) based on the ESP formalism and the fitting formula of [96]. We also show the fractional deviation of the [96] and the measured halo mass function relative to our theoretical prediction.

In Fig. (2), we compare the simulated halo mass function to the ESP prediction at redshift $z = 0$ and 1. The latter can be straightforwardly obtained from the multiplicity function $f_{\text{ESP}}(\nu_c)$ as

$$
\frac{d n_{\text{ESP}}}{d \log M} = \frac{\bar{\rho}}{M} \nu_c f_{\text{ESP}}(\nu_c, R_s) \frac{d \log \nu_c}{d \log M} \quad (66)
$$

where we used the fact that $\gamma_{\nu \mu} = \sigma_0'$ to obtain the second equality. The ESP prediction agrees with the simulations at the 10% level or better from $10^{14} \, M_\odot/h$ down to a halo mass $10^{11} \, M_\odot/h$, where the correspondence between virialized halos and initial density peaks should be rather vague. The abundance of very rare clusters with $M > 10^{14} \, M_\odot/h$ is difficult to predict because of exponential sensitivity to $\delta_{sc}$. In this respect, it might be more appropriate to work with a critical linear density $\delta_{sc} \approx 1.60$ if halos are defined with a fixed nonlinear threshold $\Delta = 200$ relative to the mean density (see e.g. for a discussion [101, 102]).
2.2 THE ESP BIASING SCHEME

The peaks approach is based on the statistics of the smoothed density field, as well as its first and second derivatives. As a result, the rms values of these quantities play an important role. In the theory, these rms values are related to integrals over the initial power spectrum. However, they depend on the form of the smoothing filter, and although it is more natural to define the density field using a TopHat filter, since it is more directly related to the physics of gravitational collapse and it is usually employed in numerical N-body simulations to define halo masses, some of the integrals do not converge with this choice of filtering. We will therefore use the following filtering scheme

$$\nu(x) = \frac{1}{\sigma_{0T}} \delta_T(x)$$

$$J_1(x) = -\frac{1}{\sigma_{2G}} \nabla^2 \delta_G(x)$$

$$\mu(x) = -\frac{d\delta_T}{dR_T}(x) ,$$

where the subscripts $T$ and $G$ indicate TopHat and Gaussian filtering, respectively. Note that, while $\nu$ and $J_1$ have unit variance, $\mu$ is not normalized. We will use the notation $\langle \mu^2 \rangle = \Delta_0^2$ in what follows.

Cross-correlations among these three variables are useful and will be denoted as

$$\langle \nu J_1 \rangle = \gamma_1 = \frac{\sigma_{1X}^2}{\sigma_{0T} \sigma_{2G}}$$

$$\langle \nu \mu \rangle = \gamma_{\nu \mu} = \frac{1}{\sigma_{0T}} \int_0^{\infty} \frac{dk}{2\pi^2} P(k) k^2 W_T(kR_T) \frac{dW_T(kR_T)}{dR_T}$$

$$\langle J_1 \mu \rangle = \gamma_{J_1 \mu} = \frac{1}{\sigma_{2G}} \int_0^{\infty} \frac{dk}{2\pi^2} P(k) k^4 W_G(kR_G) \frac{dW_T(kR_T)}{dR_T} .$$

The first-order, mixed spectral moment $\sigma_{1X}$ is

$$\sigma_{1X}^2 = \frac{1}{2\pi^2} \int dk P(k) k^4 W_T(kR_T) W_G(kR_G) ,$$

i.e. one filter is tophat and the other Gaussian.

The bias coefficients of ESP are given by (1-point) ensemble average of derivative operators, or, equivalently, of orthonormal polynomials. The type of these polynomials depend on the nature of the variables. Namely, the scalars $\nu$, $\mu$ and $J_1$ are associated with Hermite polynomials while the chi-square variable $\eta^2$ corresponds to Laguerre polynomials.

---

2 For more technical details on this technical problem, the reader may refer to [82].
polynomials $L^{(1/2)}_q(3\eta^2/2)$. The jointly distributed scalars $(j_2, j_3)$ generate the polynomials $F_{l\ell m}(5J_2, J_3)$ [66, 67], being $J_3(x) = \frac{2}{3} \text{tr}[(\zeta_{ij} + 1/3 \delta_{ij})J_1^2]$.

This structure arises owing to the fact that we identify bias parameters with the averaged response of the halo mass function to the change of a given variable of the problem. For the variables $(\nu, J_1, \mu)$, the bias factors are

$$
\sigma^2_{00} \sigma^2_{22} \sigma^2_{\mu} b_{ijk} = \frac{1}{\bar{\eta}_{ESP}} \left< \frac{\partial^{i+j+k} \eta_{ESP}}{\partial u^i \partial \bar{u}^j \partial \mu^k} \right> = \frac{1}{\bar{\eta}_{ESP}} \int \! d\mathbf{y} \, n_{ESP}(\mathbf{y}) \, H_{ijk}(\nu, J_1, \mu) \, P(\mathbf{y}) .
$$  \hspace{1cm} (72)

Here, $H_{ijk}$ are tri-variate Hermite polynomials. Since these derivatives can be re-summed into a shift operators, the $b_{ijk}$ can also be written as

$$
\sigma^2_{00} \sigma^2_{22} \sigma^2_{\mu} b_{ijk} = \frac{1}{\bar{\eta}_{ESP}} \frac{\partial^{i+j+k} \eta_{ESP}}{\partial \nu^i \partial \bar{J}_1^j \partial \mu^k} ,
$$  \hspace{1cm} (73)

where it is understood that the long-mode perturbations $\nu$, $u_l$, and $\mu_l$ shift the mean of the 1-point PDF $\mathcal{N}(\nu, J_1, \mu)$, where $\mathcal{N}$ is a Normal distribution. Similar expressions arise for $\eta^2$, $J_2$ and $J_3$. Namely, the bias factors $\chi_k$ associated with $\eta^2$ are defined as

$$
\sigma^2_{11} \chi_q = \frac{(-1)^q}{\bar{\eta}_{ESP}} \int \! d\mathbf{y} \, n_{ESP}(\mathbf{y}) \, L^{(1/2)}_q(3\eta^2/2) P(\mathbf{y})
$$

$$
= \frac{1}{\bar{\eta}_{ESP} \bar{\eta}(\eta^2)^q} \frac{\partial^{i+j+k} \eta_{ESP}}{\partial \nu^i \partial \bar{J}_1^j \partial \mu^k} ,
$$  \hspace{1cm} (74)

where, owing to rotational symmetry, derivatives are taken w.r.t. the modulus squared $\eta_i^2 = \eta_l \cdot \eta_l$ of the long-wavelength perturbation $\eta_l = (\eta_{12}, \eta_{13}, \eta_{23})$. As shown in [66], the long mode perturbation $\eta_l$ shift the chi-square distribution for $\eta(x)$ into a non-central chi-square PDF that can be expanded in Laguerre polynomials $L^{(a)}_k$ with $k = 2(\alpha + 1) = 3$ degrees of freedom. Finally, the bias $\omega_{l\ell m}$ which correspond to $J_2$, $J_3$ are given by

$$
\sigma^2_{22} \omega_{l\ell m} = \frac{1}{\bar{\eta}_{ESP}} \int \! d\mathbf{y} \, n_{ESP}(\mathbf{y}) \, F_{l\ell m}(5J_2, J_3) P(\mathbf{y})
$$

$$
= \frac{1}{\bar{\eta}_{ESP}} \frac{\partial^{i+m} \eta_{ESP}}{\partial \nu^i \partial \bar{J}_2^m} ,
$$  \hspace{1cm} (75)

where $J_{2l}$ and $J_{3l}$ are long-wavelength perturbations to the second- and third-order invariant traces of $\zeta_{ij}$. The polynomials $F_{l\ell m}$ are given by

$$
F_{l\ell m}(5J_2, J_3) = (-1)^\ell \sqrt{\frac{\Gamma(5/2)}{2^{3m} \Gamma(3m + 5/2)}} L^{(3m+3/2)}_l(s/2) P_m(x_3) ,
$$  \hspace{1cm} (76)
where \( s = 5f_2 \) and \( P_n(x) \) are Legendre polynomials. Again, the factor of \((-1)^L\) ensures that the term with highest power always has positive sign. In general however, the bias coefficients do not factorise into a product of \( b_{ijk} \), \( \chi_q \) and \( \omega_{lmr} \), and take the generic form

\[
c_{ijkm} = \frac{\left\langle n_{\text{ESP}} H_{ijl}(v, f, \mu) L_{ijl}^{-1/2}(3\eta^2/2) F_{lm}(5f_2, f_3) \right\rangle}{\sigma_0^2 \sigma_1^{2L+2L+3m} \sigma_L^{2L+2L+3m} n_{\text{ESP}}.}
\] (77)

We refer the reader to [67] for details about the construction of perturbative bias expansions.

2.2.1 The effective bias expansion

The bias factors \( c_{ijkm} \) are the coefficients of the ESP Lagrangian perturbative bias expansion \( \delta_L^{\text{ESP}}(x) \). For the sake of completeness, \( \delta_L^{\text{ESP}}(x) \) is explicitly given in Appendix \( \S9 \) up to third order in the linear density field and its derivatives. Most importantly, the bias coefficients \( c_{ijkm} \) also multiply orthonormal polynomials [66, 83, 67]. This series can be used to compute the whole hierarchy of \( N \)-point correlation functions of ESP in Lagrangian space. Furthermore, it is also valid in the presence of (weak) primordial non-Gaussianity, a subject we deal with later on.

In Fourier space, this perturbative bias expansion takes the compact form

\[
\delta_L^{\text{ESP}}(x, z_s) = \sum_{n=1}^{\infty} \frac{1}{n!} \int \frac{d^3k_1}{(2\pi)^3} \cdots \frac{d^3k_n}{(2\pi)^3} c_n^{\text{L}}(k_1, \ldots , k_n; z_s) \times \delta(k_1, z_s) \cdots \delta(k_n, z_s) e^{i(k_1 + \cdots + k_n) \cdot x},
\] (78)

where the \( n \)-order Lagrangian bias functions \( c_n^{\text{L}}(k_1, \ldots , k_n; z_s) \) sum over all the possible combinations of rotational invariants that can be generated from \( n \) powers of the linear density field \( \delta_L \) and/or its derivatives. We use the notation \( c_n^{\text{L}} \) to emphasise that the Fourier space ESP bias factors correspond exactly to the renormalised Lagrangian bias functions of the “integrated perturbation theory” (iPT) [103, 104],

\[
\left\langle \frac{\delta^n}{\delta \delta_L^z} \delta_L^{\text{ESP}}(k, z_s) \cdots \delta_L^{\text{ESP}}(k_n, z_s) \right\rangle = (2\pi)^3 \delta^3(k - k_{1 \cdots n}) c_n^{\text{L}}(k_1, \ldots , k_n; z_s).
\] (79)

Here, \( \delta_L^{\text{ESP}}(k, z_s) \) is the Fourier transform of the effective overabundance \( \delta_L^{\text{ESP}}(x, z_s) \) of the biased tracers (the excursion set peaks) in Lagrangian space. We have momentarily re-introduced the explicit redshift dependence to remind the reader that these formulae are evaluated at the virialization redshift \( z_s \). We also note the important caveat that the ESP constraint involves variables other than \( \delta_L \), so that all the Lagrangian bias functions \( c_n^{\text{L}} \) are scale-dependent.

The first order ESP Lagrangian bias function is

\[
c_1^{\text{L}}(k) = \left( b_{100} + b_{010} k^2 - b_{001} \frac{d \ln \tilde{W}_R}{dR} \right) \tilde{W}_R(k)
\] (80)
while, at second-order, we have

\[
\begin{align*}
\mathcal{C}_2^L(k_1,k_2) &= \left\{b_{200} + b_{110} \left( k_1^2 + k_2^2 \right) + b_{020} k_1^2 k_2^2 + b_{002} \right. \\
&\left. - b_{101} \left[ \frac{d \ln \tilde{W}_R}{d R} (k_1) + \frac{d \ln \tilde{W}_R}{d R} (k_2) \right] - b_{011} \left[ k_1^2 \frac{d \ln \tilde{W}_R}{d R} (k_2) + k_2^2 \frac{d \ln \tilde{W}_R}{d R} (k_1) \right] \right\} \tilde{W}_R(k_1) \tilde{W}_R(k_2) \quad .
\end{align*}
\] (81)

The third-order ESP Lagrangian bias function is spelt out in Appendix \S9. At this point, we should stress that, while we have assumed a unique smoothing kernel \(\tilde{W}\) for conciseness, the ESP implementation of \[82\] which we adopt here involves two different windows: a tophat filter for \(\nu\) and \(\mu\), and a Gaussian filter for \(\eta_i\) and \(\zeta_{ij}\). Therefore, the overall multiplicative factor of \(\tilde{W}\) should be replaced by \(\tilde{W}_T\) and \(\tilde{W}_G\) wherever appropriate. Note, however, that it should be possible to write down a model with a unique smoothing kernel which, if required, can be measured directly from simulations \[86, 105, 106\].

### 2.3 Measuring Nonlocal Lagrangian Peak Bias

In this Section, we demonstrate that the bias factors \(\chi_i\) and \(\omega_{j0}\) can be measured with a one-point statistics. We test our method on density peaks of a Gaussian random field before applying it to dark matter halos.

#### 2.3.1 Bias factors \(b_{ijk}\): Hermite polynomials

[85] showed that the bias factors of discrete tracers (relative to the mass density \(\delta\)) can be computed from one-point measurements rather than computationally more expensive \(n\)-point correlations. Their idea was implemented by \[82, 84\] to halos extracted from N-body simulations in order to test the predictions of the ESP formalism. Namely, halos were traced back to their “proto-halo” patch (since one is interested in measuring Lagrangian biases) in the initial conditions, the linear density field was smoothed on some “large scale” \(R_l\) and the quantity \(H_n(\nu_l = \delta_l/\sigma_0)\) was computed (for \(n = 1, 2\) only) at the location of each proto-halo. The average of \(H_n(\nu_l)\) over all proto-halos reads

\[
\frac{1}{N} \sum_{i=1}^{N} H_n(\nu_i) = \int_{-\infty}^{+\infty} d \nu_l N(\nu_l) \langle 1 + \delta_{hi} | \nu_i \rangle H_n(\nu_i) \quad ,
\] (82)

where \(\delta_{hi}\) is the overdensity of proto-halos. This expression assumes that the first-crossing condition can be implemented through a constraint of the form Eq.(61), so
that $P(\nu_l)$ is well approximated by a Gaussian [85]. For the ESP peaks considered here, this ensemble average reads

$$\frac{1}{n_{\text{ESP}}} \int_{-\infty}^{+\infty} d\nu_l N(\nu_l) \langle n_{\text{ESP}} | \nu_l \rangle H_{\nu}(\nu_l) = \frac{1}{n_{\text{ESP}}} \int d\nu d^{11} w n_{\text{ESP}}(w) (-\epsilon_{\nu})^n$$

$$\times \left( \frac{\partial}{\partial \nu} + \frac{\epsilon_{J_1}}{\epsilon_{\nu}} \frac{\partial}{\partial J_1} + \frac{\epsilon_{\mu}}{\epsilon_{\nu}} \frac{\partial}{\partial \mu} \right) H_{\nu}(\nu_l) P_1(w).$$

Here, $\epsilon_X$ denotes the cross-correlation between $\nu_l$ and the variables $X = (\nu, J_1, \mu)$ defined at the halo smoothing scale. The right-hand side reduces to a sum of $n$th-order bias factors $b_{ijk}$ weighted by products of $\epsilon_{\nu}$, $\epsilon_{J_1}$, and $\epsilon_{\mu}$. Relations between bias factors of a given order (which arise owing to their close connection with Hermite polynomials, see e.g.[85]) can then be used to extract a measurement of each $b_{ijk}$.

Before we generalize this approach to the chi-squared bias factors $\chi_i$ and $\omega_j$, we emphasize that, in this cross-correlation approach, the smoothing scale $R_l$ can take any value as long as it is distinct from the halo smoothing scale. [84] chose $R_l \gg R_s$ in the spirit of the peak-background split but this requirement is, in fact, not necessary as long as the correlation between the two scales is taken into account. In any case, we will stick with the notation $R_l$ for convenience.
2.3.2 Bias factors $\chi_j$ and $\omega_j$: Laguerre polynomials

The approach presented above can be generalised to $\chi^2$ distributions. The main difference is the appearance of Laguerre polynomials $L_n^{(a)}$. Consider for instance the $\chi^2$-quantity $3\eta^2$ smoothed at the scale $R_l$, i.e. $3\eta_l^2$. In analogy with Eq. (82), the ensemble average of $L_n^{(1/2)}(3\eta_l^2)$ at the peak positions is

$$\frac{1}{N} \sum_{i=1}^{N} L_n^{(1/2)}(3\eta_i^2) = \int_0^\infty d(3\eta_i^2) \chi_n^2(3\eta_i^2) \left( \frac{3\eta_i^2}{2} \right)^{1/2} \times \langle 1 + \delta_l | 3\eta_i^2 \rangle L_n^{(1/2)}(3\eta_i^2) = \int_0^\infty d(3\eta_i^2) \chi_n^2(3\eta_i^2) \left( \frac{3\eta_i^2}{2} \right)^{1/2} \times \langle 1 + \delta_l | 3\eta_i^2 \rangle L_n^{(1/2)}(3\eta_i^2) \right) .$$  

The conditional average $\langle 1 + \delta_l | 3\eta_l^2 \rangle$ reads

$$\langle 1 + \delta_l | 3\eta_l^2 \rangle = \frac{1}{n_{ESP}} \int d^4 w n_{ESP}(w) P_l(w | 3\eta_l^2) = \frac{1}{n_{ESP}} \int dJ d\nu d\mu N(\nu, j, \mu) \times \int d(3\eta_l^2) \chi_n^2(3\eta_l^2) \chi_3^2(3\eta_l^2) \chi_3^2(3\eta_l^2) \int d(3\eta_l^2) \chi_n^2(3\eta_l^2) \times \int d(\text{angles}) P(\text{angles}) n_{ESP}(w) .$$  

We substitute this relation into Eq. (84) and begin with the integration over the variable $3\eta_l^2$.

We use the following relation (which can be inferred from Eq. (7.414) of [107])

$$\int_0^\infty dx e^{-x} x^{j+n} L_n^{(a)}(x) = \frac{(-1)^n j! \Gamma(j + \alpha + 1)}{n!} \left( \frac{3\eta_l^2}{2} \right)^{\alpha} \sum_{j=0}^{\infty} \frac{j!}{(j - n)!} \left( \frac{1 - \epsilon^2}{1 - \epsilon^2} \right) L_j^{(1/2)} \left[ \frac{3\eta_l^2}{2(1 - \epsilon^2)} \right] .$$  

For simplicity, let us consider the case $n = 0, 1$ solely. For $n = 0$, the sum simplifies to

$$\sum_{j=0}^{\infty} \left( \frac{-\epsilon^2}{1 - \epsilon^2} \right)^j L_j^{(1/2)} \left[ \frac{3\eta_l^2}{2(1 - \epsilon^2)} \right] = (1 - \epsilon^2)^{3/2} \exp \left( \frac{\epsilon^2}{1 - \epsilon^2} 3\eta_l^2 \right) .$$  

51
and the integral Eq.(87) \( L_n^{(1/2)}(3\eta^2/2) \equiv 1 \) is trivially equal to \( \chi_3^2(3\eta^2) \) (as it should be, since we are essentially marginalizing over \( 3\eta^2 \)).

For \( n \geq 1 \), the sum can be evaluated upon taking suitable derivatives of the right-hand side of Eq.(88), which indeed is a generating function for the Laguerre polynomials \( L_n^{(1/2)} \). For \( n = 1 \), a little algebra leads to

\[
\sum_{j=0}^{\infty} j \left( \frac{-e^2}{1-e^2} \right)^{j-1} L_j^{(1/2)} \left[ \frac{3\eta^2}{2(1-e^2)} \right] = (1 - e^2)^{5/2} L_1^{(1/2)} \left( \frac{3\eta^2}{2} \right) \exp \left[ \left( \frac{e^2}{1-e^2} \right) \frac{3\eta^2}{2} \right].
\]

Hence, Eq. (87) with \( n = 1 \) equals \( e^2 L_1^{(1/2)}(3\eta^2/2) \chi_3^2(3\eta^2) \). Performing the remaining integrals over \( \nu, J_l, \mu \) and \( 5\zeta^2 \) (the integral over the angles is trivially unity) and taking into account the ESP peak constraint through the multiplicative factor \( n_{\text{ESP}}(w) \), Eq.(84) simplifies to

\[
\int_0^\infty d(3\eta^2) \chi_3^2(3\eta^2) \langle 1 + \delta_1 | 3\eta^2 \rangle L_1^{(1/2)} \left( \frac{3\eta^2}{2} \right) = -e^2 \sigma_{\chi_1}^2.
\]

For the variable \( 3\eta^2 \), the cross-correlation coefficient \( \epsilon \) is

\[
\epsilon^2 = \frac{\langle \eta^2 \eta_1^2 \rangle - \langle \eta^2 \rangle \langle \eta_1^2 \rangle}{\sqrt{\langle \langle \eta^4 \rangle - \langle \eta^2 \rangle^2 \rangle}} = \left( \frac{\sigma_{\chi x}^2}{\sigma_{\chi_1}^2} \right)^2,
\]

which we shall denote as \( \epsilon_1 \) in what follows. Furthermore,

\[
\sigma_{n\chi}^2 = \frac{1}{2\pi^2} \int_0^\infty dk k^{2(n+1)} P(k) W_G(kR_s) W_G(kR_l)
\]

designates the splitting of filtering scales, i.e. one filter is on scale \( R_s \) while the second is on scale \( R_l \). It should be noted that, unlike \( \sigma_{iX} \) defined in Eq.(74), both filtering kernels are Gaussian.

The derivation of the bias factors \( \omega_{J0} \) associated with the quadratic variable \( 5\zeta^2 \) proceeds analogously. In particular,

\[
\int_0^\infty d(5\zeta^2) \chi_5^2(5\zeta^2) \langle 1 + \delta_1 | 5\zeta^2 \rangle L_1^{(3/2)} \left( \frac{5\zeta^2}{2} \right) = -e^2 \sigma_{\zeta_1}^2 \omega_{10}.
\]

Here, the cross-correlation coefficient is \( \epsilon = \sigma_{\zeta_1}^2 / \sigma_{\chi_1}^2 \equiv \epsilon_2 \). Note that, in both cases, the cross-correlation coefficient drops very rapidly as \( R_l \) moves away from \( R_s \) for realistic CDM power spectra. In addition, one could in principle choose \( R_l < R_s \) (if there is enough numerical resolution) to measure \( \chi_1 \) and \( \omega_{J0} \).
Figure 4: Conditional probability distribution for the variables $3\eta_l^2$ (left panel) and $5\zeta_l^2$ (right panel) measured at the position of maxima of the linear density field smoothed with a Gaussian filter on scale $R = 5 \, h^{-1}\text{Mpc}$. *Left panel:* histograms indicate the results for $R_l = 10$, 15 and 20 $h^{-1}\text{Mpc}$, which leads to $\epsilon_1 = 0.71$, 0.44 and 0.29 as quoted on the figure. *Right panel:* histograms show the results for a fixed $R_l = 10 \, h^{-1}\text{Mpc}$ (which implies $\epsilon_2 = 0.57$) but several peak height intervals. In all cases, the solid curves are the theoretical prediction (see text) whereas the dashed (green) curves represents the unconditional distribution $\chi^2_k(y)$.

2.3.3 Test with numerical simulations

In this Section, we first validate our predictions based on peaks of Gaussian random fields with measurements extracted from random realizations of the Gaussian linear density field, and then move on to calculate $\chi_1$ and $\nu_{10}$ for $M > M_\star$ halos, where $M_\star$ is the characteristic mass of the halos as it was defined in §1.1.2, that is the mass corresponding to a peak height $\nu = 1$.

We generate random realizations of the Gaussian, linear density field with a power spectrum equal to that used to seed the N-body simulations described above. To take advantage of FFTs, we simulate the linear density field in periodic, cubic boxes of side $1000 \, h^{-1}\text{Mpc}$. The size of the mesh along each dimension is 1536. We smooth the density field on scale $R_s = 5 \, h^{-1}\text{Mpc}$ with a tophat filter and find the local maxima by comparing the density at each grid point with its 26 neighbouring values.
The excursion set peaks model

We then smooth the density field on the larger scales \( R_l = 10, 15 \) and \( 20 \) h\(^{-1}\)Mpc with a Gaussian filter and compute

\[
\eta_l^2 = \frac{1}{\sigma_l^2} (\nabla \delta_l)^2
\]

\[
\zeta^2_l = \frac{3}{2\sigma_l^2} \text{tr} \left[ \left( \partial_i \partial_j \rho_l - \frac{1}{3} \delta_{ij} \nabla^2 \delta_l \right)^2 \right].
\]

These density fields with derivatives sensitively depend on the smoothing scales used. To illustrate this we show in Fig. 3 sections of \( v_l, 3\eta_l^2 \) and \( 5\zeta_l^2 \). The sections, each of which of dimensions \( 200 \times 200 \) h\(^{-2}\)Mpc\(^2\), were generated at \( z = 99 \) with the same random seed.

The first row corresponds to \( R_s = 5 \) h\(^{-1}\)Mpc, whereas the second row displays results on the filtering scale \( R_l = 10 \) h\(^{-1}\)Mpc. We note that, for the normalized density field \( v_l \), an increase in the smoothing scale washes out the small scale features, but the large scale pattern remains. For the quadratic variable \( \eta_l^2 \) however, the resemblance between the features at the small and large filtering scale is tenuous. This is even worse for \( \zeta_l^2 \).

Compared to \( v_l \), the fields \( \eta_l^2 \) and \( \zeta_l^2 \) have one and two additional derivatives which give rise to an effective window function whose isotropic part is given by

\[
W_{\text{eff}}(k, R) = k^n e^{-(kr)^2/2},
\]

where \( n = 0, 1 \) and \( 2 \) are for \( v_l, \eta_l^2 \) and \( \zeta_l^2 \), respectively. For \( n = 0 \), the window becomes narrower as \( R_l \) increases, yet remains unity for wavenumbers \( k \lesssim 1/R_l \). \( W_{\text{eff}} \) reaches a maximum at \( \sqrt{n}/R \). Hence, for \( n = 1 \) and \( 2 \), \( W_{\text{eff}} \) selects predominantly wavemodes with \( k \sim 1/R \). Consequently, since in a Gaussian random field the wavemodes at different scales are uncorrelated, patterns in the fields \( \eta_l^2 \) and \( \zeta_l^2 \) can change drastically as \( R_l \) varies. This effect expected to be most significant for \( n = 2 \), i.e. \( \zeta_l^2 \).

For each local density maxima, we store the peak height \( v \) as well as the value of \( \eta_l^2 \) and \( \zeta^2_l \) at the peak position. The left panel of Fig.4 displays as histograms the resulting probability distribution \( P(3\eta_l^2|\text{pk}) \) for three different values of \( R_l = 10, 15 \) and \( 20 \) h\(^{-1}\)Mpc. The solid curves represent the theoretical prediction Eq.(301) with \( x = \langle 3\eta^2|\text{pk} \rangle = 0 \) and \( \epsilon_1 = 0.71, 0.44 \) and 0.29 (from the smallest to largest \( R_l \)) as was measured from the random realizations. The dashed curve is the unconditional \( \chi^2 \)-distribution with 3 degrees of freedom. The theory gives excellent agreement with the simulations. Note also that we did not find any evidence for a dependence on the peak height, as expected from the absence of a correlation between \( v \) and \( \eta_l^2 \). The right panel of Fig.4 shows results for \( \zeta_l^2 \). Here however, since the cross-correlation coefficient diminishes very quickly when \( R_l \) even slightly departs from \( R_s \), we show result for \( R_l = 10 \) h\(^{-1}\)Mpc only, which corresponds to \( \epsilon_2 = 0.57 \). In addition, because one should expect a dependence of the shape of the density profile around peaks to the peak height, we consider three different ranges of \( v \) as indicated on the figure. The solid curves indicate the theoretical prediction Eq.(301) with \( \epsilon_2 = 0.57 \) and \( x = \langle 5\zeta^2|\text{pk} \rangle \), where

\[
\langle 5\zeta^2|\text{pk} \rangle = -2\partial_x \ln \int_{v_{\text{min}}}^{v_{\text{max}}} dv \ G_0^{(a)}(\gamma_1, \gamma_1 v).
\]
Here, $G_0^{(a)}$ is the integral of $f(u, a)$ over all the allowed peak curvatures. The average $\langle 5\zeta^2 | \text{pk} \rangle$ increases with the peak height to reach 5 in the limit $\nu \to \infty$. The figure shows a clear deviation from the unconditional distribution $\chi_5^2(5\zeta^2)$ (shown as the dashed curve) and a dependence on $\nu$ consistent with theoretical predictions.

2.3.4 Dark matter halos

Having successfully tested the theory against numerical simulations of Gaussian peaks, we will now attempt to estimate the bias factors $\chi_1$ and $\omega_{10}$ associated with dark matter halos. For this purpose, we first trace back all dark matter particles belonging to virialized halos at redshift $z = 0$ to their initial position at $z = 99$. We then compute the center-of-mass positions of these Lagrangian regions and assume that they define the locations of proto-halos. We can now proceed as for the Gaussian peaks and compute $\nu$, $\eta_l^2$ and $J_2 \equiv \zeta_l^2$ at the position of proto-halos.

The quadratic bias factors $\chi_1$ and $\omega_{10}$ could be in principle computed analogously to [84], i.e. by stacking measurements of $\eta_l^2$ and $\zeta_l^2$ at the locations of proto-halos:

$$\sigma_{1s}^2 \hat{\chi}_1 = -\frac{1}{2N} \sum_{i=1}^{N} L_1^{(1/2)} \left( \frac{3\eta_i^2}{2} \right)$$

and

$$\sigma_{2s}^2 \hat{\omega}_{10} = -\frac{1}{2N} \sum_{i=1}^{N} L_1^{(3/2)} \left( \frac{5\zeta_i^2}{2} \right).$$

Here, $N$ is the number of halos, $s$ designates smoothing at the halo mass scale with a Gaussian filter $W_G$ on scale $R_G(R_T)$, whereas $l$ designates Gaussian smoothing at the large scale $R_l$. However, because the cross-correlation coefficient is fairly small unless $R_l$ is very close to $R_G$, we decided to compute $\chi_1$ and $\omega_{10}$ by fitting the probability distribution $P(3\eta_l^2|\text{halo})$ and $P(5\zeta_l^2|\text{halo})$ with the conditional $\chi^2$-distribution $\chi_k^2(y|x)$. Namely,

$$\sigma_{1s}^2 \hat{\chi}_1 = \frac{1}{2} \left( \langle 3\eta^2 | \text{halo} \rangle - 3 \right)$$

$$\sigma_{2s}^2 \hat{\omega}_{10} = \frac{1}{2} \left( \langle 5\zeta^2 | \text{halo} \rangle - 5 \right),$$

where $\langle 3\eta^2 | \text{halo} \rangle$ and $\langle 5\zeta^2 | \text{halo} \rangle$ are the best-fitting values obtained for $x$. We used measurements obtained at the smoothing scale $R_l = 10 \, h^{-1}\text{Mpc}$ only to maximize the signal.

To predict the value of $R_G$ given $R_T$, we followed [82] and assumed that $R_G(R_T)$ can be computed through the requirement that $\langle \delta_G | \delta_T \rangle = \delta_T$. This yields a prediction for the value of the cross-correlation coefficients $\epsilon_1$ and $\epsilon_2$ as a function of halo mass, which we can use as an input to $\chi_k^2(y|x)$ and only fit for $x$. However, we found that using
the excursion set peaks model

Figure 5: Conditional probability distribution for $3\eta_l^2$ (top panel) and $5\zeta_l^2$ (bottom panel) measured at the center-of-mass position of proto-halos. The filter is Gaussian with $R_l = 10 \ h^{-1}\text{Mpc}$. The various curves show the best-fit theoretical predictions for the halo mass bins considered here. Halo mass range are in unit of $10^{13} \ M_\odot /h$. Poisson errors are much smaller than the size of the data points and, therefore, do not show up on the figure.

the predicted $\epsilon_1$ leads to unphysical (negative) values for $x$ when one attempts to fit $P(3\eta_l^2|\text{halo})$. Therefore, we decided to proceed as follows:

1. Estimate both $\epsilon_1$ and $x = \langle 3\eta_l^2 | \text{halo} \rangle$ by fitting the model $\chi^2_3(y|x;\epsilon_1)$ to the measured $P(3\eta_l^2|\text{halo})$.

2. Compute $\epsilon_2$ assuming that the same $R_C$ enters the spectral moments.

3. Estimate $x = \langle 5\zeta_l^2 | \text{halo} \rangle$ by fitting the theoretical model $\chi^2_5(y|x;\epsilon_2)$ to the simulated $P(5\zeta_l^2|\text{halo})$.

We considered data in the range $0 < 3\eta_l^2 < 8$ and $0 < 5\zeta_l^2 < 12$ and gave equal weight to all the measurements (assuming Poisson errors does not affect our results significantly). Table 1 summarizes the best-fitting values obtained for four different
Table 1: Best-fit parameter values as a function of halo mass. The latter is in unit of $10^{13} \, M_\odot/h$. Note that we also list the values of $\epsilon_2$ even though it is not directly fitted to the data (see text for details).

<table>
<thead>
<tr>
<th>Halo mass</th>
<th>$\langle 3\eta^2 \rangle_{\text{halo}}$</th>
<th>$\epsilon_1$</th>
<th>$\langle 5\zeta^2 \rangle_{\text{halo}}$</th>
<th>$\epsilon_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$M &gt; 30$</td>
<td>0.71</td>
<td>0.80</td>
<td>2.98</td>
<td>(0.70)</td>
</tr>
<tr>
<td>$10 &lt; M &lt; 30$</td>
<td>1.24</td>
<td>0.66</td>
<td>4.49</td>
<td>(0.52)</td>
</tr>
<tr>
<td>$3 &lt; M &lt; 10$</td>
<td>1.62</td>
<td>0.54</td>
<td>5.82</td>
<td>(0.37)</td>
</tr>
<tr>
<td>$1 &lt; M &lt; 3$</td>
<td>1.94</td>
<td>0.49</td>
<td>6.12</td>
<td>(0.31)</td>
</tr>
</tbody>
</table>

halo bins spanning the mass range $10^{13} - 10^{15} \, M_\odot/h$, whereas the measured probability distributions together with the best-fit models are shown in Fig.5. The data is reasonably well described by a conditional $\chi^2$-distribution, but the fit is somewhat poorer when the cross-correlation coefficient is close to unity.

The second-order bias factors $\chi_1$ and $\omega_{10}$ of the dark matter halos at $z = 0$ can be readily computed from Eq.(100) using the best-fit values of $\langle 3\eta^2 \rangle_{\text{halo}}$ and $\langle 5\zeta^2 \rangle_{\text{halo}}$. The results are shown in Fig.6 as the data points. Error bars indicate the scatter among the various realizations and, therefore, likely strongly underestimate the true uncertainty. The dashed curves indicate the predictions of the ESP formalism. The measurements, albeit of the same magnitude as the theoretical predictions, quite disagree with expectations based on our ESP approach, especially $\omega_{10}$ which reverses sign as the halo mass drops below $10^{14} \, M_\odot/h$.

### 2.3.5 Interpretation of the measurements

To begin with, we note that, if halos were forming out of randomly distributed patches in the initial conditions, then both $\chi_1$ and $\omega_{10}$ would be zero since $\langle 3\eta^2 \rangle = 3$ and $\langle 5\zeta^2 \rangle = 5$ for random field points.

The measured dimensionless bias factor $\sigma_1^2 \chi_1$ is always negative, which indicates that halos collapse out of regions which have values of $\eta^2$ smaller than average. In our ESP approach, we assume that the center-of-mass position of proto-halos exactly coincides with that of a local density peak, so that $\sigma_1^2 \chi_1 \equiv -3/2$. However, simulations indicate that, while there is a good correspondence between proto-halos and linear density peaks, the center-of-mass of the former is somewhat offset relative to the peak position (see e.g. [108, 60]). To model this effect, we note that, if the proto-halo is at a distance $R$ from a peak, then the average value of $3\eta^2$ is $\langle 3\eta^2 \rangle(R) = \epsilon_1^2(R) (\langle 3\eta^2 \rangle_{\text{pk}} - 3)$ (in analogy with the fact that the average density at a distance $R$ from a position where $\delta \equiv \delta_{\text{sc}}$ is
The excursion set peaks model

Figure 6: The bias factors $\sigma_{1}^{2}\chi_{1}$ and $\sigma_{2}^{2}\omega_{10}$ of dark matter halos identified in the N-body simulations at $z = 0$ are shown as filled(green) circle and (blue) triangle, respectively. Error bars indicate the scatter among 6 realizations. The horizontal dashed (green) line at $-3/2$ and the dashed(blue) curve are the corresponding ESP predictions. The dotted (blue) curve is $\sigma_{2}^{2}\omega_{10}$ in a model where halos are allowed to collapse in filamentary-like structures. The solid curves are our final predictions, which take into account the offset between peak position and proto-halo center-of-mass (see text for details).

\[
\langle \delta \rangle(R) = \xi_\delta(R) \delta_{sc}). \quad \text{Assuming that the offset } R \text{ follows a Gaussian distribution, the halo bias factor is}
\]

\[
\sigma_{1s}^{2}\chi_{1} = -\frac{3}{2} \sqrt{\frac{2}{\pi}} \int_{0}^{\infty} \frac{dR}{\sigma} \left( \frac{R}{\sigma} \right)^{2} e^{-R^{2}/2\sigma^{2}} \chi_{1}^{2}(R). \quad (101)
\]

The rms variance $\sigma(M)$ of the offset distribution, which generally depends on the halo mass, can be constrained from our measurements of $\chi_{1}$ for dark matter halos. The best-fit powerlaw function,

\[
\sigma(M) = 2.50 \left( \frac{M}{10^{13} \, M_{\odot}/h} \right)^{0.063} h^{-1} \text{Mpc}, \quad (102)
\]

58
turns out to be a weak function of halo mass. In unit of the (tophat) Lagrangian halo radius, this translates into $\sigma / R_T \approx 0.79$ and $\approx 0.36$ for a halo mass $M = 10^{13}$ and $10^{14} \, M_\odot / h$, respectively. The resulting theoretical prediction is shown as the solid curve in Fig. 6 and agrees reasonably well with our data. This crude approximation demonstrates that an offset between the proto-halo center-of-mass and the peak position can have a large impact on the inferred value of $\chi_1$, since the latter is very sensitive to small-scale mass distribution.

Likewise, an offset between the proto-halo center-of-mass and the position of the linear density peak will also impact the measurement of $\omega_{10}$, yet cannot explain the observed sign reversal. In this regard, one should first remember that density peaks become increasingly spherical as $\nu \to \infty$. Nevertheless, while their mean ellipticity $\langle e \rangle$ and prolateness $\langle p \rangle$ converge towards zero in this limit, $\langle v \rangle = \langle u e \rangle$ approaches $1/5$ at fixed $u$ (see Eq.(7.7) of [47]). Hence, $\langle \xi^2 \rangle = \langle 3v^2 + w^2 \rangle$ does not tend towards zero but rather unity, like for random field points. Consequently, $\sigma^2 \omega_{10} \to 0$ in the limit $u \to \infty$.

Secondly, at any finite $\nu$, our ESP approach predicts that $\omega_{10}$ be negative because we have assumed that proto-halos only form near a density peak ($\lambda_3 > 0$, where $\lambda_1 \geq \lambda_2 \geq \lambda_3$ are the eigenvalues of $-\partial_i \partial_j \delta$). However, N-body simulations strongly suggest that a fraction of the proto-halos collapse along the ridges or filaments connecting two density maxima, and that this fraction increases with decreasing halo mass [60]. To qualitatively assess the impact of such primeval configurations on $\omega_{10}$, we extend the integration domain in the plane $(v, w)$ to include all the points with $\lambda_2 > 0$ and $\lambda_3 < 0$ (but still require that the curvature $u$ be positive). This way we not only consider density peaks, but also extrema that correspond to filamentary configurations.

The resulting curvature function $\tilde{G}^{(a)}(x)$ can be cast into the compact form

\[
\tilde{G}^{(a)}(x) = \frac{1}{\alpha^4} \left\{ e^{-\frac{5a^2}{10 \pi}} \left( \alpha x^2 - \frac{16}{5} \right) \right. \\
\left. + \frac{e^{-\frac{5a^2}{10 \pi}}}{\sqrt{10 \pi}} \left( 31 \alpha x^2 + \frac{32}{5} \right) + \frac{\sqrt{\alpha}}{2} (\alpha x^3 - 3x) \right. \\
\left. \times \left[ \text{Erf} \left( \sqrt{\frac{5a}{2}} x \right) + \text{Erf} \left( \sqrt{\frac{5a}{2}} x \right) - 1 \right] \right\},
\]

The dotted curve in Fig. 6 shows $\sigma^2 \omega_{10}$ when the filamentary configurations are included. While it agrees with the original ESP prediction at large halo mass, it reverses sign around $10^{14} \, M_\odot / h$ because, as the peak height decreases, configurations with $\lambda_3 < 0$ or, equivalently, large values of $\xi^2$ become more probable. The solid curve takes into account, in addition to filamentary configurations, an offset between the proto-halo and the peak position according to the simple prescription discussed above. This is our final prediction for $\sigma^2 \omega_{10}$. It is clearly at odds with the measurements, which strongly suggest that $\sigma^2 \omega_{10}$ can be very different from zero for $M \gtrsim 10^{13} \, M_\odot / h$. 
It is beyond the scope of this analysis to work out a detailed description of the measurements. Using a value of $R_G$ different than that obtained through the condition $\langle \delta_G | \delta_T \rangle = \delta_T$ has a large impact on the mass function, suggesting that it will be difficult to get a good fit of both the mass function and the bias factors $\chi_1$ and $\omega_{10}$. Before concluding however, we note that, if the Lagrangian clustering of halos also depends on $s_2(x) = s_{ij}(x)s^{ij}(x)$, where (in suitable units)

$$s_{ij}(x) = \partial_i \partial_j \phi(x) - \frac{1}{3} \delta_{ij} \delta(x),$$

(104)

where $\phi$ is the gravitational potential. We are therefore not measuring $\omega_{10}$ but some weighted and scale-dependent combination of both $\omega_{10}$ and the Lagrangian bias $\gamma_2$ associated with $s_2(x)$. Recent numerical work indeed suggests that $\gamma_2$ might be non-zero for massive halos [$109, 110, 90$]. In this regards, our approach can furnish a useful cross-check of these results since it can provide a measurement of $\gamma_2$ which is independent of the bispectrum.

2.4 Conclusions

We have described the framework of the ESP model and checked that the predicted halo mass function, from which all the bias factors can be derived, agrees well with the numerical data, Sec. §2.1. We have then shown how the biasing scheme is constructed as an effective bias expansion, Sec. §2.2, and that this bias can manifest itself not only in $n$-point statistics such as the power spectrum or bispectrum, but also in simpler one-point statistics, Sec. §2.3.

We took advantage of this to ascertain the importance of certain nonlocal Lagrangian bias factors independently of a 2-point measurement. We extended the cross-correlation technique of [85] to $\chi^2$-distributed variables, focusing on the quadratic terms $\eta^2(x)$ and $\zeta^2(x)$ which arise from the peak constraint and for which we have theoretical predictions. In principle however, our approach could be applied to measure the Lagrangian bias factor associated with any $\chi^2$-distributed variable such as the tidal shear for instance. We validated our method with peaks of Gaussian random field before applying it to a catalogue of dark matter halos with mass $M > 10^{13} \ M_\odot/h$. Including an offset between the proto-halo center-of-mass and the peak position in the modelling (motivated by the analysis of [60]), we were able to reproduce our measurements of the nonlocal bias $\sigma^2_1 \chi_1$. Our result $\chi_1 < 0$ is consistent with the findings of [60], who demonstrated that proto-halos with $M > 10^{13} \ M_\odot/h$ preferentially form near initial density peaks ($\chi_1 \equiv 0$ for a random distribution). However, we were unable to explain the measurements of $\sigma^2_1 \omega_{10}$, even with the additional assumption that a fraction of the halos collapse from filamentary-like structures rather than density peaks. We speculate that a dependence of the halo Lagrangian bias on $s_2(x)$ might be needed to explain this discrepancy.

The dependence on $\eta^2(x)$ induces a correction $-2\chi_1 (k_1 \cdot k_2)$ to the halo bias which, for collinear wavevectors $k_1$ and $k_2$ of wavenumber $0.1 \ h^{-1} \text{Mpc}$, is $\Delta b \approx 0.02 (0.05)$ and

60
\[ \approx 0.30 \ (0.88) \] for halos of mass \( M = 10^{13} \) and \( 10^{14} \) \( M_{\odot}/h \) at redshift \( z = 0 \) (z=1), respectively. Relative to the evolved, linear halo bias \( b_E^1 \equiv 1 + b_{100} \), the fractional correction is \( \Delta b / b_E^1 \sim 2\% \) and \( \sim 15\% \) for the same low and high halo mass in the redshift range \( 0 < z < 1 \). Hence, this correction can safely be ignored for \( M = 10^{13} \ h^{-1} \text{Mpc} \), but it could become relevant at larger halo masses.
While the existence of a spatial bias between the halos hosting galaxies and the underlying dark matter (DM) distribution has been firmly established since the pioneering work of [45], the presence of a halo velocity bias is still being debated. Clearly, whereas galaxy velocities can be physically biased (i.e. on an object-by-object basis) owing to differences between the dark matter and baryon velocity fields (see, e.g., Refs. [111, 112]), by construction halos locally flow with the dark matter (in Einstein’s theory of gravity), so that a halo velocity bias can only arise statistically. Namely, it must be the statistical manifestation of a selection effect, which should naturally arise since virialized structures preferentially trace overdense regions of the Universe [55, 46].

A step forward in the understanding of halo bias has been recently made in Ref. [105] where, through N-body simulations, the authors have measured a halo velocity bias

$$v_h(k, t) = b_v(k) v_{dm}(k, t), \quad b_v(k) = (1 - \frac{R_v^2 k^2}{2}),$$  \hspace{1cm} (105)$$

which appears to remain constant throughout the cosmic time $t$ until virialization. Here, $R_v$ is the typical scale of the halo velocity bias

$$R_v^2 = \frac{\sigma_v^2}{\sigma_f^2},$$  \hspace{1cm} (106)$$

where the functions $\sigma_j$ are defined as in Eq. (33). This statistical effect is consistent with the relative suppression of the halo velocity divergence power spectrum at late time subsequently reported in Ref. [113], although this type of measurement is more prone to systematics arising from sparse sampling (see e.g. Refs. [114, 61, 115]).

Whereas the finding of Ref. [105] is in full agreement with the peak model, which indeed predicts the existence of a linear, statistical halo velocity bias which remains constant with time [116, 117, 65], it seems at odds with the prediction based on the coupled-fluids approximation for the coevolution of DM and halos [110]. The latter is widely used to compute the time evolution of bias [118, 119, 120] (see also Chapter §4) and is based on the idea of following the evolution over cosmic time and in Eulerian
space of the halo progenitors - the so-called proto-halos - until their virialization. While their shapes and topology change as a function of time (smaller substructures gradually merge to form the final halo), their centre of mass moves along a well-defined trajectory determined by the surrounding mass density field. Therefore, unlike virialized halos that experience merging, by construction proto-halos always preserve their identity. Their total number is therefore conserved over time, and one can write a continuity equation and an Euler equation for their number density and velocity, respectively. Nevertheless, this approach predicts that any Eulerian velocity bias rapidly decays to unity \[ b_E(k, t) = 1 + D^{-3/2}(t)(b_v(k) - 1), \] (107)

where \( D(t) \) is the linear growth rate normalized to unity at the collapse redshift. To reconcile these two apparently contradictory results, the authors of Ref. [105] argued that the Euler equation for halos should be changed from

\[
\dot{\theta}_h + H \theta_h + \frac{3}{2} H^2 \Omega_{dm} \delta_{dm} + \cdots = 0, \quad \theta_h = \vec{\nabla} \cdot \vec{v}_h, \tag{108}
\]

which predicts the incorrect behavior (107), to

\[
\dot{\theta}_h + H \theta_h + \frac{3}{2} b_v(k) H^2 \Omega_{dm} \delta_{dm} + \cdots = 0, \quad \theta_h = \vec{\nabla} \cdot \vec{v}_h, \tag{109}
\]

where \( H \) is the Hubble rate and \( \Omega_{dm} \) parametrizes the abundance of DM and the dots stand for higher-order terms. Their physical interpretation is that the gravitational force acting on DM halos is statistically biased. Together with the Euler equation for DM

\[
\dot{\theta}_{dm} + H \theta_{dm} + \frac{3}{2} H^2 \Omega_{dm} \delta_{dm} + \cdots = 0, \quad \theta_{dm} = \vec{\nabla} \cdot \vec{v}_{dm}, \tag{110}
\]

one indeed recovers the behavior (105) and the Eulerian velocity halo bias does not decay in time to unity.

Another reason why Eq. (108) cannot describe the momentum evolution of halos is the fact that it does not differ at all from (110). Consequently, it is not possible that Eq. (108) describes clustered objects like halos (or peaks), which are different from the smooth DM distribution as they are supposed to be located at points where \( \vec{\nabla} \delta_{dm} = 0 \) and where the smoothed density contrast is larger than some value \( \nu \sigma_0 \), being \( \nu \) the peak height. In other words, Eq. (108) does not contain any information about the fact we are dealing with peaks.

The aim of this analysis is to write down an Effective Boltzmann Equation (EBE) for the single phase space density of halo centers, which can be used to derive ensemble average halo clustering statistics. Therefore, this must be an effective description, in which
the halo distribution should be thought of as being the mean-field distribution associated with a given realization of the DM distribution. This is the correct interpretation of the halo overabundance defined through a perturbative bias expansion so long as the surveyed volume remains finite, as is the case of real or simulated data \([65, 67]\). The EBE for halos will be different from that of DM. This point should not be underestimated: the DM Boltzmann equation is the starting point of any analytical approach to the description of the LSS (see e.g. \([121]\)). Here we advocate that, since the ultimate goal is the understanding of the statistical properties of what we observe (that is galaxies), one ought to adopt the EBE discussed below as the starting point for halo centers.

This is not the only reason for pursuing this project, though. LSS probes such as peculiar velocities or redshift space distortions can be used to test the law of gravity through a measurement of peculiar velocities, i.e. deviations from pure Hubble flow \([122, 123]\). However, peculiar velocities are effectively measured at the position of halo centers. Therefore, even if the latter locally flow with the DM, velocity statistics may be biased if the halo peculiar velocities do not provide a fair sample of the matter flows (for recent discussions see Refs. \([113, 124, 125, 126, 127, 128]\)). Thus far, measurements of the halo velocity power spectrum appear consistent with little or no statistical velocity bias \([114, 124]\). However, \([105, 113]\) reached somewhat different conclusions. Furthermore, the effect must vanish in the limit \(k \to 0\) so that, at low redshift, scale-dependent nonlinearities complicate the interpretation of the results.

The EBE provides a way of describing the evolution equation for biased tracers. As first suggested in Ref. \([129]\), it encodes the statistical velocity bias which originates from the fluctuations around the ensemble average of the gravitational force acting on the halo and it predicts the right value \((105)\) observed at the linear level. Therefore, the EBE explains at more fundamental level why and how one needs to modify the momentum fluid equation for the coupled-fluids of halos and DM, as suggested by Ref. \([105]\). In addition, it allows us to compute the halo velocity bias at higher orders and to derive additional stochastic components that describe in a mean-field way (i.e. at the statistical level) how biasing influence the dynamics of halo centers.

**Plan of the chapter** This chapter is organized as follows. In Sec. §3.1, we motivate our approach and introduce the various averaging we perform throughout the calculation. In Sec. §3.3 we perform a detailed computation of the left-hand side of the EBE adopting a general method based on the path integral approach for computing conditional averages. Section §3.4 is devoted to a similar discussion regarding the right-hand side of the EBE. We then conclude with some final remarks in Sec. §3.5.
3.1 From the Klimontovich-Dupree Equation to the Boltzmann Equation

Let us first explain in more details what we mean by halo “mean-field”. In a local bias approximation (e.g. [53]), one can express density fluctuations in the halo mean-field as the series expansion

$$\delta_h(\vec{r}, t) = b_1 \delta_R(\vec{r}, t) + \frac{1}{2} b_2 (\delta_R^2(\vec{r}, t) - \sigma_0^2(t)) + \ldots$$  \hspace{1cm} (111)

where the filtering scale $R$ is proportional to the halo mass [117], the bias parameters $b_1$, $b_2$ etc. are ensemble average quantities, whereas $\delta_R(\vec{r}, t)$ is one particular realization of density fluctuations in the large scale structure at cosmic time $t$. The filtering reflects the fact that we are not interested in the internal structure of the halos, only in their center-of-mass position and kinematics. From the knowledge of this halo mean-field, we can compute halo clustering statistics, such as the halo 2-point correlation, as

$$\xi_h(\vec{x}_1 - \vec{x}_2, t) = b_2 \langle \delta_R(\vec{x}_1, t)\delta_R(\vec{x}_2, t) \rangle + \ldots$$  \hspace{1cm} (112)

where $\langle \ldots \rangle$ is an average over realisations of the large scale structure.

Our goal is to write down an EBE for the single particle phase space density $f_h$ of halo centers, such that its zeroth moment yields $\delta_h(\vec{r}, t)$. This EBE must be different from the Boltzmann equation of DM: the latter is recovered only if the halos are statistically unbiased relative to the coarse-grained DM density field (which they are not). Our starting point is the Klimontovich density [130]

$$f_K(\vec{r}, \vec{v}, t) = \sum_i \delta_D[\vec{r} - \vec{r}_i(t)] \delta_D[\vec{v} - \vec{v}_i(t)],$$  \hspace{1cm} (113)

which is the single particle phase space density for DM. We have assumed that all the elementary DM particles have equal mass $m$, and we normalize our units so that $m = 1$.

The corresponding equation of motion for such a phase space density is the Klimontovich-Dupree equation

$$\frac{\partial f_K}{\partial t} + \vec{v} \cdot \frac{\partial f_K}{\partial \vec{r}} - \vec{\nabla} \Phi_K \cdot \frac{\partial f_K}{\partial \vec{v}} = 0,$$  \hspace{1cm} (114)

where

$$\vec{\nabla} \Phi_K(\vec{r}, t) = G_N \int d^3 r' d^3 v' f_K(\vec{r}', \vec{v}', t) \frac{(\vec{r} - \vec{r}')}{|\vec{r} - \vec{r}'|^3}$$  \hspace{1cm} (115)

is the corresponding gravitational force. At this stage we have not learned much, as the Klimontovich density encodes the trajectories of all the particles in the system, which is much more information than we can handle and want to keep track of (in particular, we are not interested in the internal properties of virialized structures). Moreover, eq. (114) describes one particular realization of a Universe, which may not resemble our own Universe. Therefore, we go one step further and average over realizations of Universes
that have the same large-scale structure (i.e. similar coarse-grained properties). In this way we can construct macrostates by averaging over a statistical ensemble of microstates with similar phase space density in (small) volumes containing a sufficient amount of particles.

Let us explain in more detail how this ensemble average is performed with the help of fig. 7.

Figure 7: A pictorial exemplification of the ensemble average described in the text. We consider all the configurations (microstates) for the distribution of DM particles that yield to indistinguishable large scale structure configurations (a given macrostate). (a) The distribution of DM particles is described in the phase space by the Klimontovich density $f_K(\vec{r}, \vec{v}, t)$. Four different realizations are shown. (b) After a suitable smearing (the convolution with a window function) we can describe these microstates in terms of a smooth density $\rho$. In any point $\vec{r}$, $\rho$ can vary in a range of values among the microstates: this stochasticity is encoded in the probability distribution $P(\rho, \vec{v}, t)$. (c) The macrostate is described by the phase space density $f(\vec{r}, \vec{v}, t)$, which follows from the ensemble of microstates through the ensemble average $\langle f_K(\vec{r}, \vec{v}, t) \rangle_{\text{micro}}$, or equivalently by computing the expected value of $\rho$ with respect to the probability density $P(\rho, \vec{v}, t)$.

The Klimontovich density describes a set of points in the phase space. Let us focus on a small box in the configuration space. In the part (a) of fig. 7 we show four different
realizations of this box, which yield the same LSS. Their ensemble average, which we denote by \( \langle \ldots \rangle_{\text{micro}} \), gives the mean value \( f(\vec{r}, \vec{v}, t) \):

\[
\left\langle f_K(\vec{r}, \vec{v}, t) \right\rangle_{\text{micro}} = \left\langle \sum_i \delta_D [\vec{r} - \vec{r}_i(t)] \delta_D [\vec{v} - \vec{v}_i(t)] \right\rangle_{\text{micro}} = f(\vec{r}, \vec{v}, t),
\]

(116)

An alternative way of understanding this average is to smear out the discrete distribution of points into a continuous density with an associated probability \( P_t(\rho, \vec{v}, t) \), where \( \vec{r} \) is any position inside the box, as represented in the part (b) of fig. 7. The range of values assumed by \( \rho \) over the ensemble defines the probability density \( P_t(\rho, \vec{v}, t) \) for having a velocity \( \vec{v} \) and a density \( \rho \) in a small volume centered at position \( \vec{r} \), provided that the LSS is the one fixed in (c) in fig. 7. In this case, \( f(\vec{r}, \vec{v}, t) \) is given by the mean value of \( \rho \) with respect to the probability distribution \( P_t(\rho, \vec{v}, t) \):

\[
f(\vec{r}, \vec{v}, t) = \int_0^\infty d\rho \rho P_t(\rho, \vec{v}, t) = \int_0^\infty d\rho \rho P_t(\rho | \vec{v}, t) P_t(\vec{v}, t) = \langle \rho | \vec{v} \rangle P_t(\vec{v}, t).
\]

(117)

We will drop the subscript \( \vec{r} \) here and henceforth, as it is clear that \( P(\rho, \vec{v}, t) \) is defined at a given spatial location.

The distribution function \( f(\vec{r}, \vec{v}, t) \) that characterizes the macrostates defines, upon taking moments of the velocity \( \vec{v} \), the overdensity field, the bulk velocity, etc. of a given macrostate. For instance, the matter overdensity field \( \delta(\vec{r}, t) \) is the zeroth moment

\[
\bar{\rho}(1 + \delta(\vec{r}, t)) = \int d^3v f(\vec{r}, \vec{v}, t).
\]

(118)

This distribution function satisfies the equation

\[
\frac{\partial f}{\partial t} + \vec{v} \cdot \frac{\partial f}{\partial \vec{r}} - \vec{v} \Phi \cdot \frac{\partial f}{\partial \vec{v}} = \nabla_{\vec{v}} \cdot \vec{F},
\]

(119)

where

\[
\vec{F}(\vec{r}, \vec{v}, t) = G_N \int d^3r' d^3v' f(\vec{r}', \vec{v}', t) \frac{(\vec{r} - \vec{r}')}{|\vec{r} - \vec{r}'|^3}
\]

(120)

and

\[
f_{2c}(\vec{r}, \vec{v}, \vec{r}', \vec{v}', t) = f_2(\vec{r}, \vec{v}, \vec{r}', \vec{v}', t) - f(\vec{r}, \vec{v}, t) f(\vec{r}', \vec{v}', t)
\]

(122)

1 Note that this smearing should not be confused with the filtering procedure introduced in section 3.3. The latter defines a UV cutoff in momentum space denoted by \( \Lambda \), while the former represents an average over microstates in small, real space volumes of size \( \ell \), with the requirement that \( \ell \ll 1/\Lambda \).
is the irreducible two-particle correlation function.

Equation (119) is our starting point towards a derivation of the EBE for halos, as it enables us to apply constraints involving $\delta$ etc. to define biased tracers of the large scale structure.

### 3.2 Halos as Extrema of the Smoothed Density Field

In order to write down an EBE for the halo phase space density $f_h(\vec{r}, \vec{v}, t)$, we follow the evolution over cosmic time, and in Eulerian space, of the so-called proto-halos until their virialization. The proto-halos are the progenitors of isolated DM halos, i.e. DM halos that are not contained in any larger halo. While their shape and topology change as a function of time (smaller substructures gradually merge to form the final halo), their centre of mass moves along a well-defined trajectory determined by the surrounding mass density field. Hence, unlike virialized halos that experience merging, by construction proto-halos always preserve their identity. Their total number is, therefore, conserved over time, such that we can write a continuity and Euler equation for their number density and velocity, respectively.

We will consider a specific model for the halo biasing in order to derive explicit expressions for the amplitude and scale-dependence of the effective corrections to the halo EBE. Namely, we will hereafter designate by halos (or peaks) those clustered objects that are located at the points where $\nabla \delta_R(\vec{r}) = \vec{0}$ and where the smoothed density contrast satisfies $\delta_R(\vec{r}) = \nu \sigma_0$ (we will occasionally write $\delta pk \equiv \nu \sigma_0$). The subscript $R$ represents quantities that have been smoothed on the (comoving) size $R$ of a halo. For instance,

$$\delta_R(\vec{r}, t) = \int \text{d}^3 r' \delta(\vec{r}', t) W\left(\frac{1}{R} |\vec{r} - \vec{r}'|\right), \quad (123)$$

where $W(x)$ a spherically symmetric smoothing kernel. Moreover, $\nu$ is the peak height and $\sigma_0 \equiv \sigma_0(t)$ is the variance of the DM field at time $t$ in a given region of size $R$. In general, the spectral moments of the smoothed fields are given by

$$\sigma^2_j = \int \frac{\text{d}^3 k}{(2\pi)^3} k^{2j} P(k) \tilde{W}^2(kR), \quad (124)$$

where $\tilde{W}(x)$ is the Fourier transform of $W(x)$, $P(k) = \langle |\delta(\vec{k}, t)|^2 \rangle_{\text{macro}}$ is the matter power spectrum, the ensemble average is over the macrostates and the time dependence is implicit. We will occasionally refer to those conditions as the peak constraint, although we enforce the extremum condition solely. However, notice that, in the high peak limit $\nu \gg 1$, nearly all extrema are maxima of $\delta_R$.

Last but not least importantly, the filtering of the density field reflects the fact that halos are extended objects. In our model, we will take this into account upon assuming that halo centers feel a force which is a smoothed version of the force acting on dark matter particles.
3.3 The Halo Boltzmann Equation: The Left-hand Side

Our goal is to specialize the Boltzmann equation (119) from the DM phase space density to the peak phase space density so as to describe the evolution of halos. For this purpose, \( f_{pk}(\vec{r}, \vec{v}, t) \) will designate the peak phase space density – i.e. the number density of peaks of velocity \( \vec{v} \) at the position \( \vec{r} \) and time \( t \) – such that its zeroth moment corresponds to the halo mean-field \( \delta_h(\vec{r}, t) \). The density \( f_{pk} \) can be written in terms of conditioned probabilities, as we will show in eq. (154) of section 3.4. The peak constraint in \( \vec{r} \) also affects the computation of the force on the right-hand side of Boltzmann equation, introducing conditioned probabilities. As we will discuss in section 3.4, this side of the Boltzmann equation eventually vanishes.

Let us focus now on the left-hand side of (119). There, the following operator appears:

\[
\vec{\nabla} \Phi(\vec{r}, t) \cdot \partial_{\vec{v}} f(\vec{r}, \vec{v}, t),
\]

which is the product of two operators evaluated at the same point, that is, a composite operator. It is well-known that, owing to the local product of fields making up the composite operator, new ultraviolet divergences appear and depend on the ultraviolet cut-off \( \Lambda \sim \mathcal{O}(1/R) \). The appearance of these divergences requires the introduction of new counterterms. The physical reason why the term \( 125 \) changes in the equation for halos is that the force felt by peaks will be statistically biased with respect to the DM case. The technical way by which we can reconstruct the form of \( \vec{\nabla} \Phi \cdot \partial_{\vec{v}} f \) for the halo Boltzmann equation is through renormalization.

Renormalized operators can be defined and generically expressed as linear combinations of all the bare operators of equal or lower canonical dimensionality (see, for instance, [132])

\[
[O_i(\vec{r}, t)] = \sum_j Z_{ij}(\Lambda) O_j(\vec{r}, t).
\]

Therefore, any composite operator can be expressed as a sum of operators allowed by the symmetries of the problem at hand. The renormalized operator \( [\vec{\nabla} \Phi \cdot \partial_{\vec{v}} f] \) being a scalar field, it can mix with all the operators allowed by the extended Galilean symmetry \([133]\) and must remain a function of the velocity \( \vec{v} \) and preserve the derivative \( \partial_{\vec{v}} \). We can classify the operators \( O_i \) that mix with \( \vec{\nabla} \Phi \cdot \partial_{\vec{v}} f \) according to their number of derivatives in \( \vec{v} \):

\[
[\vec{\nabla} \Phi \cdot \partial_{\vec{v}} f] = \vec{\nabla} \Phi \cdot \partial_{\vec{v}} f + Z_{01} \vec{\nabla} \delta \cdot \partial_{\vec{v}} f + \ldots + Z_{2} \vec{\nabla} \Phi \cdot \partial_{\vec{v}}^3 f + \ldots.
\]

In the picture in which \( \Lambda \) represents the smoothing scale at which the peak is defined, \( \Lambda \) never truly diverges in CDM cosmologies since it is a physical scale. However, in the formalism we introduce here we wish to appeal to the reader’s intuition of the QFT renormalisation scheme. Within this context, our terminology is similar to [131].
Once we have accomplished this classification, we select the leading operators according to the number of derivatives (we are considering large scales) and fields contained in them (adding additional fields brings to a result including more power spectra). The fields we are dealing with are $\Phi$ (which can only appear with derivatives, since the gravitational potential is not observable and does not directly affect the dynamics) and the density contrast $\delta = \nabla^2 \Phi / \alpha$, where $\alpha = 3H^2 \Omega_m/2$. We can then organise the $O_i$ according to the number of extra couples of derivatives (in order to build a vector, we must add two derivatives to $\tilde{\nabla} \Phi$ and contract two indices) and extra fields appearing in them. The result is shown in fig. 8, where moving to the right means adding two derivatives, and moving downwards means adding a field. The appearance of derivative operators makes clear that the statistical bias manifests itself only through $k$-dependent corrections and, therefore, vanishes in the limit of small momenta.

$$\begin{align*}
O_0 &= \tilde{\nabla} \Phi \cdot \partial_{\vec{v}f} \\
O_1 &= \tilde{\nabla} \delta \cdot \partial_{\vec{v}f} \\
O_2 &= \delta \tilde{\nabla} \Phi \cdot \partial_{\vec{v}f} \\
O_3 &= (\tilde{\nabla} \Phi \cdot \tilde{\nabla}) \tilde{\nabla} \Phi \cdot \partial_{\vec{v}f} \\
O_n &= \cdots
\end{align*}$$

Figure 8: Operators allowed by rotational invariance that can mix with $\tilde{\nabla} \Phi \cdot \partial_{\vec{v}f}$, classified according to their number of fields and derivatives.

The renormalization procedure requires a suitable prescription. For our purpose, it is convenient to adopt the following renormalization condition: the correlators of the renormalized quantities should reproduce the correlators at the peak,

$$\left\langle [O_1(\vec{r}, t)] \cdot [O_2(\vec{r}, t)] \cdots [O_n(\vec{r}, t)] \right\rangle = \left\langle O_1(\vec{r}, t) \cdot O_2(\vec{r}, t) \cdots O_n(\vec{r}, t) \right\rangle_{pk}. \quad (128)$$

Here and henceforth, the notation $\langle \cdots \rangle_{pk}$ will designate conditional averages at the peak position. In our problem, it is precisely the renormalization of the operator $\tilde{\nabla} \Phi(\vec{r}, t) \cdot \partial_{\vec{v}f}(\vec{r}, \vec{v}, t)$ that leads to differences in the clustering properties of halos and DM, as we will now see. Notice that, since we want to follow the evolution over cosmic time of the so-called proto-halos until their virialization, the renormalization condition has to be imposed at all times.
3.3.1 Renormalization at the linear order

In this subsection, we study the renormalization of the composite operator $\tilde{\nabla} \Phi \cdot \partial_{\tilde{u}f}(\tilde{r}, \tilde{v}, t)$ at the linear order in perturbation theory.

We begin with the computation of the average of $\tilde{\nabla} \Phi(\tilde{r}, t) \cdot \partial_{\tilde{u}f}(\tilde{r}, \tilde{v}, t)$ at the peak position:

$$\langle \tilde{\nabla} \Phi(\tilde{r}, t) \cdot \partial_{\tilde{u}f}(\tilde{r}, \tilde{v}, t) \rangle_{pk} = \partial_{\tilde{v}f_{pk}}(\tilde{r}, \tilde{v}, t) \cdot \langle \tilde{\nabla} \Phi(\tilde{r}, t) \rangle_{pk}. \quad (129)$$

This amounts to computing the correlators using the appropriate conditional probabilities.

Let us first operate at the linear level in perturbation theory such that, at early times, the DM density $\delta$ and the gravitational force $\tilde{\nabla} \Phi$ are Gaussian variables. We can then apply the theorem stated, for instance, in Refs. [47, 134], which ensures that the conditional probability of zero-mean Gaussian variables $X_A$ and $X_B$ is itself a Gaussian variable with mean

$$\langle X_B | X_A \rangle = \frac{\langle X_B \otimes X_A \rangle}{\langle X_A \otimes X_A \rangle} X_A \quad (130)$$

and covariance matrix

$$C(X_B, X_B) = \langle X_B \otimes X_B \rangle - \frac{\langle X_B \otimes X_A \rangle}{\langle X_A \otimes X_A \rangle} \langle X_A \otimes X_B \rangle. \quad (131)$$

In our case, we identify $X_A$ with $\tilde{\nabla} \delta_R$ and $X_B$ with $\tilde{\nabla} \Phi_R$. As explained above, the gravitational force acting on the halo centers is smoothed to reflect the finite extent of the halos. Since $\langle \tilde{\nabla} \Phi \otimes \delta \rangle = \tilde{0}$ because of rotational invariance, we can simply compute $\langle \tilde{\nabla} \Phi_{pk} \rangle$ as $\langle \tilde{\nabla} \Phi_R | \tilde{\nabla} \delta_R = \tilde{0} \rangle$. The mean shift of $\tilde{\nabla} \Phi$ at the peak position is given by

$$\langle \tilde{\nabla} \Phi \rangle_{pk} = \langle \tilde{\nabla} \Phi_R | \tilde{\nabla} \delta_R \rangle = \frac{\langle \tilde{\nabla} \Phi_R \cdot \tilde{\nabla} \delta_R \rangle_{pk}}{\langle \tilde{\nabla} \delta_R \rangle_{pk}^2} \tilde{\nabla} \delta_R = -\alpha \frac{\langle \delta_R^2 \rangle}{\langle \tilde{\nabla} \delta_R \rangle_{pk}^2} \tilde{\nabla} \delta_R \quad (132)$$

where we have highlighted the dependence of the spectral moments $\sigma_j$ on the UV cut-off $\Lambda$. Since at the peak $\tilde{\nabla} \delta_R$ is vanishing, no renormalization is needed. We will hereafter drop the subscript $R$ whenever $\tilde{\nabla} \Phi$ and $\delta$ appear because we will always refer to peaks, hence the smoothing is understood. Notice also that, while the renormalisation condition is imposed at all times, the scale $\Lambda$ is fixed at initial time and does not evolve.
In reality, $\Lambda$, which is related to the smoothing scale $R$ that define the Lagrangian extent of a halo, likely evolves with time (see e.g. [68]). For simplicity however, we will ignore this complication in the present work.

We recall now the prescription (128) where we used the fact that although the mean shift of the gravitational force vanishes at the peak and thus there is no extra force at the peak-by-peak level, the gravitational force at the peak receives an correction when the statistical ensemble average is taken, as exemplified by eq. (136). The extra effective force felt by the halos is purely of statistical origin and, therefore, does not violate the Equivalence Principle. Furthermore, the time-dependence of $Z_{01}$ is

\[
\left\langle \left( \nabla \Phi(\vec{r},t) \cdot \partial_{\vec{v}} f(\vec{r},\vec{v},t) \right)^2 \right\rangle_{pk} = \partial_{\vec{v}} f_{pk}(\vec{r},\vec{v},t) \cdot \partial_{\vec{v}} f_{pk}(\vec{r},\vec{v},t) \left\langle \nabla \Phi(\vec{r},t) \cdot \nabla \Phi(\vec{r},t) \right\rangle_{pk} ,
\]

where we used the fact that $\left\langle \nabla_i \Phi(\vec{r},t) \cdot \nabla_j \Phi(\vec{r},t) \right\rangle \propto \delta_{ij}$. This requires the knowledge of the covariance matrix for $\nabla \Phi$ given the constraint at the peak. At the linear order the unconstrained correlators needed to evaluate eq. (131) read

\[
\begin{align*}
\left\langle \nabla \Phi \otimes \nabla \Phi \right\rangle &= \frac{\alpha^2}{3} \sigma_{-1}^2(\Lambda) \, \mathbb{1}_{3 \times 3}, \\
\left\langle \nabla \Phi \otimes \nabla \delta \right\rangle &= -\frac{\alpha}{3} \sigma_0^2(\Lambda) \, \mathbb{1}_{3 \times 3}, \\
\left\langle \nabla \delta \otimes \nabla \Phi \right\rangle &= -\frac{\alpha}{3} \sigma_0^2(\Lambda) \, \mathbb{1}_{3 \times 3}, \\
\left\langle \nabla \delta \otimes \nabla \delta \right\rangle &= \frac{1}{3} \sigma_1^2(\Lambda) \, \mathbb{1}_{3 \times 3}.
\end{align*}
\]

The covariance matrix for $\nabla \Phi$ given the peak constraint is then equal to

\[
C \left( \nabla \Phi, \nabla \Phi \right)_{pk} = \frac{\alpha^2}{3} \left( \sigma_{-1}^2(\Lambda) - \frac{\sigma_0^4(\Lambda)}{\sigma_1^2(\Lambda)} \right) \, \mathbb{1}_{3 \times 3},
\]

and

\[
\left\langle \nabla \Phi(\vec{r},t) \cdot \nabla \Phi(\vec{r},t) \right\rangle_{pk} = \text{Tr} \left( \nabla \Phi, \nabla \Phi \right) = \alpha^2 \left( \sigma_{-1}^2(\Lambda) - \frac{\sigma_0^4(\Lambda)}{\sigma_1^2(\Lambda)} \right) .
\]

We recall now the prescription (128): at the lowest order, $[\nabla \Phi \cdot \partial_{\vec{v}} f]$ contains only the operators $O_0$ and $O_1$ (see fig. 8). Hence, we learn that the renormalized force felt by the halos at the linear order reads

\[
\left[ \nabla \Phi(\vec{r},t) \cdot \partial_{\vec{v}} f(\vec{r},\vec{v},t) \right]_{\text{first}} = \left( \nabla \Phi(\vec{r},t) + \frac{\alpha}{3} \frac{\sigma_0^2(\Lambda)}{\sigma_1^2(\Lambda)} \nabla \delta(\vec{r},t) \right) \cdot \partial_{\vec{v}} f(\vec{r},\vec{v},t).
\]
that of \( \alpha \sigma_0^2 / \sigma_1^2 \), that is, the amplitude of the resulting \( k^2 \)-correction in Fourier space does not depend on time. Therefore, this gravity bias does not decay with time, in agreement with \([65, 105]\). Note that the same result eq. (137) was found in Refs. \([116, 117, 05, 129]\), but our Boltzmann approach combined with a renormalization procedure puts it into a different perspective.

3.3.2 Renormalization at higher order with the path integral technique

In this subsection, we will compute the renormalized operator \( [\vec{\nabla} \Phi(\vec{r}, t) \cdot \partial v_f(\vec{r}, \vec{v}, t)] \) at the second order in perturbation theory. We will compute the correlators beyond the Gaussian approximation with the path integral technique (details can be found in appendix 10.1). We will restrict ourselves to the four operators \( O_0, O_1, O_2, O_3 \) defined in fig. 8.

Let us give a closer look to the renormalization prescription (128). We begin by observing that the expectation values of single operator \( \langle O_i \rangle \) are not helpful, because they are all proportional to expectation values of vector fields and thus vanish, as in (129) and (132). Therefore, in order to determine the coefficients \( Z_{ij} \) of (126), we write down the system given by the expectation values of products of two operators:

\[
\sum_{i,m} Z_{il} Z_{jm} \langle O_l O_m \rangle = \langle O_i O_j \rangle_{pk}. \tag{138}
\]

We must now specify how we define the leading and subleading orders of our computation. On the left-hand side of eq. (138), operating at second order the correlation functions \( \langle O_l O_m \rangle \) contain at most two power spectra. Therefore, at leading order we compute the correlators by including only their Gaussian components and obtaining one power spectrum in the result; the next-to-leading order includes up to two power spectra. On the right-hand side, the leading order of the computation is obtained by considering again only the Gaussian component of the fields. When computing the correlators \( \langle O_l O_m \rangle_{pk} \), the linear order is equivalent to consider only the leading order in \( \nu \sigma_0 \).
The results for the correlators without the peak constraint are \(^3\)

\[
\begin{align*}
\langle \mathcal{O}_0 \mathcal{O}_0 \rangle &= \frac{a^2}{3} \sigma_{1}^2 \delta_{ij} \\
\langle \mathcal{O}_0 \mathcal{O}_1 \rangle &= -\frac{a}{3} \sigma_{1}^2 \delta_{ij} \\
\langle \mathcal{O}_0 \mathcal{O}_2 \rangle &= \frac{a^2}{3} \mathcal{K}_1 \delta_{ij} \\
\langle \mathcal{O}_0 \mathcal{O}_3 \rangle &= -\frac{a^3}{6} \mathcal{K}_1 \delta_{ij} \\
\langle \mathcal{O}_1 \mathcal{O}_1 \rangle &= \frac{1}{3} \sigma_{1}^2 \delta_{ij} \\
\langle \mathcal{O}_1 \mathcal{O}_2 \rangle &= -\frac{a}{3} \frac{17}{27} \sigma_{0}^4 \delta_{ij} \\
\langle \mathcal{O}_1 \mathcal{O}_3 \rangle &= -\frac{a^2}{3} \mathcal{K}_2 \delta_{ij}
\end{align*}
\]

(leading order) (next-to-leading order)

where we are dropping the factors of \(\partial_{\vec{v}} f\) to simplify the notation. The correlators calculated with the peak constraint read

\[
\begin{align*}
\langle \mathcal{O}_0 \mathcal{O}_0 \rangle_{pk} &= \frac{a^2}{3} \delta_{ij} \left[ \left( \sigma_{1}^2 - \frac{\sigma_{0}^4}{\sigma_{1}^2} \right) + (\nu \sigma_0) \left( \mathcal{K}_1 \frac{1}{\sigma_0} - \frac{20}{21} \frac{\sigma_0^4}{\sigma_{1}^2} \right) \right] \\
\langle \mathcal{O}_0 \mathcal{O}_2 \rangle_{pk} &= \frac{a^2}{3} (\nu \sigma_0) \delta_{ij} \left( \sigma_{1}^2 - \frac{\sigma_{0}^4}{\sigma_{1}^2} \right) \\
\langle \mathcal{O}_0 \mathcal{O}_3 \rangle_{pk} &= \frac{a^3}{6} (\nu \sigma_0) \delta_{ij} \left( \sigma_{1}^2 - \frac{\sigma_{0}^4}{\sigma_{1}^2} \right)
\end{align*}
\]

(147) (148) (149)

We also notice that our renormalization prescription \((128)\) enforces \(\langle \mathcal{O}_1 \mathcal{O}_i \rangle_{pk} = 0 \forall i\). This corresponds to a subsystem of \((138)\) (when \(i = 1\)) whose right-hand side is \(0\). The solution of which is \(Z_{ij} = 0\) or \([\vec{\nabla} \delta \cdot \partial_{\vec{v}} f] = 0\). This is a consequence of our renormalization prescription: once we impose that at the position \(\vec{r}\) there is a peak defined by \(\vec{\nabla} \delta(\vec{r}) = \vec{0}\), the operator \(\vec{\nabla} \delta\) after renormalization cannot be anything but zero.

We are now free to impose the condition \([\mathcal{O}_i] = \mathcal{O}_i + \ldots\), i.e. the diagonal entries satisfy the condition \(Z_{ii} = 1\) for \(i \neq 1\). Let us first discuss the system for the renormalization of the three operators \{\(\mathcal{O}_0, \mathcal{O}_1, \mathcal{O}_2\}\}. We begin by inspecting the equation \(\langle \mathcal{O}_0 \mathcal{O}_0 \rangle_{pk} = Z_{00} Z_{00} \langle \mathcal{O}_0 \mathcal{O}_0 \rangle\), which contains all (and only) the coefficients \(Z_{0i}\). We know that

\[
Z_{00} = 1, \quad Z_{01} = a \frac{\sigma_0^2}{\sigma_1^2} + \mathcal{O}(\nu \sigma_0), \quad Z_{02} = \mathcal{O}(\nu \sigma_0).
\]

\(^3\) see Appendix \(10.2\) for details about the calculation of these correlators, and for the definition of the finite quantities \(\mathcal{K}_1, \mathcal{K}_2\).
We should impose the equality \( \langle O_0 O_0 \rangle_{\text{pk}} = Z_{0i} Z_{ij} \langle O_i O_j \rangle \) separately between the leading order terms and the next-to-leading order ones. We find that

\[
Z_{01} = \alpha \frac{\sigma_0^2}{\sigma_1^2} + O(\nu \sigma_0), \quad Z_{02} = (\nu \sigma_0) \left[ \frac{1}{2\sigma_0^2} + \frac{31\sigma_0^2}{6 \left(-17\sigma_0^2 + 7K_1\sigma_1^2\right)} \right] + O((\nu \sigma_0)^2).
\]

We now look at the equations \( \langle O_0 O_2 \rangle_{\text{pk}} = Z_{0i} Z_{2j} \langle O_i O_j \rangle \) and \( \langle O_2 O_2 \rangle_{\text{pk}} = Z_{2i} Z_{2j} \langle O_i O_j \rangle \), which include the unknowns \( Z_{20} \) and \( Z_{21} \). These two equations, once we select the terms consistently with their order (for example, the connected contribution of \( \langle O_2 O_2 \rangle = \langle \delta \tilde{\nabla} \Phi \delta \tilde{\nabla} \Phi \rangle \) contains 3 power spectra and should not be put together with second order quantities) we get

\[
Z_{20} = \frac{\kappa_1 \sigma_1^2 - \frac{17}{2} \sigma_0^2}{(\sigma_0^2 - \sigma_1^2 \sigma_1^2)} + (\nu \sigma_0),
\]

\[
Z_{21} = \alpha \left[ \frac{\sigma_0^2 \kappa_1 - \frac{17}{2} \sigma_0^2 \sigma_1^2}{(\sigma_0^2 - \sigma_1^2 \sigma_1^2)} \right] + \frac{1}{\sigma_1^2} \left[ -\frac{(\kappa_1 \sigma_1^2 - \frac{17}{2} \sigma_0^2)^2}{(\sigma_0^2 - \sigma_1^2 \sigma_1^2)} + \sigma_0^2 \left[ \frac{(\kappa_1 \sigma_1^2 - \frac{17}{2} \sigma_0^2)^2}{(\sigma_0^2 - \sigma_1^2 \sigma_1^2)} \right] + \frac{31\sigma_0^2}{6 \left(-17\sigma_0^2 + 7K_1\sigma_1^2\right)} \right] + (\nu \sigma_0).
\]

This concludes the determination of the renormalization coefficients at the next-to-leading order for the set of operators \( \{ O_0, O_1, O_2 \} \), keeping up to two power spectra in the (connected) correlators. We could now include the operator \( O_3 \) in the discussion, since it contains the same number of fields \( \Phi \) and derivatives acting on them as \( O_2 \). But in this way we should compute correlation functions as \( \langle O_3 O_3 \rangle_{\text{pk}} \), whose connected part contains three power spectra, and thus goes beyond the perturbative order we are considering. The difference with respect to the operator \( O_2 \) is that the correlator \( \langle O_2 O_2 \rangle_{\text{pk}} \) with the constraint \( \delta = \nu \sigma_0 \), gives \( (\nu \sigma_0)^2 \) times a correlator with one power spectrum. Now, \( \sigma_0^2 \) is of the same order of a power spectrum in a perturbative expansion (see eq. 124), thus \( \langle O_2 O_2 \rangle_{\text{pk}} \) is of a lower order with respect to \( \langle O_3 O_3 \rangle_{\text{pk}} \). Therefore, at the perturbative order we are considering, \( [\tilde{\nabla} \Phi \cdot \partial_{\tilde{\nabla}} f] \) does not mix with \( (\tilde{\nabla} \Phi \cdot \tilde{\nabla}) \tilde{\Phi} \cdot \partial_{\tilde{\nabla}} f \).

We have therefore found that the next-to-leading order renormalization correction reads

\[
\left[ \tilde{\nabla} \Phi (\vec{r}, t) \cdot \partial_{\tilde{\nabla}} f (\vec{r}, \vec{v}, t) \right]_{\text{NLO}} = \left( \nu \sigma_0 (\Lambda) \left[ \frac{1}{2\sigma_0^2 (\Lambda)} + \frac{31\sigma_0^2 (\Lambda)}{6 \left(-17\sigma_0^2 (\Lambda) + 7K_1(\Lambda)\sigma_1^2 (\Lambda)\right)} \right] \delta \tilde{\nabla} \Phi \right) \cdot \partial_{\tilde{\nabla}} f (\vec{r}, \vec{v}, t).
\]

Let us summarize the main results of this section, as they are the central point of this work: we have renormalized the composite operator \( [\tilde{\nabla} \Phi \cdot \partial_{\tilde{\nabla}} f] \) imposing the peak
constraint. Already at first order, we have found that while the mean of $\nabla \Phi$ is zero as in the unconstrained case, the variance has an additional term. The composite operator therefore renormalized in such a way as to generate the correct variance and account for this statistical effect, in (137). We then compute the next-to-leading order effect and find the corresponding renormalization, (150).

3.3.3 Non-renormalization of the DM Boltzmann equation

It is crucial to notice that for the smooth component of DM (that is without peak constraint) no renormalization is needed and the corresponding Boltzmann equation is not altered. This is consistent with the fact that we have used standard perturbation theory results for DM to evaluate the statistical correlators at the peak locations. Consistency can be explicitly checked upon noticing that, for DM, probabilities are not conditional. For instance, we get at linear order

\[
\left\langle \nabla \Phi(\vec{r}, t) f(\vec{r}, \vec{v}, t) \right\rangle = f(\vec{r}, \vec{v}, t) \left\langle \nabla \Phi(\vec{r}, t) \right\rangle = 0,
\]

\[
\left\langle \nabla \Phi(\vec{r}, t) \cdot \nabla \Phi(\vec{r}, t) f(\vec{r}, \vec{v}, t) f(\vec{r}, \vec{v}_2, t) \right\rangle = \frac{a^2}{3} \sigma^2 \left\langle f(\vec{r}, \vec{v}_1, t) f(\vec{r}, \vec{v}_2, t) \right\rangle \cdot \nabla \Phi(\vec{r}, t)
\]

and one simply finds

\[
\left[ \nabla \Phi(\vec{r}, t) \right]_{\text{linear}} = \nabla \Phi(\vec{r}, t).
\]

This remains true also at higher orders.

3.4 THE HALO BOLTZMANN EQUATION: THE RIGHT-HAND SIDE

We will now deal with the right-hand side of the Boltzmann equation, see Eqs. (121) and (122). It contains a force term that depends on the irreducible two-particle correlation (which, in turns, depends on the irreducible three-particle correlation etc.). It is the correlation encoded by $f_{2c}$ that causes the non-conservation of the one-particle phase-space distribution function in configuration space. The physical interpretation of this effect is that the gravitational interactions between particles (which induce long-range correlations) lead to clustering and, therefore, non-conservation of the one-particle phase-space distribution.

To calculate this extra force, we follow Ref. [134]. The procedure consists in relating phase-space densities to probability distributions for the mass and velocity field, which are naturally specified by models of cosmological structure formation. As explained in sec. 3.1, we can express the phase space density $f(\vec{r}, \vec{v}, t)$ in terms of the probability distribution $P(\rho, \vec{v}, t)$. This approach can be extended to the irreducible two-particle...
correlation. Inspecting eqs. (121) and (122), we observe that \( f_{2c}(\vec{r}, \vec{v}, \vec{r}', \vec{v}', t) \) is integrated over \( d^3v' \) and we need only the velocity at one point (the superscript \( ' \) indicates that quantities are evaluated at \( \vec{r}' \))

\[
f_{2c}(\vec{r}, \vec{v}, \vec{r}', \vec{v}', t) = \int d^3v' f_{2c}(\vec{r}, \vec{v}, \vec{r}', \vec{v}', t) = \int_0^\infty d\rho \rho \int_0^\infty d\rho' \rho' \left[ P(\rho, \vec{v}, \rho', t) - P(\rho, \vec{v}, t)P(\rho', t) \right] \\
= \int_0^\infty d\rho \rho \int_0^\infty d\rho' \rho' \left[ P(\rho, \vec{v}, \rho', t) - P(\rho|\vec{v}, t)P(\rho', t) \right] P(\vec{v}, t) \\
= \left[ \langle \rho \rho'|\vec{v} \rangle - \langle \rho|\vec{v} \rangle\langle \rho' \rangle \right] P(\vec{v}', t),
\]

(153)

where, in the last equality, the averages are equal-time correlators. Even though one derives these equations assuming a single-valued velocity field in each realization, they are also valid when there is a distribution of velocities at each point. We simply interpret \( \rho \) as the total mass density while \( \vec{v} \) is a single bulk velocity. Therefore, these equations are fully general [134].

Let us specialize to the case of halos. The phase-space distribution (117) can be written using Bayes’ theorem as

\[
f_{pk}(\vec{r}, \vec{v}, t) = \int_0^\infty d\rho \rho P(\vec{v}|\rho, t) P(\rho, t) = \rho_{pk} P(\vec{v}, t|pk) = \bar{\rho}(1 + \delta_{pk}) J_{\vec{v}\vec{v}}^{-3} P(\bar{\vec{v}}, t|pk) \tag{154}
\]

where, in the last passage, we have introduced the quantity

\[
\bar{\psi}(\vec{r}, t) = -\frac{\hbar}{\lambda} \bar{\Phi}(\vec{r}, t) = -\int \frac{d^3r'}{4\pi} \frac{(\vec{r} - \vec{r}')}{|\vec{r} - \vec{r}'|^3} \delta(\vec{r}', t),
\]

(155)

and \( J_{\vec{v}\vec{v}} \) is the Jacobian describing the passage from the variable \( \vec{v} \) to the variable \( \vec{v} \).

In a similar fashion, we can compute the irreducible two-point correlation function

\[
f_{2c}^{Pk}(\vec{r}, \vec{v}, \vec{r}', \vec{v}', t) = \int_0^\infty d\rho \rho \int_0^\infty d\rho' \rho' \left[ P(\vec{v}, \rho'|\rho, t)P(\rho, t) - P(\vec{v}, \rho, t)P(\rho', t) \right] \\
= \rho_{pk} \int_0^\infty d\rho \rho' \left[ P(\vec{v}, \rho'|\rho, t|pk) - P(\vec{v}, \rho, t|pk)P(\rho', t) \right] \\
= \rho_{pk} P(\vec{v}, t|pk) \int_0^\infty d\rho' \rho' \left[ P(\rho', t|\vec{v}, pk) - P(\rho', t) \right] \\
= \bar{\rho}^2(1 + \delta_{pk}) J_{\vec{v}\vec{v}}^{-3} \left[ \langle \delta(\vec{r}', t) | \bar{\psi}(\vec{r}, t), pk \rangle - \langle \delta(\vec{r}', t) \rangle \right] \\
= \bar{\rho} f_{pk}(\vec{r}, \vec{v}, t) \left\langle \delta(\vec{r}', t) | \bar{\psi}(\vec{r}, t), pk \right\rangle.
\]

(156)

In the last equality, we have substituted \( \langle \delta(\vec{r}', t) \rangle = 0 \) as this average vanishes in the absence of peak constraint. The corresponding force felt by the halos is therefore

\[
\vec{F}_{pk} = G_N\bar{\rho} f_{pk}(\vec{r}, \vec{v}, t) \int d^3r' \left\langle \delta(\vec{r}', t) | \bar{\psi}(\vec{r}, t), pk \right\rangle \cdot \frac{\vec{r} - \vec{r}'}{|\vec{r} - \vec{r}'|^3}.
\]

(157)
The computation of the force (157) is reported in Appendix 10.3 and we provide only the final result up to second-order

\[
\bar{F}_{\text{pk}}(\vec{r}, \vec{v}, t) = -4\pi G_N \bar{\rho}_f \mathcal{P}_{\text{pk}}(\vec{r}, \vec{v}, t) \left[ \bar{\psi}(\vec{r}, t) + \left( \bar{\psi}(\vec{r}, t) - \frac{\sigma_0^2}{\sigma_1^2} \bar{\psi} \delta(\vec{r}, t) \right) \frac{\delta(\vec{r}, t)}{\sigma_0^2 - \sigma_1^2 \sigma_{-1}^2} \sigma_0^2 \right.
\]

\[
\times \left( \left( \delta(\vec{r}, t) \bar{\psi}(\vec{r}, t) \cdot \bar{\psi}(\vec{r}, t) \right) - \int d\ell \frac{\ell_i}{|\ell|} \left( \delta(\vec{r}', t) \delta(\vec{r}, t) \psi_i(\vec{r}, t) \right) \right)
\]

\[
+ \left( \bar{\psi}_1 - \frac{\sigma_0^2}{\sigma_1^2} \bar{\psi} \bar{\nu} \delta(\vec{r}, t) \right) \frac{\delta(\vec{r}, t)}{\sigma_0^2 - \sigma_1^2 \sigma_{-1}^2} \left( \int d\ell \frac{\ell_i}{|\ell|} \left( \delta(\vec{r}', t) \delta(\vec{r}, t) \psi_{i+1}(\vec{r}, t) \right) \right)
\]

\[
- \left( \delta(\vec{r}, t) \bar{\psi}(\vec{r}, t) \cdot \bar{\nabla} \delta(\vec{r}, t) \right) \right] ,
\]

(158)

where \( \vec{l} = (\vec{r}' - \vec{r}) \). Using the results for the cumulants reported in (327), (328), (340), (341) we find (notice that the last two lines in (158) cancel each other)

\[
\bar{F}_{\text{pk}}(\vec{r}, \vec{v}, t) = -4\pi G_N \bar{\rho}_f \mathcal{P}_{\text{pk}}(\vec{r}, \vec{v}, t) \left[ \bar{\psi}(\vec{r}, t) + \left( \bar{\psi}(\vec{r}, t) - \frac{\sigma_0^2}{\sigma_1^2} \bar{\psi} \delta(\vec{r}, t) \right) \right.
\]

\[
\times \frac{\delta(\vec{r}, t)}{\sigma_0^2 - \sigma_1^2 \sigma_{-1}^2} \sigma_0^2 \left( \frac{3}{2} - \frac{10}{21} \sigma_0^2 \sigma_{-1}^2 \right) \right].
\]

(159)

Two comments are in order. Firstly, we can easily recover the DM case away from the peaks, either by computing directly the force from expression (156) with aid of \( \langle \delta(\vec{r}', t) \bar{\psi}(\vec{r}, t) \rangle \), or by eliminating the peak constraint taking \( \sigma_0^2, \sigma_1^2 \gg 1 \). In both cases, we obtain

\[
\int d^3 v \bar{v} \bar{\nabla} \bar{\psi} \cdot \bar{F}_{\text{pk, dm}}(\vec{r}, \vec{v}, t) = \bar{0},
\]

(161)

in agreement with the Equivalence Principle. Technically, this cancellation follows from the fact that integrating out the velocities is equivalent to an integration over the configurations of the field \( \bar{\psi} \). As the phase-space distribution is proportional to \( P(\bar{\psi}, t) \), we obtain

\[
\int d^3 \psi P(\bar{\psi}, t) \bar{\psi} = \langle \bar{\psi} \rangle = \left\{ \begin{array}{ll}
- \alpha \frac{\sigma_0^2(\Lambda)}{\sigma_1^2(\Lambda)} \bar{\nabla} \delta \bigg|_{\text{pk}} = \bar{0}, & \text{(for peaks)} , \\
\bar{0} & \text{(for DM)} ,
\end{array} \right.
\]

(162)

where, for peaks, we made use of eq. (132). This equation seems to suggest that one would break the equivalence principle if the constraint were different from \( \bar{\nabla} \delta = \bar{0} \), as
may be the case for other tracers of the large scale structure. However, isotropy and homogeneity would enforce a constraint on the modulus of $\hat{\nabla} \delta$ only, e.g. $|\hat{\nabla} \delta| \leq c$ for some constant $c > 0$. Therefore, the integral over $\bar{\Psi}$ would also vanish in those cases, and the Equivalence Principle would still be satisfied, as it should be.

### 3.5 Discussion and Conclusion

In this analysis we have derived an effective Boltzmann equation that describes the dynamics of the halo mean-field phase space. This equation is necessarily different from the Boltzmann equation of DM as halos are not statistically unbiased relatively to the coarse-grained DM density field. Through a renormalization procedure of composite operators, we have obtained up to second-order in perturbation theory the statistically biased gravitational force felt by halos. We conclude that, at this order, the Boltzmann Equation for halos reads

$$
\frac{\partial f_{pk}}{\partial t} + \bar{v} \cdot \frac{\partial f_{pk}}{\partial \bar{v}} - \left( \hat{\nabla} \Phi_R(\bar{r}, t) + \alpha \frac{\sigma_0^2(R)}{\sigma_1^2(R)} \nabla \delta_R(\bar{r}, t) + \nu \sigma_0^2(R) + 6 \left[ -17 \sigma_0^6(R) + 7 K_1(R) \sigma_1^2(R) \right] \delta_R(\bar{r}, t) \cdot \hat{\nabla} \Phi_R(\bar{r}, t) \right) \cdot \frac{\partial f_{pk}}{\partial \bar{v}} = 0 ,
$$

(163)

where the extra contributions to $\hat{\nabla} \Phi_R(\bar{r}, t)$ are purely of statistical origin: they do not arise because halos feel a different gravitational force than dark matter, as this would violate the equivalence principle, but rather as a result of imposing the peak constraint on their statistics\(^4\). This should be contrasted to the DM Boltzmann equation

$$
\frac{\partial f}{\partial t} + \bar{v} \cdot \frac{\partial f}{\partial \bar{r}} - \hat{\nabla} \Phi(\bar{r}, t) \cdot \frac{\partial f}{\partial \bar{v}} = 0 .
$$

(164)

We emphasize again that $f_{pk}$ is the one-particle phase space density associated with the mean-field distribution of halos that corresponds to a particular $f$, that is, a particular realization of the large scale structure. The biased gravitational force experienced by the halo centers imprints a signature, among others, in the peculiar velocities of virialized halos. This generates the first-order statistical velocity bias which has been discussed in [117, 65, 105] and measured in [60, 105]. However, tidal forces etc. will also be affected. Our approach can in principle be applied to work out these corrections at any order in perturbation theory.

\(^4\) An alternative, perhaps more intuitive, interpretation of this result is the following: instead of modifying the Boltzmann equation for halos with an effective force field, one may interpret the results of equations (137) and (150) as the corrections to the probability distribution function (PDF) of the force acting on halos. In this picture, the Boltzmann equation for halos (163) would look exactly as the one of DM, equation (164), but with a conditional PDF of the form of the terms in bracket in equation (163).
To give intuition about the magnitude of the statistical correction to the gravitational force in eq. (163), we plot in fig. 9 the amplitude of the second and third terms in parenthesis

\[
\xi_1(R) = \frac{\alpha \sigma_0^2(R)}{\sigma_1^2(R)}, \quad \xi_2(R) = \nu \sigma_0(R) \left[ \frac{1}{2\sigma_0^2(R)} + \frac{31\sigma_0^4(R)}{6[-17\sigma_0^6(R) + 7K_1(R)\sigma_1^2(R)]} \right]
\]

as a function of the Lagrangian scale \( R = 1/\Lambda \) assuming the biasing is described by the peak constraint. The magnitude of the second-order statistical correction also rises with \( R \), in agreement with the expectation that these effects should increase with halo mass.

Measurements from N-body simulations are challenging since the effect must vanish in the limit \( k \to 0 \). Furthermore, the sparsity of massive halos could bias the measure-
ments when the purpose is to define volume-weighted statistics (see e.g. [135] for recent discussions). Note, however, that number-weighted statistics are most suited to extract these effects as they are the quantities being observed.

There is some confusion in the literature regarding the nature of this effect. A physical bias will immediately arise as soon as there are at least two different fluids since, in this case, velocities do not need to share the same value. This bias is effective on an object-by-object basis. This is obviously the case of e.g. baryons and DM. Furthermore, this will also be the case of the halos and DM since, owing to the finite halo size, the force acting on the halo center-of-mass, $\vec{\nabla}\Phi_R$, is different from the force acting on DM particles, $\vec{\nabla}\Phi$. Such a coarse-graining procedure can introduce additional terms in the fluid equations, as was already pointed out in [136]. This possibility was recently reconsidered in [137]. If we ignore in Eq.(163) all the statistical corrections but the term $\vec{\nabla}\Phi_R$, then our results are identical to his, except for the time constancy of our $R$. As shown in [137], a time-dependent filter $W(\vec{r}/R, t)$ would yield extra terms in the moments of the Boltzmann equation owing to the time derivative $\partial f/\partial t$.

However, [137] did not include the statistical corrections, which are the focus of this work. In fact, our results would hold even in the extreme case where no extra filtering is applied to the macrostate. This illustrates the fact that the statistical effects we are discussing are not caused by the finite extent of the halos. Regarding the existence of these statistical effects, we note that, while a time-dependent window $W(\vec{r}/R, t)$ mimicking the collapse of halos can induce $k$-dependent contributions at low redshift, corrections are small at high redshift [137]. Therefore, this cannot be the explanation for the suppression of the proto-halo velocity power spectrum measured in [60, 105].

To conclude, note that halo number conservation implies that we follow the collapse of Lagrangian “peak-patch”, as coined by [88], into parent or isolated DM halos. Therefore, we do not explicitly track the merging history of these host halos. Although one could naively apply the extended Press-Schechter (EPS) approach [48] to determine the merger history, it would be very interesting to assess whether the statistical effects considered here also affect merger rates etc. However, it is not obvious how to do this within our current formulation, especially since the smoothing scale $R$ is linked to the parent halo mass and, thus, is not free unlike in EPS theory. Our assumption that the density peak retain their initial critical height (i.e. $\nu$ does not depend on time) is another limitation of our approach. A better treatment would consistently track the time evolution of both $R$ and $\nu$ such that the mass enclosed within $R$ always equates the host halo mass. We leave all this for future work.
Cosmological parameter estimation from galaxy clustering data is hampered by galaxy biasing, i.e. by the fact that galaxies do not perfectly trace the underlying mass distribution. Various theoretical arguments and outcomes of numerical simulations both suggest that, on sufficiently large scales, the halo overdensity $\delta_h(x, \tau)$ can be written as a generic function $f[\delta(x, \tau)]$ of the mass density perturbation $\delta(x, \tau)$. This function can be Taylor-expanded, with the unknown coefficients in the series defining the bias parameters

$$\delta_h(x, \tau) = f[\delta(x, \tau)] = b_1(\tau)\delta(x, \tau) + \frac{b_2(\tau)}{2} (\delta^2(x, \tau) - \langle\delta^2(x, \tau)\rangle) + ... \quad (167)$$

The local model [53, 118] described in, e.g., (167) is however incomplete: there is indeed no a priori reason why the halo density contrast should be only a local function of the matter density contrast.

We have argued in the previous chapter how, already at early times, non local Lagrangian bias parameters arise from simple physical principles. In this chapter, however, we want to discuss another source of non-locality: indeed, already at second-order in perturbation theory, the gravitational evolution generates a term quadratic in the tidal tensor and, therefore, non-local in the density field. This term is absent in the initial conditions. This point was made in Refs. [138, 139, 140, 141] for the matter density contrast and subsequently investigated in the context of halo bias in Refs. [63, 142, 64, 65, 103, 110, 109, 143]. In Refs. [64, 143] in particular, it was pointed out that the symmetries (essentially extended Galilean and Lifshitz symmetries) present in the dynamical equations for the halo and DM systems allow to construct a set of invariant operators which should appear in the halo bias expansion, precisely because they are allowed by the symmetries of the problem. These invariants lead to non-local bias contributions, which numerical simulations have already detected at the quadratic level and found to agree well with the prediction of perturbation theory [109, 110]. In this chapter we compute the non-local bias coefficients in the basis of invariant non-local bias operators at third-order in perturbation theory by using the Lagrangian Bias parameters generated within the peak-background split model.
Although the magnitude of the non-local bias terms is small relative to the linear halo bias, upcoming large scale structure data will be sensitive to them. In particular, the impact of the non-local bias, if not accounted for, could mimic the $k$-dependent suppression of the growth rate in cosmologies with massive neutrinos [144, 145, 146, 147, 148, 149] (see [150, 151] for detailed reviews of the subject). Their influence on the spatial distribution of DM halos has been recently scrutinized in a series of papers [152, 153, 154] using large N-body simulations that incorporate massive neutrinos as an extra set of particles (see also the recent work of [155]). Massive neutrinos generate a scale-dependent bias in the power spectrum of DM halos. As we will see, this effect is somewhat degenerate with the signature left by the various non-local bias terms, which must be taken into account in order to reproduce the N-body data with good accuracy (i.e. at the $\leq 5\%$ level).

**Plan of the chapter** This chapter is organized as follows. In Sec. 4.1, we present our computation of the non-local bias at third-order in perturbation theory as well as the calculation of the Lagrangian bias parameters through the peak-background split model. In Sec. 4.2 and Sec. 4.3 we discuss the halo bias in the presence of massive neutrinos and argue that contributions from nonlocal Lagrangian and gravity bias must be accounted for in order to fit the numerical data. We show our results in Sec. 4.4.

### 4.1 Non-local bias up to third-order

In this section we extend the analysis of [64, 109, 110] and compute the non-local bias coefficients up to third order in perturbation theory. Our starting point is the evolution over cosmic time and in Eulerian space of the halo progenitors - the so-called proto-halos - until their virialization. The basic idea is that, while their shapes and topology change as a function of time (smaller substructures gradually merge to form the final halo), their centre of mass moves along a well-defined trajectory determined by the surrounding mass density field [120]. Therefore, unlike virialized halos that undergo merging, by construction proto-halos always preserve their identity. Their total number is therefore conserved over time, such that we can write a continuity equation for their number density

$$\delta_{h}(x, \tau) + \vec{\nabla} \cdot [(1 + \delta_{h}(x, \tau)) v(x, \tau)] = 0. \tag{168}$$

We may subtract from it the DM mass conservation equation of motion

$$\delta(x, \tau) + \vec{\nabla} \cdot [(1 + \delta(x, \tau)) v(x, \tau)] = 0, \tag{169}$$

where we have assumed unbiased halo velocity (we will relax this assumption later), to get
\[ \delta_h(x, \tau) - \delta(x, \tau) + \vec{\nabla} \cdot [(\delta_h(x, \tau) - \delta(x, \tau)) \mathbf{v}(x, \tau)] = 0. \]  

(170)

This is the fundamental equation which we will solve order by order in perturbation theory, and which gives rise to a non-local bias expansion. For the sake of simplicity, we will restrict ourselves to a matter-dominated Universe and denote by \( x \) the comoving spatial coordinates, \( \tau = \int dt/a \) the conformal time where \( a \) the scale factor in the FRW metric, and \( \mathcal{H} = \frac{d\ln a}{d\tau} \) the conformal expansion rate. In addition, \( \delta(x, \tau) = (\rho(x, \tau)/\bar{\rho} - 1) \) is the overdensity over the mean matter density \( \bar{\rho} \), \( \delta_h(x, \tau) \) stands for the halo counterpart, and \( \mathbf{v}(x, \tau) \) is the common peculiar velocity. In the following we will also denote by \( \Phi(x, \tau) \) the gravitational potential induced by density fluctuations. Since the calculation is rather long, we leave the details in the appendix §11.1, where we review the first- and second-order calculations before deriving the third-order nonlocal biases.

We compute the relation between the halo density contrast and the DM field by solving Eq. (170) at third order to be

\[ \begin{align*}
\delta_h(x, \tau) &= (1 + b_1^L(\tau))\delta(x, \tau) + \left( \frac{1}{2} b_2^L(\tau) + \frac{4}{21} b_1^L(\tau) \right) \delta^2(x, \tau) - \frac{2}{7} b_1^L(\tau) s^2(x, \tau) \\
&\quad + \left[ \frac{1}{3!} b_3^L(\tau) - \frac{13}{21} \left( \frac{1}{2} b_2^L(\tau) + \frac{2}{21} b_1^L(\tau) \right) \right] \delta^3(x, \tau) \\
&\quad - \frac{1}{2} b_1^L(\tau) \psi - \frac{4}{7} \left( \frac{1}{2} b_2^L(\tau) + \frac{2}{81} b_1^L(\tau) \right) \delta(x, \tau) s^2(x, \tau) \\
&\quad - \frac{5}{7} b_1^L(\tau) s(x, \tau) \cdot t(x, \tau) + \frac{2}{7} b_1^L(\tau) s^3(x, \tau),
\end{align*} \]

(171)

where we define the non local operators \( s_{ij}, t_{ij} \) and \( \psi \) in Eqs. (350), (363) and (371) respectively.

The halo density contrast can thus be written as

\[ \delta_h(x, \tau) = b_1 \delta(x, \tau) + \frac{1}{2!} b_2 \delta^2(x, \tau) + \frac{1}{3!} b_3 \delta^3(x, \tau) \\
+ \frac{1}{2} b_2 s^2(x, \tau) + b_\psi \psi(x, \tau) + b_{st} s(x, \tau) \cdot t(x, \tau) + \cdots, \]

(172)

where the relevant bias coefficients are

\[ b_1 = 1 + b_1^L, \quad b_2 = b_2^L + \frac{8}{21} b_1^L, \quad b_3 = -\frac{4}{7} b_1^L, \quad b_\psi = -\frac{1}{2} b_1^L, \quad b_{st} = -\frac{5}{7} b_1^L, \]

(173)

### 4.1 Non-local Lagrangian bias and velocity bias

So far, we have not considered the possibility that the Lagrangian bias factors may be scale-dependent. We have discussed in the previous Chapter §2 that the peak constraint
induces $k$-dependent corrections at all orders in the bias coefficients. Moreover, peak velocities are statistically biased at linear order (see Chapter §3), and this bias propagates to higher order owing to gravity mode-coupling [65, 117]. We have seen that in the framework of the ESP model, the scale-independent Lagrangian bias factors $b_n^m$ should be replaced (in Fourier space) by the scale-dependent functions $c_n^m(k_1, \ldots, k_n, \tau_i)$. In this analysis, we consider a simplified version of the scheme discussed in Sec. §2.2: we neglect the first crossing variable $\mu$ and we restrict to a flat, constant barrier, $B(R) = \delta_{\text{sc}}$. At the lowest orders, Eqs. (80) and (81) are modified into

\begin{align}
  c_1^1(k, \tau_i) &= b_1^{10}(\tau_i) + b_1^{01}(\tau_i)k^2, \\
  c_2(k_1, k_2, \tau_i) &= \left\{ b_2^{20}(\tau_i) + b_2^{11}(\tau_i) (k_1^2 + k_2^2) + b_2^{02}(\tau_i)k_1^2k_2^2 - 2\chi^{10}(\tau_i) (k_1 \cdot k_2) + \chi^{11}(\tau_i) \left[ 3 (k_1 \cdot k_2)^2 - k_1^2k_2^2 \right] \right\}.
\end{align}

As a result, the local Lagrangian bias expansion generalizes to

\begin{align}
  \delta_h(x, \tau_i) &= b_1^{10}(\tau_i)\delta(x, \tau_i) - b_1^{01}(\tau_i)\nabla^2\delta(x, \tau_i) + \frac{1}{2} b_2^{10}(\tau_i) \delta^2(x, \tau_i) \\
  &\quad - b_2^{11}(\tau_i)\delta(x, \tau_i)\nabla^2\delta(x, \tau_i) + \frac{1}{2} b_2^{02}(\tau_i) \left[ \nabla^2\delta(x, \tau_i) \right]^2 \\
  &\quad + \chi_1^{10}(\tau_i)\left( \nabla\delta \right)^2(x, \tau_i) + \frac{1}{2} \chi_1^{11}(\tau_i) \left[ 3\partial_i\partial_j\delta - \delta_{ij}\nabla^2\delta \right]^2(x, \tau_i) + \ldots,
\end{align}

where the various bias factors $b_n^m$ and $\chi_n^m$ can be obtained from a peak-background split applied to the halo mass function (see Sec. §2.2 for details).

In addition, the Zel’Dovich gravity mode-coupling kernels $F_n^{ZA}(k_1, \ldots, k_n)$ should be replaced by

\begin{align}
  F_n^{ZA}(k_1, \ldots, k_n) &= F_n^{ZA}(k_1, \ldots, k_n) \times b_v(k_1) \ldots b_v(k_n).
\end{align}

Here, $b_v(k) = 1 - R_v^2k^2$ is the linear velocity bias we have discussed in Chapter §3. The linear peak velocities can be thought of as arising from the continuous, local relation [117, 65]

\begin{align}
  \mathbf{v}_h(x, \tau_i) &= \mathbf{v}(x, \tau_i) - R_v^2\nabla\delta(x, \tau_i).
\end{align}

\footnotetext{1}{To some degree of approximation, this is equivalent to consider only Gaussian smoothing for all the variables, and not a mixed TopHat-Gaussian scheme. In this case, the variables $f_1 \approx \mu$ are basically the same variable.}
This development could be generalized to include all the higher order terms consistent with the symmetry of the problem. However, since we are mainly interested in the first-order scale-dependent corrections that dominate at relatively small \( k \), we postpone such a study to future work. Eq. (177) reflects the fact that, in the peak approach, the linear velocity bias remains constant throughout time [65].

At the first order, Eq. (345) generalizes easily to

\[
\delta_h^{(1)}(x, \tau) \simeq \left( 1 + b_{01}^{L}(\tau) \right) \delta^{(1)}(x, \tau) + \left( R_0^2 - b_{01}^{L}(\tau) \right) \nabla^2 \delta^{(1)}(x, \tau),
\]

(179)

which follows from the linearized continuity equation \( \delta^{(1)} \propto -\nabla \cdot \mathbf{v}^{(1)} \). This result reproduces the findings of [65], who explicitly took into account the peak constraint. Note also that \( b_{01}^{L}(\tau) = b_{01}^{L}(\tau_1)(a(\tau_1)/a(\tau)) \). As can be seen, the amplitude of the contribution proportional to \( k^2 \) scales as \( R_0^2 - b_{01}^{L}(\tau) \), which is generally non-zero. Therefore, we shall expect this \( k^2 \)-dependence to appear at sufficiently small scales in the halo bias.

At the second order, the halo overabundance \( \delta_h(x, \tau) \) with scale-dependent spatial and velocity bias can be computed by combining the results of Ref. [103] (derived in the Zel’dovich approximation) with those of [117] (derived at higher order in Lagrangian PT). The second-order halo overdensity takes the form

\[
\begin{align*}
\delta_h(x, \tau) &= \frac{1}{2} \int \frac{d^3q_1}{(2\pi)^3} \int \frac{d^3q_2}{(2\pi)^3} \left\{ \mathcal{F}_{2A}^Z(q_1, q_2) + \frac{a(\tau_i)}{a(\tau)} \left[ \mathcal{F}_{2A}^Z(q_1, q_2, \tau_i) + \mathcal{F}_{1A}^Z(q_2) c_1(q_1, \tau_i) \right] \right. \\
&\quad + \frac{a^2(\tau_i)}{a^2(\tau)} c_2(q_1, q_2, \tau_i) + \frac{2}{7} - \frac{3}{7} \left( \mu^2 - \frac{1}{3} \right) \right\} \delta^{(1)}(q_1, \tau) \delta^{(1)}(q_2, \tau) \delta_D(k - q_1 - q_2),
\end{align*}
\]

(180)

where we have momentarily set \( \delta^{(1)} \equiv \nabla^2 \phi^{(1)} \) so that it resembles the well-known PT expression. The various contributions can be reduced to a combination of local and nonlocal quantities analogously to the calculation performed above. For instance,

\[
(1/2)(a(\tau_i)/a(\tau)) [\mathcal{F}_{2A}^Z(q_1) c_1(q_2, \tau_i) + 1 \leftrightarrow 2]
\]

(181)

includes a term proportional to

\[
(1/2)b_{01}^{L}(\tau) \left[ q_2^2 \mathbf{k} \cdot \mathbf{q}_1 \frac{q_1}{q_1} + 1 \leftrightarrow 2 \right]
\]

(182)

which, after some manipulations, becomes

\[
\begin{align*}
b_{01}^{L}(\tau) \nabla^2 \delta^{(1)} &\equiv -b_{01}^{L}(\tau) \nabla^2 \delta^{(2)}(k, \tau) + \frac{9}{7} b_{01}^{L}(\tau) \delta^{(1)}(\nabla^2 \delta^{(1)}) + \frac{18}{7} b_{01}^{L}(\tau) \left( \nabla^2 \delta^{(1)} \right)^2 \\
&+ \frac{6}{7} b_{01}^{L}(\tau) \left( \frac{2}{3H^2} \partial_i \partial_j \Phi - \frac{1}{3} \delta_{ij} \delta \right) \left( \delta_{ij} \delta - \frac{1}{3} \delta_{ij} \nabla^2 \delta \right) \\
&+ \frac{4}{7} b_{01}^{L}(\tau) \left( \frac{2}{3H^2} \partial_i \partial_j \partial_k \Phi - \frac{1}{3} \delta_{ij} \partial_k \delta \right)^2.
\end{align*}
\]

(183)
We have restored the factors of $3/(2H^2)$ to be consistent with the units employed throughout this chapter. The first term in the right-hand side adds to $-b_{01}^L \nabla^2 \delta^{(1)}$ in Eq. (179) to yield $-b_{01}^L \nabla^2 \delta$, where $\delta$ is the mass density fluctuations at second order. The following two terms contribute to the Eulerian biases $b_{11}$ and $b_{02}$. The last two terms are new nonlocal bias contributions, which are however heavily suppressed relative to $\delta^2(x, \tau)$ as they carry additional powers of $k$. The remaining terms in the right-hand side of Eq. (180) can be written analogously. For the purpose of the present work however, Eq. (179) and its $k$-dependent contribution at linear order are sufficient. Therefore, we generalize Eq. (172) to

$$\delta_h(x, \tau) = b_{10} \delta(x, \tau) - b_{01} \nabla^2 \delta(x, \tau) + \frac{1}{2!} b_{20} \delta^2(x, \tau) + \frac{1}{2} b_{s2} s^2(x, \tau) + b_{s1} s(x, \tau) \cdot t(x, \tau) + \cdots,$$

(184)

where

$$b_{10} = 1 + b_{10}^L, \quad b_{01} = -R_2^2 + b_{01}^L, \quad b_{20} = b_{20}^L + \frac{8}{21} b_{10}^L,$$

$$b_{s2} = -\frac{4}{7} b_{10}^L, \quad b_{s1} = -\frac{1}{2} b_{10}^L, \quad b_{s1} = -\frac{5}{7} b_{10}^L.$$

(185)

This is the model we will consider hereafter in this section. As we will see shortly, a scale-dependent bias at linear order appears necessary to explain recent numerical measurements of halo bias with massive neutrinos.

### 4.2 Perturbative Approach to Cosmologies with Massive Neutrinos

Refs. [152, 153, 154] investigated the impact of massive neutrinos on the spatial distribution of dark matter halos and galaxies using large box-size N-body simulations that incorporate massive neutrinos as an extra set of particles. They found that massive neutrinos generate an additional scale-dependence in the halo power spectrum on midly nonlinear scales, in agreement with previous theoretical predictions [156], when defining the bias with respect to the whole dark matter component. On the other hand, if bias is defined with respect to the only cold dark-matter, this scale dependence is weaker. In this Section, we will compare the scale-dependence induced by massive neutrinos with that generated by gravitational mode-coupling and Lagrangian halo bias. We will show that the latter is substantially steeper, so that it should be relatively easy to isolate the contribution of massive neutrinos from a measurement of the halo power spectrum.

To model the impact of massive neutrinos on the clustering of dark matter halos, we follow [156, 157, 158, 159] and assume that the latter trace the cold Dark Matter (CDM) plus baryons fluctuation field, with a linear growth rate suppressed in a scale-dependent way by the massive neutrinos. This approximation is motivated by the smallness of the
neutrino masses we consider. For $\sum m_{\nu} < 0.6$ eV, the neutrino free-streaming scale is sufficiently large that the neutrino perturbations remain in the linear regime up to late time. It has been shown to work well in Refs. [152, 153, 154]. We thus write the total dark matter perturbation as a weighted sum of cold dark matter (we will ignore the effect of baryons in what follows, except for the fact that our CDM transfer function is a weighted sum of the baryons + CDM transfer functions) and neutrino fluctuations,

$$\delta_m = (1 - f_\nu)\delta_c + f_\nu \delta_\nu,$$  \hspace{1cm} (186)

where the neutrino overdensity $\delta_\nu$ is in the linear regime, and the neutrino fraction $f_\nu$ is

$$f_\nu = \frac{\Omega_\nu}{\Omega_c + \Omega_\nu}. \hspace{1cm} (187)$$

Replacing $\delta(x, t)$ by $\delta_c(x, t)$ in the right-hand side of Eq.(184), the halo-mass cross-power spectrum reads

$$P_{hm}(k) = (b_{10} + b_{01}k^2) P_{\text{cm}}^{\text{NL}}(k) + \Delta P_{hm}(k) + P_{cc}(k)I_3(k), \hspace{1cm} (188)$$

where

$$P_{\text{cm}}^{\text{NL}} = \frac{P_{\text{cm}}^{\text{NL}} - f_\nu P_{\text{cm}}}{1 - f_\nu} = \frac{P_{\text{cm}}^{\text{NL}} - f_\nu (1 - f_\nu) P_{cc} - f_\nu^2 P_{\nu\nu}}{1 - f_\nu}, \hspace{1cm} (189)$$

and

$$\Delta P_{hm}(k) = (1 - f_\nu) b_{20} \int \frac{d^3 q}{(2\pi)^3} P_{cc}(q) P_{cc}(|\vec{k} - \vec{q}|)F_2(q, \vec{k} - \vec{q}) + (1 - f_\nu) b_{s2} \int \frac{d^3 q}{(2\pi)^3} P_{cc}(q) P_{cc}(|\vec{k} - \vec{q}|)F_2(q, \vec{k} - \vec{q})S(q, \vec{k} - \vec{q}) \hspace{1cm} (190)$$

Here, $P_{cc}$, $P_{\nu\nu}$ and $P_{c\nu}$ are the linear CDM, neutrinos power spectrum and the CDM-neutrinos cross-power spectrum, respectively. We have adopted the notation of Refs. [64, 160] for our definition of

$$I_3(k) = \frac{32}{105} (1 - f_\nu) \left( b_{s1} - \frac{5}{2} b_{s2} + \frac{16}{21} b_\psi \right) \int \text{d}lnr \Delta^2_{cc}(kr) I_R(r), \hspace{1cm} (191)$$

being

$$I_R(r) = I(r) + \frac{5}{6}, \hspace{1cm} (192)$$

$$I(r) = \frac{105}{32} \int_{-1}^1 d\mu D_2(k - q, k)S(q, k - q)$$
and

\begin{align*}
F_2(q_1, q_2) &= \frac{5}{7} + \frac{1}{2} \frac{q_1 \cdot q_2}{q_1 q_2} \left( \frac{q_1}{q_2} + \frac{q_2}{q_1} \right) + \frac{2}{7} \left( \frac{q_1 \cdot q_2}{q_1 q_2} \right)^2, \\
S(q_1, q_2) &= \left( \frac{q_1 \cdot q_2}{q_1 q_2} \right)^2 - \frac{1}{3} , \\
D_2(q_1, q_2) &= \frac{2}{7} \left[ S(q_1, q_2) - \frac{2}{3} \right].
\end{align*}

The latter are the second-order perturbative kernels. Note that, below the neutrino free-streaming scale, the cross-power spectrum \( P_{\text{cm}}^{\text{NL}} \) is enhanced by a factor of \((1 - f_\nu)^{-1}\) relative to \( P_{\text{mm}}^{\text{NL}} \). Following Ref. [64], one should interpret \( b_{\text{st}} \) etc. as “renormalized” bias parameters. Eq. (173) leads to the following relation

\begin{equation}
\frac{32}{105} \left( b_{\text{st}} - \frac{5}{2} b_{\text{st}} + \frac{16}{21} b_\psi \right) = \frac{32}{315} b_{10}^{L} \equiv b_{\text{NL}}.
\end{equation}

Further details regarding the evaluation of these integrals can be found in Appendix \S\ 11.2.

The nonlinear mass power spectrum \( P_{\text{mm}}^{\text{NL}} \) and the linear \( P_{\text{cc}}, P_{\nu\nu}, P_{\nu v} \), Eqs. (188)-(189) and (190), in the presence of massive neutrinos are obtained from the CLASS code [161] (see also Ref. [162] for an overview). We refer the reader to [163] for details about this implementation.

### 4.3 Bias Parameters

To evaluate the halo-matter power spectrum Eq. (188), we need predictions for the values of the bias parameters \( b_{10}^{L}, b_{01}^{L}, b_{20}^{L} \), together with the scale \( R_v \) that quantifies the magnitude of the \( k \)-dependent velocity bias.

For this purpose, we compute all the Lagrangian bias factors using the excursion set peak (ESP) mass function (see Chapter \S\ 2), which has been shown to agree well with simulated halo mass functions constructed with a spherical overdensity (SO) criterion, Figure 2. Following [157, 154], we replace the average mass density \( \rho_m \) by \( \rho_{\text{cdm}} \) and the variance of mass fluctuations \( \sigma_m \) by \( \sigma_{\text{cc}} \) in the ESP halo mass function. While we refer the reader to the aforementioned references for details, it is worth stressing the following points:

1. The bias parameters depend both on redshift and halo mass. To follow the analysis of [152, 153, 154] as closely as possible, we average the Lagrangian bias factors over a suitable range of halo masses,

\begin{equation}
b_{ij} = \frac{\int_{M_{\text{min}}}^{M_{\text{max}}} b_{ij}^L(M) \tilde{n}_{\text{ESP}}(M)/M \, dM}{\int_{M_{\text{min}}}^{M_{\text{max}}} \tilde{n}_{\text{ESP}}(M)/M \, dM},
\end{equation}

90
where \( n_{\text{ESP}} \) is the ESP halo mass function, Eq. (66). The mass range is \( 2 \times 10^{13} h^{-1} M_\odot < M < 3 \times 10^{15} h^{-1} M_\odot \). In Table 1 some typical values for the bias parameters for different neutrino masses.

\[
\begin{array}{|c|c|c|c|c|}
\hline
\sum m_\nu (\text{eV}) & b_{10}^L & b_{01}^L & R_v^2 & b_{20}^L \\
\hline
0 & 0.68 & 12.32 & 10.41 & -0.32 \\
0.1 & 0.73 & 12.43 & 10.38 & -0.28 \\
0.2 & 0.80 & 12.55 & 10.33 & -0.22 \\
0.3 & 0.87 & 12.68 & 10.29 & -0.15 \\
0.6 & 1.15 & 13.13 & 10.17 & 0.24 \\
\hline
\end{array}
\]

Table 2: Lagrangian bias factors as computed from the ESP mass function for the cosmological models considered here (they are labeled according to the sum of neutrino masses). The bias parameters, defined relative to the linear density field extrapolated at \( z = 0 \), are weighted over the mass range \( 2 \times 10^{13} h^{-1} M_\odot < M < 3 \times 10^{15} h^{-1} M_\odot \).

2. Here and henceforth, we will use the word “local” when we refer to the simplest bias prescription without any scale dependence. In this case, we set \( b_{s2} = b_\psi = b_\mathrm{st} = b_{01} = 0 \). Conversely, our scale-dependent bias prescription is summarized by Eq. (185) with \( R_v \) computed from Eq. (106). We only retain the nonlocal peak bias \( b_{01}^L \) since i) it is the only nonlocal Lagrangian term for which we have the time evolution and ii) several second-order peak bias factors induce \( k^2 \)-corrections (see e.g. [65, 103]) that are at least partly degenerate with \( (b_{01}^L - R_v^2)k^2 \) over the range of wavenumber considered here.

3. The scale dependencies induced by the peak constraint also propagate to \( b_{s2}, b_\psi, b_\mathrm{st} \) etc. for which the Lagrangian to Eulerian mapping can be derived upon analyzing terms like Eq. (183). Since \( b_{s2} \) etc. are small however, we do also not expect significant corrections on the scales of interest.

To compare our predictions with the numerical results of [152, 153, 154], we use a set of \((\Omega_m, \Omega_\nu)\) such that the total mass density \( \Omega_m = \Omega_c + \Omega_\nu \) is held fixed to \( \Omega_m = 0.2708 \) while the CDM and neutrino density are varied. The sum of neutrino masses is \( \sum m_\nu = 0, 0.1, 0.2, 0.3 \) and \( 0.6 \text{ eV} \), such that \( \Omega_\nu \) varies between 0 and 0.0131. We have also adopted \( h = 0.7 \) for the Hubble rate, \( n_s = 1 \) for the scalar spectral index, and \( A_s = 2.43 \times 10^{-9} \).
non local halo bias with and without massive neutrinos

Figure 10: Halo bias at redshift $z = 0$ as a function of wavenumber in the case $\sum m_\nu = 0$. The different terms that contribute to the scale-dependence bias in Eq. (271) are labelled according to the bias parameter they are proportional to (see text). The solid black curve represents the sum of all the contributions, Eq. (271). All the bias factors have been computed consistently from the ESP halo mass function and the relations Eq. (185).

for the amplitude of primordial scalar perturbations. The normalization amplitude $\sigma_8$ changes in accordance with the sum of neutrino masses $^2$.

4.4 RESULTS

The quantity we focus on is the halo bias defined as the ratio of the halo-matter cross-power spectrum to the matter auto-power spectrum,

$$
\frac{P_{hm}}{P_{mm}} = \frac{P_{hm}(k)}{P_{mm}(k)} = \frac{b_{10} + b_{01}k^2}{P_{mm}(k)} + \Delta P_{hm}(k) + P_{cc}(k) I_3(k).
$$

$^2$ The simulations were seeded with a transfer function output from a Boltzmann code at $z = z_i$ rather than at $z = 0$, where $z_i$ is the starting redshift, and subsequently ignore the contribution of radiation to the Hubble rate. Therefore, for the sake of comparison we set $\rho_\gamma = 0$ in the CLASS code when $0 \leq z \leq z_i$. This translates into a few percent increase in the linear growth rate at $z = 0$, and mainly has an impact on our prediction of $b_{10}$.

92
Figure 11: Halo bias (left) and fractional scale-dependence (right) at $z = 0$ as a function of wavenumber for values of $\sum m_\nu = 0, 0.1, 0.2, 0.3$ and 0.6 eV. In the left panel, only the local bias terms are included in the predictions. In the right panel, the models with non-zero neutrino masses still assume local bias, whereas the solid (black) curve represents the case in which massive neutrinos are absent but all non-local terms are accounted for. The non-local bias contributions induced by gravity and by the peak constraint generate a sharp rise beyond $k \sim 0.1 \hMpc$ substantially steeper than the effect of non-zero neutrino mass.

It is expected to be constant (matching the value of $1 + b_{10}^{L}$) on large scales, with a scale dependence arising at smaller scales. The origin of this scale-dependence is twofold: the non-locality discussed in Sec. §4.1 and the suppression of the linear growth rate induced by the presence of massive neutrinos. Therefore, we begin by exploring the contributions generated by the non-locality arising from gravity and the peak constraint.

Fig. 10 displays the various scale-dependent contributions to the halo bias Eq. (271) when the neutrinos are massless, i.e. $\sum m_\nu = 0$. We have labelled the curves according to the bias parameters that weight each scale-dependent contribution. Namely, $b_{20}$ and $b_{22}$ denote the two terms of Eq.(190) while $b_{NL}$, defined in Eq. (194), indicate the term proportional to $I_{3}(k)$ in Eq. (191). Note that $P_{\text{NL}}^{CN}(k) = P_{\text{NL}}^{mm}(k)$ in this case. Clearly, the dominant contribution arises from the peak constraint through the curvature term $(b_{01}^{L} - R_{v}^{2})k^{2}$, which is of positive sign for the mass considered here (see Table 1). It is only partly compensated by the other ones, which are negative or close to zero.

The left panel of Fig. 11 displays the scale-dependence of the halo bias at $z = 0$ as a function of neutrino masses. For the sake of illustration, we have assumed the simplest local bias model specified above, in which only $b_{10} = 1 + b_{10}^{L}$ and $b_{20} = b_{20}^{L} + 8/21 b_{10}^{L}$.
Figure 12: Halo bias at $z = 0$ as a function of wavenumber. Data points are from [152]. In the upper left panel, we show for $\sum_i m_{\nu_i} = 0.3$ eV the local bias prediction as the dashed (blue) curve, and our full non-local model as the solid (orange) curve. The difference between the solid (orange) and the dotted (green) curve represents the effect of turning off the contribution $b_{01}k^2$ arising from the peak constraint. In the upper right panel, we compare our non-local prediction with the numerical data for $\sum_i m_{\nu_i} = 0$, $0.3$ and $0.6$ eV. The lower panels show the fractional deviation between theory and simulations. Note that these predictions have no free parameter.

are different from zero. In the right panel of Fig.11, we show the scale-dependence relative to the $\sum m_{\nu} = 0$ case. This should be compared to the solid (black) curve, which represents the scale-dependence obtained when massive neutrinos are absent but all the non-local terms are included. These terms generate a sharp rise beyond $k \sim 0.1 \, h \, \text{Mpc}^{-1}$ which is much steeper than the effect of varying the sum of neutrino masses. Therefore, a measurement of the scale-dependence of bias over a range of wavenumbers should help disentangling the contribution of massive neutrinos from that induced by the non-local terms.

In Figs 12 and 13, we compare our predictions for the halo bias at $z = 0$ and $0.5$ with the N-body measurements of [152]. In both figures, the left panel displays the case of a neutrino sum of $0.3$ eV. The dashed (blue) curve is the prediction in the simplest local bias model, which falls short of explaining the rise on scales $k \gtrsim 0.1 \, h \, \text{Mpc}^{-1}$. Our full non-local prediction shown as the solid (red) curve agrees with the numerical data within $3\%$ for $k \lesssim 0.3 \, h \, \text{Mpc}^{-1}$. Most of the difference with the local bias prediction arises from including the peak constraint through the contribution $b_{01}k^2$, which is the difference between the solid (red) and dotted (magenta) curve. The right panel compares our full non-local model with the numerical results for $\sum_i m_{\nu_i} = 0$, $0.3$ and $0.6$ eV. In all
cases the agreement is at the $\sim 3\%$ level down to $k = 0.3 \, h \, \text{Mpc}^{-1}$. We emphasize again that our theoretical predictions have no free parameter: all the bias factors are consistently determined from the ESP halo mass function and from the relations Eq.\((185)\). However, the fact that the magnitude of the $k^2$ correction is consistent with $b_{L01}^2 - R_v^2$ may be coincidental as we expect similar contributions from higher-order Lagrangian bias. Since we do not have as yet Eulerian expressions for these additional bias contributions, we defer a thorough discussion of this issue to future work.
Part II

CONSTRAINING PRIMORDIAL NON-GAUSSIANITY THROUGH THE LARGE SCALE STRUCTURE OF THE UNIVERSE
Inflation (see [164] for a review) has become the dominant paradigm for understanding the initial conditions for the large scale structure (LSS) formation and for the Cosmic Microwave Background anisotropy (CMB). In the inflationary picture, primordial density fluctuations are created from quantum fluctuations “redshifted” out of the horizon during an early period of superluminal expansion of the Universe, where they are “frozen”. Perturbations at the surface of last scattering are observable as temperature anisotropy in the CMB. The last and most impressive confirmation of the inflationary paradigm has been recently provided by the data of the Planck mission which has marked a crucial step in the precision era of the CMB measurements in space ([32]).

Despite the simplicity of the inflationary paradigm, the mechanism by which the cosmological curvature perturbation is generated is not yet fully established. In the single-field models of inflation, the observed density perturbations are induced by fluctuations of the inflaton field itself. An alternative to the standard scenario is represented by the curvaton mechanism ([165], [166], [167]) where the final curvature perturbations are produced from an initial isocurvature perturbation associated to the quantum fluctuations of a light scalar field (other than the inflaton), the curvaton, whose energy density is negligible during inflation. The curvaton isocurvature perturbations are transformed into adiabatic ones when the curvaton decays into radiation much after the end of inflation. Alternatives to the curvaton model are those models characterised by the curvature perturbation being generated by an inhomogeneity in the decay rate ([168], [169]) of the particles responsible for the reheating after inflation. Other opportunities for generating the curvature perturbation occur at the end of inflation ([170], [171]) and during preheating ([172]). A precise measurement of the spectral index $n_s$ of the comoving curvature perturbation $\zeta$ provides a powerful constraint inflation models. However, both single-field models and alternative mechanisms, like the curvaton, predict a value of the spectral index very close to unity. Furthermore, the lack of a gravity-wave signal in CMB anisotropies would not give us any information about the perturbation generation mechanism, since alternative mechanisms predict an amplitude of gravity waves far too small to be detectable by future experiments aimed at observing the $B$-mode of the CMB polarisation.
There is, however, a third observable which will prove fundamental in providing information about the mechanism chosen by nature to produce the structures we see today. It is the deviation from a Gaussian statistics, i.e., the presence of higher-order connected correlation functions of the perturbations. Indeed, a possible source of non-Gaussianity (NG) could be primordial in origin, being specific to a particular mechanism for the generation of the cosmological perturbations (for a review see [173]). This is what makes a positive detection of NG so relevant: it might help discriminating among competing scenarios which, otherwise, would might remain indistinguishable.

While a full characterization of Gaussian statistics, by definition, require the only study of the two-point correlation function of the fluctuations, non Gaussian ones retain statistical information in higher order correlation functions. Detecting a possible primordial source of non-Gaussianity in the cosmological perturbations is one of the main targets of current and future experiments measuring the properties of the CMB anisotropies and of the large-scale structure. Being CMB fluctuations still in the linear regime, any measured deviation from Gaussianity must be of primordial origin. On the other hand, perturbations at late time have become very non-linear and consequently, even if the primordial perturbations were Gaussian distributed, non-Gaussianity is generated, so that the study of primordial signatures of non-Gaussianity in the large scale structure of the late Universe is more challenging. Despite this complication, the properties of the clustering of galaxies have been identified to be a powerful probe of NG thanks to the fact that NG introduces a scale-dependent bias between the power spectra of halos and dark matter [49, 50, 51] (see Sec. §5.3).

Different mechanisms for the production of primordial perturbations may generate non-Gaussianity which peaks at different shapes of the correlation functions. For example, single-field models may generate a sizable non-Gaussianity which peaks at equilateral shapes of the bispectrum, i.e. in configurations in which the three momenta have the same amplitude [174]. In this thesis, we consider models that produce primordial comoving curvature perturbations characterized by the so-called “local”-type of NG, defined as

$$\zeta(x) = \zeta_G(x) + \frac{6}{5} f_{\text{NL}} \left[ \zeta_G^2(x) - \langle \zeta_G^2(x) \rangle \right],$$

(197)

where $\zeta_G$ is Gaussian distributed and we have introduced the non-linearity parameter, $f_{\text{NL}}$. Hence $f_{\text{NL}}$ is defined in terms of the three-point correlator, the bispectrum, of the primordial comoving curvature perturbation in the so-called squeezed limit

$$f_{\text{NL}} = \frac{5}{12} \frac{\langle \zeta_{k_1} \zeta_{k_2} \zeta_{k_3} \rangle}{P_{k_1} P_{k_2} P_{k_3}} \quad (k_1 \ll k_2 \sim k_3),$$

(198)

while the four-point correlator, the trispectrum, defines another non-linearity parameter, $\tau_{\text{NL}}$, in the so-called collapsed limit

$$\tau_{\text{NL}} = \frac{1}{4} \frac{\langle \zeta_{k_1} \zeta_{k_2} \zeta_{k_3} \zeta_{k_4} \rangle}{P_{k_1} P_{k_2} P_{k_3} P_{k_4}} \quad (k_{12} \simeq 0).$$

(199)
We have normalised the correlators with respect to the power spectrum of the curvature perturbation,
\[
\langle \zeta_{\vec{k}_1} \zeta_{\vec{k}_2} \rangle = (2\pi)^3 \delta(\vec{k}_1 + \vec{k}_2) P^\zeta_{\vec{k}_1}
\]
and used the notation \( \vec{k}_{ij} = (\vec{k}_i + \vec{k}_j) \). In all single-field models of inflation the bispectrum is suppressed in the squeezed limit and is non vanishing only when the spectral index deviates from unity, \( f_{\text{NL}} = 5/12(1 - n_s) \simeq 0.02 \) (see [175], [176],[177], [178]). A convincing detection of NG in the squeezed limit, \( f_{\text{NL}} \gg 1 \), would therefore rule out all single-field models (one should be aware though that, in single-field models of inflation, a large NG can be generated in shapes others than the squeezed, e.g. in the equilateral configurations). However, such a detection would not rule out multi-field models of inflation where the NG is seeded by light fields other than the inflaton.

The parameter \( f_{\text{NL}} \) is currently constrained by the Planck mission to be in the range \((-0.8 \pm 5.0) \) at 68% CL [179, 56] and \((28 \pm 23) \) by the large-scale structure [51], while the parameter \( \tau_{\text{NL}} \) needs to be in the range \( (\tau_{\text{NL}} < 2800) \) at 95% CL [180].

**Plan of the chapter** This chapter works as an introduction to the subsequent analyses of Chapters \( \S 6 \) and \( \S 7 \). In Sec. \( \S 5.1 \) we introduce the Suyama - Yamaguchi inequality, a useful relation between \( f_{\text{NL}} \) and \( \tau_{\text{NL}} \). In Sec. \( \S 5.2 \) we discuss the possibility that the non Gaussianity parameters have some scale-dependence. In Sec. \( \S 5.3 \) we review the most promising way of constraining these parameters through observations of the large scale structures of the Universe, through the scale dependence of halo bias on large scales.

### 5.1 The Suyama - Yamaguchi Inequality

Measuring the amplitudes of both the bispectrum and the trispectrum is extremely interesting as, if only one degree of freedom is responsible for the perturbations, then there is a well-defined relation between the NG parameters, \( \tau_{\text{NL}} = \left( \frac{6}{5} f_{\text{NL}} \right)^2 \). On the contrary, if more than one field is responsible for the cosmological perturbations generated through the inflationary dynamics, then there exists an inequality, the so-called Suyama-Yamaguchi (SY) inequality ([52], see also [181], [182] and more recently [183, 184]). Based on the conditions that 1) scalar fields are responsible for generating curvature perturbations and that 2) the fluctuations in the scalar fields at the horizon crossing are scale invariant and Gaussian, Suyama and Yamaguchi proved the inequality

\[
\tau_{\text{NL}} \geq \left( \frac{6}{5} f_{\text{NL}} \right)^2.
\]

The condition 2) amounts to assuming that the connected three- and four-point correlations of the light fields vanish and that the NG is generated at super-horizon scales.
This is quite a restrictive assumption. However, based on the operator product expansion, which is particularly powerful in characterising in their full generality the squeezed limit of the three-point correlator and the collapsed limit of the four-point correlator, it was shown that the SY inequality holds also for NG light fields ([183]). This is consequence of fundamental physical principles (like positivity of the two-point function) and its hard violation would require some new non-trivial physics to be involved.

The observation of a strong violation of the inequality would then have profound implications for inflationary models. It would imply either that multi-field inflation cannot be responsible for generating the observed fluctuations independently of the details of the model, or that some new non-trivial (ghost-like) degrees of freedom play a role during inflation ([183]).

5.2 THE RUNNING OF PRIMORDIAL NON-GAUSSIANITY

Even though the definitions (198) and (199) are widely used to model NG in the primordial perturbations, it is just the first step one can make on this matter. One, more general, definition of the bispectrum and trispectrum could include a scale-dependence in the non-linearity parameters \( f_{NL} \) and \( \tau_{NL} \). This step is well-motivated by the theoretical predictions of some models [185, 186, 187, 188, 189]. The running with physical scale of the NG parameters \( f_{NL} \) and \( \tau_{NL} \) has been the subject of an intense recent research [190, 191, 192, 193, 194, 195, 196, 197].

To account for the running of \( f_{NL} \) in its full generality one can adopt for example the parametrization used in Ref. [198] (see also Ref. [187])

\[
B_\xi(k_1, k_2, k_3) = \frac{6}{5} \left[ \zeta_{f_{NL}, m}(k_3) \xi_m(k_1) \xi_m(k_2) P_\xi(k_1) P_\xi(k_2) + \text{cyc.} \right],
\]

where

\[
\zeta_{f_{NL}, m}(k) = \zeta_{f_{NL}, m}(k_0) \left( \frac{k}{k_0} \right)^{n_{f_{NL}, m}}.
\]

Here \( \zeta_{f_{NL}}(k) \) parametrizes the (self-)interactions of the fields and \( \xi_{m}(k) \) the ratio of the contribution of each field, being \( n_{m} \) the associated index. From this general parametrization, we can also easily extend the one for the trispectrum

\[
T_\xi(k_1, k_2, k_3, k_4) = 4 \left[ \zeta_{\tau_{NL}, m}(k_5, k_4) \xi_m(k_1) \xi_m(k_2) \xi_m(k_3) P_\xi(k_1) P_\xi(k_2) P_\xi(k_3) P_\xi(k_4) + \text{cyc.} \right],
\]

where

\[
\zeta_{\tau_{NL}, m}(k_i, k_j) = \zeta_{\tau_{NL}, m}(k_0) \left( \frac{k_i k_j}{k_0^2} \right)^{n_{\tau_{NL}, m}}.
\]
In the single-field limit, $\zeta_{NL}(k_i, k_j) = \frac{36\pi}{25} \zeta_{f_{NL}}(k_i) \zeta_{f_{NL}}(k_j)$ and $\xi_{m}(k) = 1$. According to this parametrization, in the case of a multi-field inflation, we have three free parameters, $n_{f_{NL}}, n_{m}$ and $n_{\tau_{NL}}$, which describe the scale dependence of the non-linearity parameters $f_{NL}$ and $\tau_{NL}$ and of the dimensionless power spectra. In order to decrease the complexity of the analysis, from now on we make the assumption that $n_{m}$ is significantly smaller than unity. By doing so, we are left with the following parametrization of the non-linear parameters

$$f_{NL}(k) = f_{NL}^* \left( \frac{k}{k_*} \right)^{n_{f_{NL}}},$$

(206)

and

$$\tau_{NL}(k_i, k_j) = \tau_{NL}^* \left( \frac{k_i k_j}{k_*^2} \right)^{n_{\tau_{NL}}}. \quad (207)$$

A local bias analysis performed in Ref. [190] showed that high-redshift surveys ($z > 1$) covering a large fraction of the sky corresponding to a volume of about $100 h^{-3}$ Gpc$^3$ might provide a $1\sigma$ error on the running $f_{NL}$ parameter of the order of $0.4(50/f_{NL})$.

In the case in which the perturbations are sourced by a single field, then a well-defined relation between the running spectral indices holds,

$$n_{f_{NL}} = n_{\tau_{NL}} \quad (208)$$

and the indices are therefore not independent. It is interesting however to consider the possibility that $f_{NL}$ and $\tau_{NL}$, and therefore their spectral indices too, are not related to each other, thus leaving open the possibility that the perturbations are originated from a multi-field scenario.

## 5.3 Scale Dependent Bias from the Inflationary Bispectrum

Upcoming large scale structure surveys will take over the hunt for primordial non-Gaussianity (PNG) from CMB experiments. The recent (individual) limits on the non-linear parameter $f_{NL}$ from measurements of galaxy clustering and the integrated Sachs-Wolfe (ISW) effect are already at the level of the CMB pre-Planck constraints, i.e. $\Delta f_{NL} \sim 80$ (see [199, 200, 201]). Forecasts for future Euclid-like galaxy surveys show that a measurement of the large scale galaxy power spectrum alone can constrain $f_{NL} \sim$ a few (e.g. [196, 202, 203, 204, 205]), whereas intensity mappings of the 21cm emission line of high-redshift galaxies could achieve $\Delta f_{NL} \sim 1$ (e.g. [206]).

One of the most powerful large scale structure probes of PNG to date is the galaxy power spectrum. The most prominent effect is a scale dependence at very large scales.
proportional to \( f_{NL} \), which was first observed in N-body simulations in [49]. The amplitude \( b_{NG} \) of the scale-dependent bias induced by a primordial bispectrum of the local shape (i.e. \( f_{NL} \xi^2 \) as in Eq. (197)) was found to be proportional to the first-order bias, i.e. \( b_{NG} = \delta_c b_1 \) and exhibiting a distinct scale dependence at large scales [50, 51].

Using the peak-background split argument (see Sec. §1.2) [51] argued that the amplitude of the non-Gaussian bias is proportional to the logarithmic derivative of the halo mass function w.r.t. \( \sigma_8 \) (or any proxy for the normalisation amplitude), i.e. \( b_{NG} = b_{NG}^{\text{pbs}} \), where

\[
b_{NG}^{\text{pbs}} = \frac{\partial \ln n_h}{\partial \ln \sigma_8}.
\]  

[207] generalised the peak-background split approach to include the non-Markovian (in the excursion set sense) and non-universality of the mass function. Given the importance of this effect in what we discuss in the next chapters, we spend some time in reviewing its theoretical derivation, following [51].

The starting point of a peak-background split ansatz is separating long- and short-wavelengths mode of the density field \( \delta \), treating the former as a constant amplitude rescaling of the latter, which determines the small scales processes. The basic assumption for being able to perform such a separation is that the two regimes are decoupled. This is the case for Gaussian distributed fields, as we considered in Sec. §1.2. In this second part of the thesis, we want to consider the possibility of the primordial perturbation to exhibit a certain (small) degree of primordial non-Gaussianity (PNG), to be distinguished by the non-Gaussianity generated on the density field \( \delta \) at late time by the non-linear gravitational interactions.

By definition, the scales of a non-Gaussian field are coupled, this is evident in, e.g., Eq. (197). In this framework, it can be convenient to work with the primordial gravitational perturbation \( \Phi(x) = \frac{3}{5} \zeta(x) \), so that the local model for PNG reads

\[
\Phi(x) = \phi_G(x) + 2 f_{NL} (\phi_G^2(x) - \langle \phi_G^2 \rangle).
\]  

The density field is connected to \( \Phi(x) \) through the Poisson equation,

\[
\delta(k) = \mathcal{M}(k) \Phi(k),
\]  

where

\[
\mathcal{M}(k) = \frac{2c^2 k^2 T(k) D(z)}{3 \Omega_m H_0^2}
\]  

and \( c \) is the speed of light, \( T(k) \) is the transfer function, \( \Omega_m \) is the ratio of the matter density component over the critical density today, \( H_0 \) is the Hubble parameter today and \( D(z) \) is the linear growth factor at redshift \( z \). If the initial conditions are non-Gaussian, as in Eq. (210), the separation of scales has to be performed not at the level of the density field \( \delta \), but on Gaussian potential fluctuations\footnote{Let us drop the subscript \( G \), it is understood that \( \phi \) indicates the Gaussian perturbation and \( \Phi \) the non-Gaussian one.}, \( \phi_G \),

\[
\phi = \phi_L + \phi_S.
\]
and consequently we get

\[ \Phi = \phi_L + f_{NL} \phi_L^2 + (1 + 2f_{NL}\phi_L)\phi_S + f_{NL}\phi_S^2 + \text{const.} \]  

(214)

On sufficiently large scales, the relation

\[ \delta_L(k) = \mathcal{M}(k)\phi_L(k) \]  

(215)

holds, while the short-wavelength modes of the density field are affected, at first order, as

\[ \delta_S \approx \mathcal{M}(1 + 2f_{NL}\phi_L)\phi_S, \]  

(216)

which may be interpreted as a local modulation of the amplitude of matter fluctuations by the long modes proportional to \( f_{NL} \),

\[ \sigma_S \rightarrow \sigma_S(1 + 2f_{NL}\phi_L) \equiv \hat{\sigma}_S. \]  

(217)

As a result, the local number density of halos of mass \( M \) is not a function of \( \delta_L \) only, but also of this local modulation, so that Eq. (21) is modified, at first order, into

\[ \delta_h(x) \approx -\frac{1}{n_h} \frac{d\bar{n}_h}{d\delta_L} \delta_L(x) + 2f_{NL} \frac{d\phi_L}{d\delta_L} \frac{d\ln \bar{n}_h}{d\ln \sigma_8} \delta_L(x) + ... , \]

\[ = b_1^G \delta_L(x) + 2f_{NL} \frac{d\phi_L}{d\delta_L} b_{1,NG} \delta_L(x) + ... \]  

(218)

where \( \sigma_8 \) is the variance of the density field for a smoothing at \( R = 8 \) Mpc/h, the typical reference amplitude of matter fluctuations and we define the first order bias in Fourier space in the presence of primordial non-Gaussianity

\[ b_1(M, k) \equiv b_1^G(M) + \Delta b_{1,NG}(M, k) = b_1^G(M) + \frac{2f_{NL}}{\mathcal{M}(k)} b_{1,NG}^G(M) \]  

(219)

being

\[ b_{1,NG} \equiv \frac{\partial \ln \bar{n}_h}{\partial \ln \sigma_8}. \]  

(220)

This result is general, at first order, as we did not make assumptions other than the scale separation in \( \phi_G \). For universal mass functions of the form of Eq. (15),

\[ \bar{n}_h(M) = \frac{\rho}{M^2} f(v) \frac{d\ln v}{d\ln M}, \]  

(221)

as in the case of the Press & Schechter mass function, Eq. (14), the logarithmic derivative of the mass function with respect to the normalization amplitude \( \sigma_8 \) takes the simple form

\[ \frac{\partial \ln \bar{n}_h}{\partial \ln \sigma_8} = \frac{\partial \ln \bar{n}_h}{\partial \ln v} \frac{d\ln v}{d\ln \sigma_8} = -\frac{v}{\bar{n}_h} \frac{\partial \bar{n}_h}{\partial v} = \delta_{sc} b_1^G, \]  

(222)
The effect summarized above can be computed through various methods such as high peaks ([50], [198]) or multivariate bias expansions ([208], [209]) but, to date, the peak-background split provides the most accurate estimate of the effect ([51], [210], [211], [212], [207]). A generalization of the non-Gaussian contribution to the linear bias, in the approximation of universal mass functions, including higher order correlation functions of $\Phi$ was calculated in [211],

$$\Delta b_1(k) = \frac{4}{(N-1)!} \frac{\mathcal{F}_s^{(N)}(k,z)}{\mathcal{M}_s(k,z)}$$

$$\times \left[ b_{N-2}\delta_{sc} + b_{N-3} \left( N - 3 + \frac{d\ln\mathcal{F}_s^{(N)}(k,z)}{d\ln\sigma_s} \right) \right] ,$$

where $b_N$ are Lagrangian bias parameters as in Eq. (22) and $z$ is the redshift and we are now stressing with the subscript $s$ that quantities are smoothed on some scale $R_s$. We have defined the function

$$\mathcal{F}_s^{(N)}(k,z) = \frac{1}{4\sigma_s^2 P_\Phi(k)} \left[ \prod_{i=1}^{N-2} \int \frac{d^3k_i}{(2\pi)^3} \mathcal{M}_s(k_i,z) \right] \mathcal{M}_s(q,z)$$

which is a projection factor whose $k$-dependence is dictated by the exact shape of the $N$-point function $\xi^{(N)}_\Phi$ of the gravitational potential. For the local constant-$f_{NL}$ model, the factor $\mathcal{F}_s^{(3)}$ is equal to $f_{NL}$ in the low $k$-limit (squeezed limit), so that the logarithmic derivative of $\mathcal{F}_s^{(N)}$ w.r.t. the rms variance $\sigma_s$ of the small-scale density field vanishes on large scales. For all other models of primordial non-Gaussianity however, this term is significant for most relevant peak heights and becomes negligible in the high peak limit only ([213]). Since the halo mass function may not be universal, the non-Gaussian bias correction should in principle be computed by taking derivative of the Gaussian halo mass function w.r.t. mass ([207]). However, because it is difficult to estimate such a mass derivative from real data, we will use Eq.(223), which is valid for a universal mass function. Nevertheless, one should bear in mind that non-universality can induce additional corrections at the $\sim 10\%$ level ([207, 214]). We discuss this case in Chapter §7. Note also that path integral extensions of the excursion set formalism (see [76]) suggest that memory terms (involving $N$-point correlators of the density field smoothed on any scale between $R_s$ and $R_l$) could also contribute at some level ([215, 216]).
Specialising the above result to the bispectrum and trispectrum shapes of Eqs. \((198)\) and \((199)\), the non-Gaussian bias correction reads

\[
\Delta b_1(k) = 2f_{\text{NL}} \frac{\delta_{sc} b_1}{\mathcal{M}_s(k)} + \frac{1}{2} \left( g_{\text{NL}} + \frac{25}{27} \tau_{\text{NL}} \right) \times \frac{\sigma_s^2 S^{(3)}_{s,\text{loc}}}{\mathcal{M}_s(k)} \left[ b_2 \delta_{sc} + b_1 \left( 1 + \frac{d \ln S^{(3)}_{s,\text{loc}}}{d \ln \sigma_s} \right) \right],
\]

where \(b_1, b_2\) are the first- and second-order Lagrangian bias parameters, \(g_{\text{NL}}\) is another NG coefficient parametrising the NG arising from a cubic third-order term in the curvature perturbation \(\zeta\) and \(S^{(3)}_{s,\text{loc}}(M)\) is the skewness of the density field in a local, quadratic non-Gaussian model with \(f_{\text{NL}} = 1\). Strictly speaking, this expression is valid in the limit \(k \ll 1\) only since we have ignored the \(k\)-dependence of \(\mathcal{F}_3^{(4)}\). However, deviations become significant only for \(k \gtrsim 0.1\) where the non-Gaussian signal is negligible and the signal-to-noise saturates ([217]). In linear theory, the product \(\sigma_s S^{(3)}_{s,\text{loc}}(M)\) is independent of redshift. Therefore, at fixed values of \(b_1\) and \(b_2\), the non-Gaussian correction induced by \(f_{\text{NL}}\) scales as \(D(z)^{-1}\), whereas that induced by \(g_{\text{NL}}\) and \(\tau_{\text{NL}}\) does not have any extra dependence on redshift.
Cosmological observations have the potential to test fundamental physics beyond what is accessible with laboratories on Earth. In particular, cosmological perturbations are believed to have been seeded during an inflationary phase which may have occurred at an energy scales potentially as high as $10^{14}$ GeV. The measurement of two-point correlation functions and of higher order clustering statistics can teach us about the interaction of the inflaton and the field content of the Universe during that period by constraining the PNG. The Planck satellite \[56\] has already put constraints on PNG, but there is still much space for interesting phenomenology, especially if we can access the regime where the amplitude of the PNG is an order of magnitude smaller than the current limits. This is precisely the goal of future missions designed to measure the three-dimensional structure of the Universe, such as the ESA mission EUCLID \[42\], the Square Kilometre Array \[218\] and the NASA SPHEREx mission \[205\], a proposed all-sky spectroscopic survey satellite.

In this chapter, we present some analyses on how well we can hope to constrain such signatures through the analysis of the Fisher information content from the two-point statistics of halos and DM. In particular, we discuss forecasts on the precision with which next generation galaxy redshift surveys will be able to constrain the non-linearity parameters $f_{\text{NL}}$ and $\tau_{\text{NL}}$ and their spectral indices $n_{f_{\text{NL}}}$ and $n_{\tau_{\text{NL}}}$.

**Plan of the chapter** This chapter is organized as follows. The first Sec. §6.1 deals with forecasting the precision with which we will be able to constrain the SY inequality with a EUCLID-like survey, while in Sec. §6.2 we estimate similar forecasts for the running of PNG, from measurements of the scale dependent bias.

### 6.1 Testing the Suyama-Yamaguchi inequality

Testing the SY inequality with future LSS observations and, therefore, the validity of multi-field inflationary models is the subject of this section.

Measurements of the galaxy power spectrum have been exploited to set limits on primordial non-Gaussianity competitive with those inferred from CMB observations (\[51\], \[109\])
As we have seen, a large value of $f_{NL}$ in the squeezed limit implies that the cosmological perturbations are generated within a multi-field model of inflation where the NG is sourced by light fields other than the inflaton. An inescapable consequence of the SY inequality, Eq. (201), is that the NG is also characterised by a large trispectrum in the collapsed limit. Therefore, investigations that take advantage of the scale-dependent effects of NG on the clustering of dark matter halos should in principle take into account both $f_{NL}$ and $\tau_{NL}$. However, since the contribution from the latter is suppressed by $10^{-4}(\tau_{NL}/f_{NL})$, setting limits on $f_{NL}$ under the assumption $\tau_{NL} = 0$, as done in the literature, should be a good approximation unless $|\tau_{NL}| \gg f_{NL}^2$.

We want to answer the following question: what values of $f_{NL}$ and $\tau_{NL}$ have to be measured in order to either confirm or disprove the SY inequality? As we shall see, even though the contributions from $f_{NL}$ and $\tau_{NL}$ are degenerate in the non-Gaussian halo bias, combining multiple halo mass bins can greatly help breaking the degeneracy. As we shall demonstrate, testing multi-field models of inflation at the 3-$\sigma$ level would require, for a EUCLID-like survey, a detection of a four-point correlator amplitude in the collapsed limit of the order of $\tau_{NL} \sim 10^5$ given a measurement of a local bispectrum at the level of $f_{NL} \sim 10$. Conversely, we will argue that disproving multi-field models of inflation would require a detection of $|f_{NL}|$ at the level of $80$ or larger if dark matter halos can be resolved down to a mass $10^{10} M_\odot / h$.

### 6.1.1 The Fisher matrix formalism applied to galaxy surveys

An important feature of the NG bias correction we presented in Sec. §5.3, namely

$$\Delta b_1(k) = 2f_{NL} \frac{\delta_{sc} b_1}{M_s(k)} + \frac{1}{2} \left( g_{NL} + \frac{25}{27} \tau_{NL} \right) \times \frac{\sigma^2 S_{s,loc}^{(3)}}{M_s(k)} \left[ b_2 \delta_{sc} + b_1 \left( 1 + \frac{d \ln S_{s,loc}^{(3)}}{d \ln \sigma_s} \right) \right],$$

is that its scale-dependence is degenerate in $f_{NL}$, $\tau_{NL}$ and $g_{NL}$ in the large scale limit, since all the $k$-dependence is then located in $1/M_s(k) \sim 1/k^2$. This degeneracy can be partly broken by considering galaxy populations tracing halos of different mass and, possibly, at different redshifts. While recent studies have analysed the problem of detecting NG through future large-scale surveys combining a number of observational datasets with simple models where only $f_{NL}$ is nonzero and the other two nonlinear parameters are set to zero, we will assume here that both $f_{NL}$ and $\tau_{NL}$ are non-vanishing since we aim at testing the SY inequality, Eq. (201). We will however set $g_{NL}$ to zero.\footnote{Notice that this assumption also gets rid of potentially large one-loop corrections to the SY inequality ([184]). These corrections would be anyway below the errors we will estimate on $\tau_{NL}$ even for $g_{NL}$ as large as $10^6$.}
We refer the reader to [221] for a recent study in which both $f_{NL}$ and $g_{NL}$ are nonzero. Finally, it is worth mentioning that our Eq. (226) is different from the expression given in [222], who neglected the mass-dependence of the skewness.

In order to assess the ability of forthcoming experiments to test the SY inequality through the measurement of the large scale bias, we make use of the Fisher information content on $f_{NL}$ and $\tau_{NL}$ from the two-point statistics of halos and dark matter in Fourier space. The Fisher matrix formalism has been extensively applied to predict how well galaxy surveys will constrain the nonlinear parameter $f_{NL}$ (e.g., [49], [223], [224]). In particular, combining differently biased tracers of the same surveyed volume and weighting halos by mass can help mitigate the effect of cosmic variance and shot noise and, therefore, reduce the uncertainty on $f_{NL}$ ([225], [51], [226], [227]).

Here and henceforth, we closely follow the notation of [227] and define the halo overdensity in Fourier space as a vector, every element corresponding to halos with different mass bins

$$\delta_h = (\delta_h(M_1), \delta_h(M_2), \cdots, \delta_h(M_n))^\top.$$  \hfill (227)

Assuming the halos to be locally biased and stochastic tracers of the dark matter density field $\delta$, we can write the overdensity of halos as

$$\delta_h = \mathbf{b} \delta + \epsilon,$$ \hfill (228)

where $\mathbf{b}$ is a vector whose $i$-component is the (Eulerian) bias of the $i$-th sample,

$$b_i^E(k, M_i, z) = 1 + b_1(M_i, z) + \Delta b_1(k, M_i, z),$$ \hfill (229)

and $\epsilon$ is a residual noise-field with zero mean. We assume that it is uncorrelated with the dark matter.

Computing the Fisher information requires knowledge of the covariance matrix of the halo samples,

$$C_h = \langle \delta_h \delta_h^\top \rangle = \mathbf{b} \mathbf{b}^\top P + E.$$ \hfill (230)

The brackets indicate the average within a $k$-shell in Fourier space. $P = \langle \delta^2 \rangle$ is the nonlinear dark matter power spectrum which, on large scales, can be assumed independent of $f_{NL}$ and $\tau_{NL}$ and $E = \langle \epsilon \epsilon^\top \rangle$ is the shot-noise matrix. We will follow the general treatment of [227] and assume that $E$ is not simply diagonal with entries consistent with Poisson noise (see Sec. §6.1.2 for explicit expressions). In order to simultaneously constrain $f_{NL}$ and $\tau_{NL}$, it is pretty clear that at least two different halo samples are required to break some of the parameter degeneracies, since the bias coefficients $b_1$, $b_2$, the rms variance $\sigma_s$ and the skewness $S_{s,loc}^{(3)}$ have distinct mass dependencies (as is apparent from the numerical fits of [228] or [229]). More precisely, the Fisher matrix takes the following general form

$$F_{ij} = V_{\text{surv}} f_{\text{sky}} \int \frac{dk}{2\pi^2} \frac{k^2}{2} \text{Tr} \left( \frac{\partial C_h}{\partial \theta_i} C_h^{-1} \frac{\partial C_h}{\partial \theta_j} C_h^{-1} \right),$$ \hfill (231)
where \( \theta_i \) are the parameters whose error we wish to forecast. The integral over the momenta runs from \( k_{\text{min}} = \pi / (V_{\text{surv}})^{1/3} \) to \( k_{\text{max}} = 0.1 \, \text{Mpc}^{-1} / h \), where \( V_{\text{surv}} \) is the surveyed volume and \( f_{\text{sky}} \) is the fraction of the sky observed. For illustration purposes, we will adopt the specifications of an EUCLID-like experiment: \( V_{\text{surv}} f_{\text{sky}} = 25 \, \text{Gpc}^3 / h^3 \) at median redshift \( z = 0.7 \). In particular, we are not making use of the characteristics of the Euclid photometric/imaging survey. We will ignore the redshift evolution and assume that all the surveyed volume is at that median redshift. In principle however, it should be possible to extract additional information on the non-Gaussian bias from the redshift dependence of the survey. For a single mass bin, the four entries of the Fisher matrix have the same \( k \)-dependence at low-\( k \). As a consequence, the determinant is very close to zero and, therefore, yields large (marginalised) errors. In this case, it is impossible to test the SY inequality regardless the characteristics of the halo sample, unless one has some prior on one of the parameters.

In the general case of \( N \) halo populations, the entries of the halo covariance matrix read \((a, b = 1, \cdots, N)\)

\[
C_{ab} = b_i^E(k, M_a, z)b_j^E(k, M_b, z)P(k) + E_{ab}.
\]

The derivative of the halo covariance matrix with respect to some parameter \( \theta \) is

\[
\frac{\partial C_h}{\partial \theta} = (b_{\theta}b_{\theta}^\top + bb^\top)c_h^{-1}P + b_{\theta}^\top c_h^{-1}P.
\]

On inserting the expression (233) into Eq. (251), we can write down the Fisher matrix for two generic parameters \( \theta_i \) \((i = 1, 2)\) as

\[
F_{ij} = V_{\text{surv}} f_{\text{sky}} \int \frac{dk^2}{2\pi^2} \frac{P^2}{2} \text{Tr} \left[ (bb_i^\top + b_{\theta}b_{\theta}^\top)c_h^{-1}(bb_j^\top + b_{\theta}b_{\theta}^\top)c_h^{-1} \right].
\]

The elements of the Fisher matrix can be easily expressed in terms of the following quantities

\[
\alpha = b^\top c_h^{-1}b P
\]

\[
\beta_i = b_{\theta}^\top c_h^{-1}b_{\theta} P
\]

\[
\gamma_{ij} = b_{\theta}^\top c_h^{-1}b_{\theta} P.
\]

After some algebra, we obtain

\[
F_{ij} = \frac{\alpha \gamma_{ij} + \beta_i \beta_j + \alpha (\alpha \gamma_{ij} - \beta_i \beta_j)}{(1 + \alpha)^2},
\]
which generalises the calculation reported in [227]. Note that, in what follows, \( \theta_1 = f_{NL} \) and \( \theta_2 = \tau_{NL} \).

One should bear in mind the caveat that the present Fisher matrix analysis assumes Gaussian uncertainties, even though it is likely that the estimators \( f_{NL} \) and \( \tau_{NL} \) have non-Gaussian distributions. One possible way of testing this assumption would be to generate Monte-Carlo simulations of the halo samples, but this is beyond the scope of this analysis.

### 6.1.2 Halo model predictions

Even though the halo model makes a number of predictions that are not physically sensible (such as a white noise contribution in the limit \( k \to 0 \) of the cross halo-mass power spectrum), it was shown to furnish a very good fit to the eigenvalues and eigenvectors of the halo stochasticity matrix ([232]). In this model, the shot-noise matrix can be cast into the closed form expression \( E \),

\[
E = \bar{n} \mathbf{I} - \mathbf{b} \mathbf{M}^\top - \mathbf{M}^\top \mathbf{b}.
\] (238)

Here, \( \mathbf{M} = \mathbf{M}/\bar{n}_m - \mathbf{b} \langle n \mathbf{M}^2 \rangle/2\bar{n}_m^2 \), \( \mathbf{M} \) is a vector whose entries are the halo masses and \( n = n(M) \) is the number density of halos of mass \( M \). The Poisson expectation is recovered upon setting \( M = 0 \). In the limit of \( N \gg 1 \) halo mass bins with identical number density \( \bar{n} \), we can replace the scalar products by integrals. A straightforward calculation shows that the coefficients \( \alpha, \beta_i \) and \( \gamma_{ij} \) can be rewritten as

\[
\alpha = \frac{\langle b^2 \rangle}{\left( \bar{n}_\text{tot}^{-1} - \langle M b \rangle \right)^2 - \langle b^2 \rangle \langle M^2 \rangle} \bar{n}_\text{tot}^{-1} P
\] (239)

\[
\beta_i = \frac{\langle b b_i \theta \rangle \left( \bar{n}_\text{tot}^{-1} - \langle M b \rangle \right) + \langle b^2 \rangle \langle M b \theta \rangle}{\left( \bar{n}_\text{tot}^{-1} - \langle M b \rangle \right)^2 - \langle b^2 \rangle \langle M^2 \rangle} P
\] (240)

\[
\gamma_{ij} = \langle b_{\theta} b_{\theta} \rangle \bar{n}_\text{tot} P
\] (241)

\[
+ \frac{\langle b^2 \rangle \langle M b_{\theta} \rangle \langle M b_{\theta} \rangle + \langle M^2 \rangle \langle b b_{\theta} \rangle \langle b b_{\theta} \rangle}{\left( \bar{n}_\text{tot}^{-1} - \langle M b \rangle \right)^2 - \langle b^2 \rangle \langle M^2 \rangle} \bar{n}_\text{tot}^2 P
\]

\[
+ \frac{\langle b b_{\theta} \rangle \langle M b_{\theta} \rangle + \langle b b_{\theta} \rangle \langle M b_{\theta} \rangle}{\left( \bar{n}_\text{tot}^{-1} - \langle M b \rangle \right)^2 - \langle b^2 \rangle \langle M^2 \rangle} \bar{n}_\text{tot}^{-1} \bar{n}_\text{tot} P
\]

\[
+ \frac{\langle b^2 \rangle \langle M b_{\theta} \rangle \langle M b_{\theta} \rangle}{\left( \bar{n}_\text{tot}^{-1} - \langle M b \rangle \right)^2 - \langle b^2 \rangle \langle M^2 \rangle} \bar{n}_\text{tot}^{-1} P
\]

\[
+ \frac{\langle b b_{\theta} \rangle \langle M b_{\theta} \rangle + \langle b b_{\theta} \rangle \langle M b_{\theta} \rangle}{\left( \bar{n}_\text{tot}^{-1} - \langle M b \rangle \right)^2 - \langle b^2 \rangle \langle M^2 \rangle} \bar{n}_\text{tot}^{-1} \bar{n}_\text{tot} P
\]

\[
+ \frac{\langle b^2 \rangle \langle M b_{\theta} \rangle \langle M b_{\theta} \rangle}{\left( \bar{n}_\text{tot}^{-1} - \langle M b \rangle \right)^2 - \langle b^2 \rangle \langle M^2 \rangle} \bar{n}_\text{tot}^{-1} P
\]
where
\[
\langle xy \rangle \equiv \frac{1}{\bar{n}_{\text{tot}}} \int_{M_{\text{min}}}^{M_{\text{max}}} dM n(M) x(M) y(M) \quad (242)
\]
\[
\bar{n}_{\text{tot}} \equiv \int_{M_{\text{min}}}^{M_{\text{max}}} dM n(M) = N \bar{n} \quad (243)
\]

Here, \( n(M) \) is the halo mass function, which we assume to be of the \([71]\) form with \( p = 0.3, q = 0.73 \) and a normalisation \( A = 0.322 \). This yields \( \langle n M^2 \rangle / \rho_m^2 = 75.9 \text{Mpc}^3 / h^3 \) at redshift \( z = 0.7 \).

### 6.1.3 Results and conclusions

We first compute the uncertainties on \( f_{\text{NL}} \) and \( \tau_{\text{NL}} \) from two different tracer populations and for a shot-noise matrix consistent with Poisson noise, i.e. \( \mathbf{E} = \text{diag}(1/n(M_1), 1/n(M_2)) \).

We consider a nearly unbiased sample with average mass \( M \sim 10^{12} M_\odot / h \) and a high mass sample with \( M = 10^{14} M_\odot / h \). Table 3 summarizes the characteristics of these populations.
Table 3: Average host halo mass, number density, (Lagrangian) linear and quadratic bias factors for the low- and high-mass halo samples used in Fig. 14.

<table>
<thead>
<tr>
<th></th>
<th>( M(M_\odot/h) )</th>
<th>( \bar{n} \left(h^3\text{Mpc}^{-3}\right) )</th>
<th>( b_1 )</th>
<th>( b_2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Halo 1</td>
<td>( 10^{12} )</td>
<td>( 7 \times 10^{-4} )</td>
<td>0.2</td>
<td>-0.2</td>
</tr>
<tr>
<td>Halo 2</td>
<td>( 10^{14} )</td>
<td>( 3 \times 10^{-6} )</td>
<td>2.5</td>
<td>4.5</td>
</tr>
</tbody>
</table>

For a given mass \( M \), the second-order Lagrangian bias parameter \( b_2(M) \) is computed from the Sheth-Tormen multiplicity function, whereas the skewness \( S_{s,\text{loc}}^{(3)} \) is computed from the phenomenological relation given in Sec. §6.1.1. Fig. 14 shows the resulting 68, 95 and 99\% confidence contours for the parameters \( f_{\text{NL}} \) and \( \tau_{\text{NL}} \) when the fiducial model assumes \( f_{\text{NL}} = \pm 10 \) and \( \tau_{\text{NL}} = 2 \times 10^4 \). The 1-\( \sigma \) errors are \( \sigma_{f_{\text{NL}}} \simeq 23 \) and \( \sigma_{\tau_{\text{NL}}} \simeq 2.0 \times 10^5 \). We have tried different combinations of halo populations and found that the errors do not change significantly. At this point, we would conclude that halo bias alone cannot yield interesting constraints on \( \tau_{\text{NL}} \) and \( f_{\text{NL}} \). The situation changes dramatically when the surveyed halos are divided into \( N \gg 1 \) populations of increasing mass, with equal number density. In Fig. 16, symbols represent the halo model prediction for the 1-\( \sigma \) uncertainties \( \sigma_{f_{\text{NL}}} \) and \( \sigma_{\tau_{\text{NL}}} \) in the limit of infinitely many halo bins. The shot-noise matrix now takes the form Eq. (238). Red triangles indicate \( \sigma_{f_{\text{NL}}} \) in a one-parameter model with \( f_{\text{NL}} = 0 \) (left panel) and \( f_{\text{NL}} = 10 \) (right panel). Filled and empty squares represent \( \sigma_{f_{\text{NL}}} \) and \( \sigma_{\tau_{\text{NL}}} \) in a two-parameters model with \( (f_{\text{NL}}, \tau_{\text{NL}}) = (0, 0) \) (left panel) and \( (f_{\text{NL}}, \tau_{\text{NL}}) = (10, 2 \times 10^4) \) (right panel). Results are shown as a function of the mass of the smallest halos resolved in the survey. Compared to the previous configuration, significant gains are already achieved for \( M_{\text{min}} \approx 10^{13}M_\odot/h \). While the constraint on \( f_{\text{NL}} \) is somewhat degraded if one allows for a non-zero \( \tau_{\text{NL}} \), the 1-\( \sigma \) uncertainty on \( \tau_{\text{NL}} \) is of the order of \( (10^3 - 10^4) \), an order of magnitude better than in the case of two galaxy populations. Table 4 gives the 1-\( \sigma \) errors for \( M_{\text{min}} = 10^{13} \) and \( 10^{11}M_\odot/h \).

Table 4: 1-\( \sigma \) errors obtained with \( N \gg 1 \) halo mass bins with \( M > M_{\text{min}} \). Top and bottom rows show results for \( M_{\text{min}} = 10^{13} \) and \( 10^{11}M_\odot/h \), respectively.

<table>
<thead>
<tr>
<th>( f_{\text{NL}} )</th>
<th>( \tau_{\text{NL}} )</th>
<th>( f_{\text{NL}} )</th>
<th>( \tau_{\text{NL}} )</th>
<th>( f_{\text{NL}} )</th>
<th>( \tau_{\text{NL}} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( 0 )</td>
<td>no</td>
<td>( 0 )</td>
<td>no</td>
<td>( 0 )</td>
<td>( 2 \times 10^4 )</td>
</tr>
<tr>
<td>( f_{\text{NL}} )</td>
<td>( \sigma_{f_{\text{NL}}} )</td>
<td>( \sigma_{f_{\text{NL}}} )</td>
<td>( \sigma_{f_{\text{NL}}} )</td>
<td>( \sigma_{f_{\text{NL}}} )</td>
<td>( \sigma_{f_{\text{NL}}} )</td>
</tr>
<tr>
<td>( 0 ) &amp; ( 0 )</td>
<td>2.9</td>
<td>3.4</td>
<td>3.6</td>
<td>4.2</td>
<td>6.8</td>
</tr>
<tr>
<td>( 0 ) &amp; ( 2 \times 10^4 )</td>
<td>3.4 \times 10^4</td>
<td>3.6 \times 10^4</td>
<td>4.2 \times 10^4</td>
<td>6.8 \times 10^4</td>
<td>4.2 \times 10^4</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>( \sigma_{\tau_{\text{NL}}} )</th>
<th>( \sigma_{\tau_{\text{NL}}} )</th>
<th>( \sigma_{\tau_{\text{NL}}} )</th>
<th>( \sigma_{\tau_{\text{NL}}} )</th>
<th>( \sigma_{\tau_{\text{NL}}} )</th>
<th>( \sigma_{\tau_{\text{NL}}} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( f_{\text{NL}} )</td>
<td>( \tau_{\text{NL}} )</td>
<td>( f_{\text{NL}} )</td>
<td>( \tau_{\text{NL}} )</td>
<td>( f_{\text{NL}} )</td>
<td>( \tau_{\text{NL}} )</td>
</tr>
<tr>
<td>( 0 ) &amp; ( 0 )</td>
<td>0.9</td>
<td>0.9</td>
<td>0.9</td>
<td>0.9</td>
<td>0.9</td>
</tr>
<tr>
<td>( 0 ) &amp; ( 2 \times 10^4 )</td>
<td>1.9 \times 10^4</td>
<td>0.9 \times 10^4</td>
<td>1.9 \times 10^4</td>
<td>0.9 \times 10^4</td>
<td>1.9 \times 10^4</td>
</tr>
</tbody>
</table>
How well can we test the SY inequality with halo bias? Fig. 15 displays, as a function of $f_{\text{NL}}$, the minimum value of $\tau_{\text{NL}}$ for which the difference $(\tau_{\text{NL}} - (36/25)f_{\text{NL}}^2)$ is greater than its 1-$\sigma$, 2- and 3-$\sigma$ error which, for Gaussian-distributed $f_{\text{NL}}$ and $\tau_{\text{NL}}$, reads

$$
\sigma_{\tau_{\text{NL}}}^2 - \frac{36}{25}\sigma_{f_{\text{NL}}}^2 = \sigma_{\tau_{\text{NL}}}^2 + 2 \left(\frac{36}{25}\right)^2 \sigma_{f_{\text{NL}},\tau_{\text{NL}}}^2 + 4 \left(\frac{36}{25}\right)^2 \sigma_{f_{\text{NL}}^2,\tau_{\text{NL}}}^2 - 4 \left(\frac{36}{25}\right) \sigma_{f_{\text{NL}}^2,\tau_{\text{NL}}}^2 f_{\text{NL}},
$$

where $\sigma_{f_{\text{NL}}}^2$, $\sigma_{f_{\text{NL}}^2,\tau_{\text{NL}}}^2$ and $\sigma_{f_{\text{NL}},\tau_{\text{NL}}}^2$ are the entries of the inverted Fisher matrix and $f_{\text{NL}}$, $\tau_{\text{NL}}$ are the values of the fiducial model assumed. The various curves indicate the halo model prediction for $N \gg 1$ halo populations with a minimum resolved mass $M_{\text{min}} = 10^{13} M_\odot / h$. For instance, if a non-vanishing value of $f_{\text{NL}} = 10$ is measured in the future, then the contribution induced by the collapsed limit of the trispectrum must be detected with an amplitude of at least $\tau_{\text{NL}} \sim O(1) \times 10^5$ in order to have a 3-$\sigma$ detection of the SY inequality with the non-Gaussian halo bias. Of course, these values are only indicative since the analysis is performed with the restrictive assumption of Gaussian errors. Finally, we can also assess how well halo bias can probe the violation of the SY inequality. As stated above, the observation of a strong violation would have profound implications for inflationary models as it implies either that multi-field inflation, independently of the details of the model, cannot be responsible for generating the observed fluctuations, or that some new non-trivial (ghost-like) degrees of freedom play a role during inflation. Measuring a violation essentially consists in a simultaneous detection of a non-zero value of $f_{\text{NL}}$ and a (non-zero) small enough value of $\tau_{\text{NL}}$. Here, we have simply estimated the smallest $|f_{\text{NL}}|$ such that $(36/25)f_{\text{NL}}^2$ is larger than the 3-$\sigma$ error on $\tau_{\text{NL}}$. Having found that, for the current observationally allowed range of $f_{\text{NL}}$, the error of $\tau_{\text{NL}}$ does not significantly change if we set in all runs $\tau_{\text{NL}} = 0$, we have thus computed $\sigma_{\tau_{\text{NL}}}$ assuming a vanishing value of $\tau_{\text{NL}}$. A comparison of $3\sigma_{\tau_{\text{NL}}}$ with $(36/25)f_{\text{NL}}^2$ shows that, for a minimum halo mass $M_{\text{min}} = 10^{13} M_\odot / h$, the SY inequality cannot be tested with the non-Gaussian halo bias solely for realistic values of $f_{\text{NL}}$. Even if halos are resolved down to $10^{10} M_\odot / h$ is $3\sigma_{\tau_{\text{NL}}} < (36/25)f_{\text{NL}}^2$ satisfied only for $|f_{\text{NL}}|$ larger than $\sim 80$. Summarising, a large NG in the squeezed limit implies that the cosmological perturbations are generated by some light scalar field other than the inflaton. The SY inequality (201) inevitably imposes that a large trispectrum in the collapsed limit is also present. However, the contribution of $\tau_{\text{NL}}$ to the non-Gaussian halo bias is suppressed by $10^{-4}(\tau_{\text{NL}}/f_{\text{NL}})$ and strongly degenerate with that induced by $f_{\text{NL}}$. Notwithstanding this, we have shown that multi-tracer methods can exploit the distinct mass-dependence of the $f_{\text{NL}}$- and $\tau_{\text{NL}}$-induced bias corrections to reduce the 1-$\sigma$ uncertainty down to $\sigma_{\tau_{\text{NL}}} \lesssim 10^4$ (and simultaneously achieve $\sigma_{f_{\text{NL}}} \sim 1 - 5$) for a survey covering half of the sky up to $z \approx 1$. The exact values depend on the mass $M_{\text{min}}$ of the least massive halos observed. Our results on the capability of testing the SY inequality through the NG scale-dependent bias are summarised in Fig. 15. The latter shows that testing the
6.2 Testing the Running of Primordial Non-Gaussianity

The goal of this section is to provide some useful forecasts on the spectral indices $n_{f_{\text{NL}}}$ and $n_{\tau_{\text{NL}}}$ from the possible physical imprints that NG can leave on the halo bias (see §5.3 for an introduction). The halo bias power spectrum with Gaussian initial conditions can be simply expressed at lowest order in terms of a linear (Eulerian) bias parameter

$$P_h(k) = \left( b_E^2 \right)^2 P_m(k),$$

(245)

SY inequality at the level of 3-$\sigma$ would require detecting $\tau_{\text{NL}}$ at the level of $10^5$ for the minimum resolved mass $M_{\text{min}} = 10^{13} M_\odot / h$. Conversely, testing the violation of the SY inequality requires both a much smaller resolved mass, $M_{\text{min}} = 10^{10} M_\odot / h$ and a large bispectrum, $|f_{\text{NL}}| \gtrsim 80$. As mentioned above, all these results are valid provided that $g_{\text{NL}} = 0$ and that the nonlinear parameters $f_{\text{NL}}$ and $\tau_{\text{NL}}$ estimated from the data are Gaussian-distributed. Relaxing these assumptions will be the subject of future work.

Figure 15: Testing the validity of the SY inequality with measurements of the non-Gaussian bias. The various curves are halo model predictions for $N \gg 1$ halo mass bins with $M_{\text{min}} = 10^{13} M_\odot / h$. At fixed value of $f_{\text{NL}}$, they indicate the minimum $\tau_{\text{NL}}$ required in order to have a measurement of the SY inequality at the 1-, 2- and 3-$\sigma$ confidence level.
where $P_m(k)$ is the dark matter power spectrum. In Sec. §5.3 we have calculated the first order correction in the case of constant valued non-linearity parameters $f_{NL}$ and $\tau_{NL}$, Eq. (226). However, this does not hold for scale-dependent primordial non-Gaussianity. In this case, we use expressions (202) and (204) for the bispectrum and trispectrum to evaluate the derivative of $F^s(N)$ with respect to $\sigma_s$.

For generic primordial 3- and 4-point functions, the non-Gaussian halo power spectrum reads

$$P_h(k) = \left[ \left( b_1^f \right)^2 + 4 b_1^f b_1 \delta_{sc} \frac{F(n_{fNL}, M)}{M_R(k)} \right] P_m(k) + \frac{25}{27} b_1^f \left[ b_2 \delta_{sc} \sigma_s^2 + b_1 \left( 1 + \frac{d\ln T_1}{d\ln \sigma_s} \right) \right] \frac{\mathcal{B}(n_{\tauNL}, M)}{M_R^2(k)} \left[ 1 + \frac{d\ln T_1}{d\ln \sigma_s} \right] + \frac{25}{9} b_1^f \frac{\mathcal{B}_2(n_{\tauNL}, M)}{M_R^2(k)} P_m(k),$$

(246)
where, on large scales, the last term in the square brackets can generate stochasticity between the halo and mass density fields if $\tau_{\text{NL}}$ is different from $(6f_{\text{NL}}/5)^2$ \cite{233, 234, 222, 235}. We have defined the quantities

$$\mathcal{F}(n_{\text{fil}}, M) = \frac{1}{\sigma_F^2} \int \frac{dq}{2\pi^2} q^2 \mathcal{M}^2_R(q) P(q) f_{\text{NL}}(q),$$  \hspace{1cm} (247)

$$T_1(n_{\text{fil}}, M) = \frac{6}{\sigma_F^2} \int \frac{d^3q}{(2\pi)^6} \mathcal{M}_R(q_1) \mathcal{M}_R(q_2) \mathcal{M}_R(q_3) P(q_1) P(q_2) \tau_{\text{NL}}(q_1, q_2),$$  \hspace{1cm} (248)

$$T_2(n_{\text{fil}}, M) = \frac{1}{\sigma_F^2} \int \frac{dq_1 dq_2}{(2\pi^2)^2} q_1^2 q_2^2 \mathcal{M}^2_R(q_1) \mathcal{M}^2_R(q_2) P(q_1) P(q_2) \tau_{\text{NL}}(q_1, q_2).$$  \hspace{1cm} (249)

We have used the definitions (206) and (207) to obtain these expressions. We have also emphasized the dependence on the parameters $n_{\text{fil}}$ and $n_{\text{fil}}$, as well as the halo mass $M$, which, for the top-hat filter, is related to the smoothing radius $R$ through $R = (3M/4\pi \bar{\rho}_m)^{1/3}$. The values of $f_{\text{NL}}^*$ and $\tau_{\text{NL}}^*$ at the pivot wavenumber $k_* = 0.045$ Mpc$^{-1}$ are assumed to be known. In the particular case of scale-independent $f_{\text{NL}}$ and $\tau_{\text{NL}}$, i.e. $n_{\text{fil}} = n_{\text{fil}} = 0$, we recover the expressions given in Refs. \cite{222} and \cite{235}.

In order to assess the ability of forthcoming experiments to probe the scale dependence of the non-linearity parameters $f_{\text{NL}}$ and $\tau_{\text{NL}}$ through a measurement of the large scale bias, we use the Fisher information content on $f_{\text{NL}}$ and $\tau_{\text{NL}}$ (see e.g. \cite{190, 193, 194, 197, 196, 198} for application to the scale-dependence of $f_{\text{NL}}$) in the two-point statistics of halos and dark matter in Fourier space.

We adopt a similar approach to Sec. \S6.1, but in this case using a simple population of tracers (see below) and a constant Poisson noise, which give a covariance matrix of the halo samples of the form,

$$C_\delta(k, M, z) = b^2(k, M, z) P_m(k) + \frac{1}{n},$$  \hspace{1cm} (250)

where $\bar{n}$ is the mean number density of the survey. In order to constrain $n_{\text{fil}}$ and $n_{\text{fil}}$, we assume that we have already measured $f_{\text{NL}}^*$ and $\tau_{\text{NL}}^*$ with infinite precision, that is, without uncertainty. Moreover, since we are interested in investigating the possibility of a detection of the spectral indices, we take $n_{\text{fil}} = n_{\text{fil}} = 0$ throughout as fiducial values. The Fisher matrix is defined as follows

$$F_{ij} = V_{\text{surv}} f_{\text{sky}} \frac{dk}{2\pi^2} \frac{1}{2C_\delta^2} \frac{\partial C_\delta}{\partial \theta_i} \frac{\partial C_\delta}{\partial \theta_j},$$  \hspace{1cm} (251)

where $\theta_i$ are the parameters whose error we wish to forecast, $V_{\text{surv}}$ is the surveyed volume and $f_{\text{sky}}$ is the fraction of the sky observed. The integral over the momenta runs from $k_{\text{min}} = 2\pi/(V_{\text{surv}})^{1/3}$ to $k_{\text{max}} = 0.03$ Mpc$^{-1}$/h, above which the non-Gaussian bias becomes smaller than contributions from second-order bias and nonlinear gravitational evolution. For illustration, we adopt the specifications of a wide-angle, high-redshift
Table 5: 1-σ errors for the population considered in the two different sets of $f_{\text{NL}}^\ast$ and $\tau_{\text{NL}}^\ast$ in Fig. 17.

<table>
<thead>
<tr>
<th>$f_{\text{NL}}^\ast$</th>
<th>$\tau_{\text{NL}}^\ast$</th>
<th>$\sigma_{n_{\text{f}_{\text{NL}}}}$</th>
<th>$\sigma_{n_{\text{f}_{\text{NL}}}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>$5 \times 10^3$</td>
<td>0.30</td>
<td>0.23</td>
</tr>
<tr>
<td>50</td>
<td>$5 \times 10^4$</td>
<td>0.15</td>
<td>0.08</td>
</tr>
</tbody>
</table>

survey such as BigBOSS or EUCLID: $V_{\text{surv}} f_{\text{sky}} = 50 \, \text{Gpc}^3/\hbar$ at median redshift $z = 0.7$. Furthermore, we ignore redshift evolution and assume that all the surveyed volume is at the median redshift.

We compute the uncertainties on $n_{f_{\text{NL}}}$ and $n_{\tau_{\text{NL}}}$ from a single population of tracers consisting of all halos of mass larger than $10^{13} \, M/\hbar$. Computing the Lagrangian bias factors from a Sheth-Tormen mass function \cite{81} leads a linear and quadratic Lagrangian bias $b_1 = 0.7$ and $b_2 = -0.4$. We take the number density to be $\bar{n} = 10^{-4} \, \text{Mpc}^3/\hbar$.

Fig. 17 shows the resulting 68, 95 and 99\% confidence contours for the parameters $n_{f_{\text{NL}}}$ and $n_{\tau_{\text{NL}}}$ when we assume two different combinations of $f_{\text{NL}}^\ast$ and $\tau_{\text{NL}}^\ast$. The 1-σ errors are displayed in Table 5. In the specific case in which only one degree of freedom is responsible for the perturbations, we can use the relation $\tau_{\text{NL}}(k_i, k_j) = \frac{36}{25} f_{\text{NL}}^\ast(k_i) f_{\text{NL}}^\ast(k_j)$, which leaves us with only one parameter, $n_{f_{\text{NL}}}$, describing the scale dependence of the
primordial NG. The 1-$\sigma$ error for $n_{f_{\text{NL}}}$ as a function of $f_{\text{NL}}^*$ is shown in Fig. 18. This result can be compared with those of previous work. For a fiducial value of $f_{\text{NL}}^* = 50$ in particular, we find an error of $\Delta n_{f_{\text{NL}}} \sim 0.2$ in the case of multi-field models, and $\Delta n_{f_{\text{NL}}} \sim 0.1$ in the case of single-field models. For single-field models, this is a factor of $O(3)$ lower than the forecast error found in Ref. [190] for a survey like EUCLID. We attribute this difference to the fact that we have considered the higher-order term $O(f_{\text{NL}}^2)$ in the halo bias and to the parametrization $f_{\text{NL}}(k) = f_{\text{NL}}(k_*) (K^{1/3}/k_*)^{n_{f_{\text{NL}}}}$ considered in Ref.[190] for the running of $f_{\text{NL}}$. In this regards, note that $K \equiv k_1 k_2 k_3$ gives a contribution to the scaling of the external momentum, leading to a suppression (for a positive $n_{f_{\text{NL}}}$) or enhancement (for a negative $n_{f_{\text{NL}}}$) of the signal with respect to our parametrization in Eq. (206). We have checked that, if we use the parametrization and restrict ourselves to the $O(f_{\text{NL}})$ contribution to the halo bias, we are able to reproduce their results. As noted in the introduction, the parametrization used here seems to be motivated by various theoretical predictions (see for example [187, 198]).
6.2.1 Conclusion

Even a tiny level of non-Gaussianity in the cosmological perturbations can tell us a lot about the dynamics of the inflationary Universe. In this section, we have focused on local non-Gaussianity, which is a generic prediction of multifield inflationary models where cosmological perturbations are sourced by light scalar fields other than the inflaton. We have considered the possibility that the non-linear parameter $f_{NL}$ is scale-dependent and, extending the previous literature, we have also assumed that $\tau_{NL}$ may be scale-dependent. This is an unavoidable consequence when only a single field other than the inflaton generates the perturbation as the spectral indices $n_{f_{NL}}$ and $n_{\tau_{NL}}$ are equal. We have assessed the ability of a large-scale galaxy survey to constrain the scale dependence of $f_{NL}$ and $\tau_{NL}$ imprinted in the non-Gaussian halo bias. Assuming the detection of a non-vanishing $f_{NL}$ and $\tau_{NL}$, we find for a large-scale survey like EUCLID that the spectral indices could be measured with an accuracy of $O(0.3)$ for $f_{NL} = 20$ and $\tau_{NL} = 5000$. This limit could be improved by suitably combining the information from several tracers (e.g. [227]).
FACING THE CHALLENGE OF PRECISION MEASUREMENTS OF
THE AMPLITUDE OF THE NON-GAUSSIAN BIAS

If inflation generated a physical coupling between short and long-wavelength perturbations of the gravitational potential, this would induce a characteristic scale dependence on the halo bias that cannot be mimicked by astrophysical effects, since the latter do not generate a mode coupling in the gravitational potential. While models of inflation in which more than one field is involved in the generation of perturbations allow for such a coupling, single-field models of inflation predict that this effect is absent \([176, 177, 236, 237]\). However, they could generate a potentially large PNG of the equilateral type for instance \([174]\) which would not show up in the \(k\)-dependence of the large scale galaxy power spectrum, yet leave an imprint in the galaxy bispectrum. Being able to rule out one of these scenarios would be a great step forward in the understanding of the early Universe.

As exciting as the prospect of detecting a sizable signature of PNG may be, as of today the limits are well consistent with pure Gaussian primordial perturbations, the strongest being the one coming from the Planck satellite \([56]\), \(f_{\text{NL}} = 0.8 \pm 5.0\) at 68\% CL, for the local-type. Even though this value has to be confirmed by other experiments, observing \(f_{\text{NL}} \gtrsim 10\) seems now very unlikely. Current forecasts on \(f_{\text{NL}}\) measures from the scale dependent bias look, however, very promising: for instance, the combination of the galaxy power spectrum and bispectrum leads to a forecasted 68\% CL error \(\sigma(f_{\text{NL}}) = 0.2\) (for the local type) in the case of SPHEREx \([205]\). Previsions from the EUCLID survey \([42]\) are instead, for the galaxy power spectrum only, \(\sigma(f_{\text{NL}}) \approx 4\) and the SKA survey \([202]\) has similar figures.

We have stressed in several points of this thesis that modelling the halo biasing is complicated: this relation is non-linear and stochastic and analytic approaches as well as numerical ones need to rely on assumptions that often are trustful in simple limits and within a certain degree of accuracy. If our goal on PNG is to reach the sensitivity to observe \(f_{\text{NL}} \simeq O(1)\) and this measure is based on the observation of the two- and three-point correlation function of biased tracers, we need to handle the biasing solidly. This task is helped by the fact that the most distinctive signature of PNG on the galaxy power spectrum appears at large scales, where non-linear effects and the scale dependence of
bias are very weak. On the other hand, the galaxy bispectrum can be affected by PNG at any scales and shape (see [238] for a review).

In this chapter, we want to address some of the problems that we need to face for a precision measurement of the effect of PNG on the large scale structure of the Universe.

**Plan of the chapter** This chapter is organized as follows. In Sec. §7.1 we study the effect of local PNG in the power spectrum of halos in the framework of the ESP model. In Sec. §7.2, we study the same effect in numerical N-body simulations, by measuring the cross halo-matter power spectrum, while in Sec. §7.3 we deal with the bispectrum in the squeezed configuration, predicting its behavior up to one loop in the presence of PNG of the local type and comparing the calculation with N-body simulations measurements in the initial conditions.

### 7.1 Non Gaussian Bias in the Framework of the ESP Model

As we already discussed in Sec. §2.1, both analytic models of halo collapse and numerical simulations support the fact that, at a given halo mass $M$, the linear threshold for halo collapse is not the deterministic constant $\delta_{sc}$ predicted by spherical collapse [88, 81, 239, 86, 87]. Owing to the tidal shear and other nonlinear effects, the average linear threshold for collapse is a monotonically increasing function of decreasing halo mass. Furthermore, this collapse threshold fluctuates from halo to halo because it is strongly sensitive to the local density and shear configuration. To the best of our knowledge however, the effect of a moving barrier on the amplitude of non-Gaussian bias has thus far been discussed only in [240] and [216]. [240] argued that the formula of [49] remains valid if one substitutes $\delta_{sc}b_1 \rightarrow \delta_{ec}b_1$, where $\delta_{ec}$ is the threshold for ellipsoidal collapse. [216] investigated the effect of an ellipsoidal barrier on the non-Gaussian bias within the path integral approach to excursion set (see [76]). They found that the non-Gaussian bias amplitude is generally different from $\delta_{ec}b_1$. However, both papers did not consider the stochasticity of the collapse barrier.

In this section, we will explore the effect of a realistic, stochastic moving barrier on the non-Gaussian bias of dark matter halos within the peak theory framework. Clearly, the ESP mass function, Eq. (61), is not universal since $f_{ESP}$ is a complex function of $\nu_c$ and the spectral moments $\sigma_i$. In addition, random walks associated with excursion set peaks are non-Markovian owing to the shape of the Top-hat and Gaussian filters.

As shown in [83], the non-Gaussian bias of excursion set peaks induced by a primordial non-Gaussianity of the form $f_{NL}\phi^2$, where the nonlinear parameter $f_{NL}$ is scale-independent, has an amplitude given by

$$b_{NG} = \sigma_0^2 b_{200} + 2\sigma_1^2 b_{110} + \sigma_2^2 b_{020} + 2\sigma_1^2 \chi_{10} + 2\sigma_2^2 \chi_{01} + \Delta_0^2 b_{002} - (\sigma_0^2)'b_{101} - (\sigma_1^2)'b_{011}.$$ (252)
7.1 Non-Gaussian Bias in the Framework of the ESP Model

Figure 19: Dimensionless second-order bias factors in the excursion set peak approach for the constant, deterministic barrier $B(\sigma) = \delta_{sc}$. Dotted curves represent negative values. $b_{NG}$ shown as the thick solid curve is the sum of all these contributions (Eq. (252)) and, according to the peak approach, is equal to the amplitude of the non-Gaussian bias. For clarity, we have not shown the bias factors $b_{101}$, $b_{011}$ and $b_{002}$ which arise from the first-crossing constraint.

Here, a prime denote a derivative w.r.t. $R_s$. $b_{ijk}$ and $\chi_{ij}$ are the ESP peak bias factors that can be derived from the ESP peak “localised” number density using a peak-background split argument (see for details and notation Sec. 2.2). This is particularly interesting because the right-hand side was obtained from the “effective” bias expansion introduced in Sec. 2.2 (see [83] for the original derivation). In Fig. 19, some of the second order bias factors together with the resulting behaviour of $b_{NG}$ are shown for the constant barrier $B(\sigma_0) = \delta_{sc}$ as a function of the peak significance $\nu_c$.

For this constant, deterministic barrier, [83] demonstrated that the amplitude $b_{NG}$ of the non-Gaussian bias satisfies

$$b_{NG} = \delta_{sc} b_1 = b^{\text{obs}}_{NG},$$

(253)
where we have defined the peak-background split prediction, which we derived in Sec. §5.3, as

$$b_{NG}^{\text{pbs}} = \frac{\partial \ln \tilde{n}_b}{\partial \ln \sigma}.$$  

(254)

Notice that here $b_N \equiv b_{N00}$ is the $k$-independent piece of the $N$th-order Lagrangian, Gaussian bias (the usual Lagrangian bias parameters in the standard local bias model, see Eq. (22)). Under the approximation of a constant barrier, peak theory thus predicts that the amplitude of the non-Gaussian bias is equally given by the sum of quadratic bias factors Eq.(252), the original result $\delta_{ec} b_1$ of [49] or the peak-background split expectation $b_{NG}^{\text{pbs}}$ obtained by [51]. We tested this equivalence numerically and found that it indeed holds. The thin, indistinguishable curves in Fig.20 show the various predictions. At this point, it is worth noticing that, although the excursion set peak mass function is not universal (it depends distinctly on $\delta_{ec}$ and the spectral moments $\sigma_i$), the logarithmic derivative of $\tilde{n}_b$ w.r.t. $\sigma$ is nonetheless equal to $\delta_{ec} b_1$. This follows from the fact that the $\sigma_i \delta$ conspire to appear only in ratios such as $\gamma_1 = \sigma^2_1 / (\sigma_0 \sigma_2)$ or in $\nu_c = \delta_{ec} / \sigma_0$.

Thus far however, we have followed [83] and assumed a constant barrier $B \equiv \delta_{sc}$. How does the relation Eq.(253) change when we take into account the scatter and mass-dependence of the collapse barrier through the square-root stochastic barrier Eq.(59)? To answer this question, we have simply computed $b_{NG}$ and $b_1 = b_{100}$ from the bias factors derived from the ESP multiplicity function Eq.(61) as in Sec. §2.2. We have also evaluated $b_{NG}^{\text{pbs}}$ numerically from the predicted halo mass function (we have again explicitly taken the numerical derivative of $\tilde{n}_b$ w.r.t. $\sigma_8$). The results are shown in Fig.20 as the thick solid curves. They can be summarised as follows:

$$b_{NG} \neq \delta_{ec} b_1 = b_{NG}^{\text{pbs}}.$$  

(255)

While $b_{NG}$ agrees with the two other quantities at the high mass end, where all the predictions converge towards the high-peak result of [50], it becomes increasingly larger as the halo mass decreases. For the lognormal distribution of $\beta$ adopted here, deviations are quite substantial. Namely, for $M = 10^{14}$ and $10^{13} h^{-1} M_\odot$, the predicted non-Gaussian bias amplitude $b_{NG}$ is $\sim 10\%$ and $\sim 40\%$ larger than the peak-background split amplitude $b_{NG}^{\text{pbs}}$. Upon turning the scatter in $\beta$ on and off, we have found that the latter is driving the difference between $b_{NG}$ and $b_{NG}^{\text{pbs}}$ for $\nu_c \gtrsim 2$. At higher peak heights, the discrepancy originates mainly from the fact that the barrier is not flat.

[240] advocated the replacement $\delta_{sc} b_1 \rightarrow \delta_{sc} b_1$ to account for the mass-dependence of the linear collapse threshold. We have found that substituting $\delta_{sc}$ either by the mean barrier $\delta_{sc} + \alpha_0 (\beta)$ or by the square-root of $((\delta_{sc} + \sigma_0 \beta)^2)$ does improve the agreement with $b_{NG}$, yet the match is far from perfect, especially around $\nu_c \sim 1$. Furthermore, we do not select our peaks according to their formation history. Hence, this has nothing to do with the assembly bias effect pointed out in [51], for which the extended Press-Schechter formalism of [48] furnishes a good description [241]. Finally, [216] pointed out that $b_{NG}$ generally differs from $\delta_{ec} b_1$ (but note that they did not discuss the validity of $b_{NG}^{\text{pbs}}$). However, they found a much larger effect than we did (see their Fig.3).
7.1 Non Gaussian Bias in the Framework of the ESP Model

7.1.1 A closer look at the peak prediction

For simplicity, let us momentarily ignore the variable \( \mu \) as it is not essential for understanding why a square-root stochastic barrier induces a difference between \( b_{\text{NG}} \) and the peak-background split prediction (it is enough to retain the correlation between \( v \) and \( u \)). In this case, the non-Gaussian bias amplitude takes the form

\[
b_{\text{NG}} = \sigma_0^2 b_{20} + 2\sigma_1^2 b_{11} + \sigma_2^2 b_{02} + 2\sigma_1^2 \chi_{10} + 2\sigma_2^2 \chi_{01} \ . \tag{256}
\]

The peak bias factors \( b_{ij} \) (associated with \( v \) and \( u \)) and \( \chi_{kl} \) (associated with the \( \chi^2 \)-distributed variables) can all be computed by generalising the peak-background split argument to variables other than the density [66]. In particular, since the peak height \( v(x) \) is correlated with \( u(x) \) (at the same position \( x \)), we have

\[
\sigma_0^2 \sigma_2^2 b_{ij} = \frac{1}{n_{pk}} \int \!^{d10} y \ n_{pk}(y) \ H_{ij}(v, u) \ P_1(y) \ . \tag{257}
\]

Here, \( n_{pk}(y) \) is the "localised" number density of BBKS peaks (as we momentarily ignore the first-crossing constraint), \( y \) is a vector of 10 variables and \( H_{ij}(v, u) \) are bivariate Hermite polynomials and \( P_1(y) \) is of the form of Eq. (39). When stochasticity in the barrier is taken into account, \( n_{pk}(y) \) contains a multiplicative factor of \( \delta_D(v(x) - v_c - \beta) \).

In Eq. (256), the contribution \( 2\sigma_1^2 \chi_{10} + 2\sigma_2^2 \chi_{01} \) does not depend on the properties of the collapse barrier because the \( \chi^2 \)-distributed variables do not correlate with \( v(x) \) at a given position \( x \). Therefore, we should focus on the piece proportional to \( b_{ij} \).

On writing the bivariate Gaussian as

\[
\mathcal{N}(v, u) = \frac{\exp\left[-\frac{v^2 + u^2 - 2\gamma_{vu} vu}{2(1 - \gamma_{vu}^2)}\right]}{2\pi \sqrt{1 - \gamma_{vu}^2}} = \frac{e^{-Q(v,u)/2}}{2\pi \sqrt{1 - \gamma_{vu}^2}} \ ,
\]

the sum \( \sigma_0^2 b_{20} + 2\sigma_1^2 b_{11} + \sigma_2^2 b_{02} \) simplifies to (without writing down the integrals over \( \beta \) and \( u)\)

\[
\sigma_0^2 b_{20} + 2\sigma_1^2 b_{11} + \sigma_2^2 b_{02} \sim \frac{(v_c + \beta)^2 + u^2 - 2\gamma_{vu}(v_c + \beta)u}{1 - \gamma_{vu}^2} - 2 \mathcal{N}(v_c + \beta, u) \sim \left(2Q(v_c + \beta, u) - 2\right) \mathcal{N}(v_c + \beta, u) . \tag{259}
\]

This should be compared to the full expression of the logarithmic derivative \( \partial \ln \tilde{n}_h / \partial \ln \sigma_8 \). The latter requires evaluating the derivatives of the multiplicity function, which is an integral of the bivariate Gaussian \( \mathcal{N}(v_c + \beta, u) \) over \( \beta \) and \( u \) similar to Eq. (61). Therefore,
the logarithmic derivative of the halo mass function w.r.t. $\sigma_8$ results in a term of the form

$$\frac{\partial}{\partial \sigma_8} N(v, u) = \frac{\partial}{\partial v} N(v, u) \frac{dv}{d\sigma_8} + \frac{\partial}{\partial u} N(v, u) \frac{du}{d\sigma_8}$$

$$= \left[ - \left( \frac{v - \gamma_{vu} u}{1 - \gamma_{vu}} \right) \frac{dv}{d\sigma_8} - \left( \frac{u - \gamma_{vu} v}{1 - \gamma_{vu}} \right) \frac{du}{d\sigma_8} \right] \times N(v, u).$$

Note that $\gamma_{vu}$ does not contribute since it is invariant under a (scale-independent) rescaling of $\sigma_8$. Now, we use the fact that $v \equiv v_c + \beta$, with $\beta$ independent of $\sigma_8$ and $u \propto 1/\sigma_2$. Hence, $dv/d\sigma_8 = -v_c/\sigma_8$ and $du/d\sigma_8 = -u/\sigma_8$. Substituting these derivatives in the previous expression, we arrive at

$$\frac{\partial}{\partial \sigma_8} N(v_c + \beta, u)$$

$$= \frac{1}{\sigma_8} \left[ (v_c + \beta - \gamma_{vu} u) v_c + (u - \gamma_{vu} (v_c + \beta)) u \right] \times N(v_c + \beta, u).$$

We should now compare the square brackets in Eq.(259) with that of Eq.(261). We note that, in Eq.(259), there is an additional factor of $-2$ inside the brackets which disappears when one takes into account the first-crossing constraint. So, the key difference is the fact that, for $\partial N/\partial \ln \sigma_8$, the square brackets reduce to $2Q(v_c + \beta, u)$ as in Eq.(261) only if $\beta \ll v_c$, a condition which is only satisfied in the high peak limit $v_c \gg 1$. This is the reason why, in Fig.20, $b_{NG}$ increasingly differs from $b_{pbs}^{\text{NG}}$ as $v_c$ decreases. We also note that Eqs. (259) and (261) will differ even in the absence of scatter in the moving barrier (i.e. $\langle \beta^2 \rangle = \langle \beta \rangle^2$).

The peak model and peak-background split predictions will agree for a moving barrier only if $dv/d\sigma_8 = -(v_c + \beta)/\sigma_8$ or, equivalently, if $\beta \propto \sigma_0^{-1}$. This implies that the deviation from $\delta_{sc}, \sigma_0 \beta$, does not depend on $\sigma_0$. However, numerical simulations [74, 87] clearly indicate that the scatter in the barrier increases with decreasing halo mass and is approximately proportional to $\sigma_0$ (hence the designation square-root barrier). Therefore, we shall expect $b_{NG} \neq b_{pbs}^{\text{NG}}$ for actual (SO) dark matter halos if excursion set peak theory accurately describes their clustering properties.
7.1.2 The squeezed limit of the galaxy bispectrum

Retaining terms up to the fourth-point function and working within the usual local bias approximation \( \delta_h(x) = b_1 \delta(x) + (1/2) b_2 \delta^2(x) + \ldots \), the halo bispectrum with primordial non-Gaussianity of the local type is given by \([242, 243, 244]\)

\[
B_h(k_1, k_2, k_3) = 2b_1^2 \left[ \left( f_{NL} \frac{M(k_3)}{M(k_1) M(k_2)} + F_2(k_1, k_2) \right) \right. \\
\left. \times P(k_1)P(k_2) \right] + b_2 b_1^2 P(k_1) P(k_2) \\
+ \frac{1}{2} b_2^2 b_1 \int \frac{d^3 \mathbf{q}}{(2\pi)^3} T(\mathbf{q}, k_1 - \mathbf{q}, k_2, k_3) \\
+ (2 \text{ cyc.}) .
\]

(262)

Here, \( M(k) \propto k^2 \) is the transfer function between linear density and potential perturbations and \( T \) is the matter trispectrum. We have also omitted the filtering kernels as they are not essential for the purpose of this discussion. In the squeezed configurations, the two dominant contributions are the first and fourth term in the right-hand side. The first is proportional to \( f_{NL} \) whereas the fourth contains a contribution from the linearly evolved primordial trispectrum proportional to \( f_{NL}^2 \), and a cross-correlation between the primordial bispectrum and the nonlinearly evolved density field proportional to \( f_{NL} \).

For peaks, the analyses of \([66, 83]\) and the correspondence with the Integrated Perturbation Theory (iPT) framework \([103, 214]\) indicate that the fourth term shall be replaced by the more general expression

\[
\frac{1}{2} c_1(k_2) c_1(k_3) \int \frac{d^3 \mathbf{q}}{(2\pi)^3} c_2(\mathbf{q}, k_1 - \mathbf{q}) T(\mathbf{q}, k_1 - \mathbf{q}, k_2, k_3) ,
\]

(263)

where the linear and quadratic Lagrangian peak bias parameters \( c_n(k_1, \ldots, k_n) \) are given by

\[
c_1(k) \equiv (b_{10} + b_{01} k^2) 
\]

(264)

and

\[
c_2(k_1, k_2) \equiv \left\{ b_{20} + b_{11} (k_1^2 + k_2^2) + b_{02} k_1^2 k_2^2 \\
- 2 \chi_{10} (k_1 \cdot k_2) + \chi_{01} \left[ 3 (k_1 \cdot k_2)^2 - k_1^2 k_2^2 \right] \right\} .
\]

(265)
Here again, we have ignored the first-crossing constraint and omitted multiplicative factors of filtering kernels for sake of conciseness (see Sec. §2.2 for the full expressions). Restricting ourselves to the contribution of the primordial trispectrum, terms of the form

\[ f_{\text{NL}}^2 c_1(k_2) c_1(k_3) M(k_2) M(k_3) \left[ P_\phi(k_2) + P_\phi(k_3) \right] \]

\[ \times P_\phi(k_1) \int \frac{d^3 q}{(2\pi)^3} M^2(q) c_2(q, -q) P_\phi(q) . \]  

arise in the squeezed configurations \( k_1 \to 0 \). Since \( M^2(k) P_\phi(k) \equiv P(k) \), where \( P_\phi(k) \) is the power spectrum of the Gaussian part of the primordial curvature perturbation, the integral over \( c_2(q, -q) \) simplifies to (after re-introducing the filtering kernels)

\[ \int \frac{d^3 q}{(2\pi)^3} c_2(q, -q) P(q) = \sigma_0^2 b_{20} + 2 \sigma_1^2 b_{11} + \sigma_2^2 b_{02} \]

\[ + 2 \sigma_1^2 \chi_{10} + 2 \sigma_2^2 \chi_{01} \]

\[ \equiv b_{\text{NG}} . \]

Therefore, this suggests that some of the terms proportional to \( \sigma_0^2 b_{20} \) in a calculation which assumes the standard local bias (e.g. [243, 244]) are, in fact, proportional to \( b_{\text{NG}} \). Since \( b_{\text{NG}} \) is noticeably different than \( \sigma_0^2 b_{20} \) (see Fig.19), this will of course have a large impact on the magnitude of the PNG signal and its dependence on halo mass. However, we stress again that, unless the barrier is flat and deterministic, \( b_{\text{NG}} \) cannot be replaced by the peak-background split expression \( b_{\text{pbs}} \) in the framework of the ESP model with a stochastic barrier. It will also be useful to compare the predictions of the peak approach with e.g. the models of [245, 217], which are based on the multivariate bias scheme of [209]. We leave all this for future work.

7.1.3 Final remarks

In this section, we have studied the prediction of the ESP model with a stochastic moving barrier on the amplitude of the scale dependent bias, which arises in the presence of PNG. In particular, we have made a detailed comparison of this prediction with the one obtained from a pure peak-background split argument, Eq. (254), and the result obtained with universal mass functions, \( \delta_{sc} b_1 \). These predictions, when calculated within the ESP framework, are in disagreement with each other and the fractional departure is expected to increase with decreasing halo mass in the proportions shown in Fig.20. In light of the ESP model assumptions, this disagreement is meaningful only for dark matter halos closely related to an initial density peak, which is approximately the case for \( M \gtrsim M_* \). We also stress that our prediction is strictly valid for SO halos only since the excursion set peak model used in the present analysis was calibrated with SO halos identified with a fixed overdensity \( \Delta = 200 \) relative to the background.
Our analysis has also revealed that, in the ESP framework, $\delta_{sc}b_1 = b_{NG}$ even though the ESP mass function is not universal. As we have shown, this follows from the fact that, in peak theory, $\bar{n}_h$ depends only on ratios of the spectral moments $\sigma_i$ in addition to $\nu_c = \delta_{sc}/\sigma_0$. Note, however, that this equality may hold only for square-root barriers. Furthermore, the functional dependence $\bar{n}_h(\nu_c, \gamma_{\nu\mu}, \ldots)$ may be very peculiar to the peak approach. Hence, it is unclear whether the clustering of actual dark matter halos satisfies $\delta_{sc}b_1 = b_{NG}$.

These findings raise some concern on how to correctly model the effect of primordial non Gaussianity in the halo bias, a prediction which will be of fundamental importance for constraining $f_{NL} \approx O(1)$ with future galaxy surveys. In particular, although it has been shown that the stochastic moving barrier of the form of Eq. (59) provides a good description of the clustering of dark matter halos, Figure 2, its implementation causes the ESP prediction of the non Gaussian bias amplitude to be different from what we would expect from a peak-background split calculation, whose derivation is general. We
have argued that the reason for this disagreement is the way we model the stochasticity: the response of the halo mass function to the change of the local matter amplitude can not capture any pure stochastic effect, as it is the case of our variable $\beta$. Hence, we are raising a question on how to properly model the collapse barrier including stochastic effects induced by the triaxiality of collapse. An improvement would be possible by explicitly including the effect of the tidal shear in the collapse barrier [246]. Moreover, in [246] it also been argued that, by correctly implementing the collapse time in the identification of halos at late time, the scatter around a mean barrier is significantly reduced.

Implementing such effects is beyond the scope of this analysis and we leave it for future work.

7.2 Non Gaussian Bias from N-Body Simulations

In Sec. 5.3 of the previous chapter, using a peak-background split ansatz, we have shown that a local-type non Gaussianity in the primordial gravitational potential induces a local modulation of the amplitude of matter fluctuations proportional, at first order, to the non-linearity parameter $f_{NL}$, thereby sourcing a scale dependent bias in the distribution of DM halos of the form [49, 50, 51]

$$\frac{\langle \delta_h \delta_m \rangle}{\langle \delta_m \delta_m \rangle} \approx b_1^C + 2f_{NL} \frac{d \phi_L}{d \delta_L} \frac{\partial \ln \bar{n}_h}{\partial \ln \sigma_8} + ... \quad (268)$$

We have also argued that for universal mass functions and in the spherical collapse approximation, the non Gaussian bias amplitude is proportional to the first order Lagrangian bias,

$$\frac{\partial \ln \bar{n}_h}{\partial \ln \sigma_8} \overset{\text{univ.}}{\longrightarrow} \delta_{sc} b_1^{\text{Lag}}(M). \quad (269)$$

The assumption of universality of the mass function has long been studied and its validity is still under debate [96, 247, 248] (see also [249] for a discussion about universality in non-Gaussian simulations). Moreover, it is not clear yet to what extent even a small deviation from universality may affect the non Gaussian bias amplitude and therefore induce corrections in the relation of Eq. (269) (see [207, 214, 249]). Quantifying these corrections is particularly relevant for galaxy redshift surveys which aim for precision measurements of the parameter $f_{NL}$, for which the goal is to reach the sensitivity of order unity. The forecasts we discussed up to now in this thesis, as well as the ones estimated by the major collaborations aiming at this measure (see introduction to this Chapter §7) are currently based on the limit of Eq. (269).

The goal of this section is therefore to verify the validity of the limit of Eq. (269) by measuring the effect directly in N-body simulations which include only DM particles. We will adopt the following strategy:
1. Run 3 sets of simulations with Gaussian initial conditions and identical cosmologies, but for different values of the matter amplitude \( \sigma_8 \);

2. Run 2 sets of simulations with non-Gaussian initial condition of the local type with positive and negative non-linearity parameter\(^1\) \( |f_{\text{NL}}| = 250 \);

3. Estimate numerically the logarithmic derivative of the halo mass function \( n_h \) w.r.t \( \sigma_8 \) using the 3 sets of simulation of point 1), plus 3 more sets with a smaller simulated volume to check for convergence of the estimation at lower mass.

4. Measure the linear Eulerian bias \( b_1^{\text{Eul}} = 1 + b_1^{\text{Lag}} \) from the Gaussian simulations;

5. Measure the scale dependence of the halo power spectrum at large scales in the presence of primordial non-Gaussianity by estimating the cross halo-matter power spectrum \( \langle \delta_h \delta_m \rangle / \langle \delta_m \delta_m \rangle \) in the non-Gaussian simulations.

Theoretical and numerical arguments suggest that the largest deviations from universality are to be expected in the range of masses around and below \( M_* \), where the collapse process becomes most affected by environmental quantities (e.g. tidal shears) and becomes triaxial. We will therefore try to investigate the signal as close as possible to this ranges.

<table>
<thead>
<tr>
<th>runs</th>
<th>N particles</th>
<th>L box (Gpc/h)</th>
<th>( \sigma_8 )</th>
<th>( f_{\text{NL}} )</th>
<th>( n_s )</th>
<th>( h )</th>
<th>( \Omega_m )</th>
<th>( \Omega_\Lambda )</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>1536^4</td>
<td>2.0</td>
<td>0.83</td>
<td>0</td>
<td>0.967</td>
<td>0.7</td>
<td>0.3</td>
<td>0.7</td>
</tr>
<tr>
<td>6</td>
<td>1536^5</td>
<td>2.0</td>
<td>0.85</td>
<td>0</td>
<td>0.967</td>
<td>0.7</td>
<td>0.3</td>
<td>0.7</td>
</tr>
<tr>
<td>4</td>
<td>1536^5</td>
<td>2.0</td>
<td>0.87</td>
<td>0</td>
<td>0.967</td>
<td>0.7</td>
<td>0.3</td>
<td>0.7</td>
</tr>
<tr>
<td>6</td>
<td>1536^5</td>
<td>2.0</td>
<td>0.85</td>
<td>-250</td>
<td>0.967</td>
<td>0.7</td>
<td>0.3</td>
<td>0.7</td>
</tr>
<tr>
<td>1</td>
<td>1536^5</td>
<td>1.0</td>
<td>0.83</td>
<td>0</td>
<td>0.967</td>
<td>0.7</td>
<td>0.3</td>
<td>0.7</td>
</tr>
<tr>
<td>1</td>
<td>1536^5</td>
<td>1.0</td>
<td>0.85</td>
<td>0</td>
<td>0.967</td>
<td>0.7</td>
<td>0.3</td>
<td>0.7</td>
</tr>
<tr>
<td>1</td>
<td>1536^5</td>
<td>1.0</td>
<td>0.87</td>
<td>0</td>
<td>0.967</td>
<td>0.7</td>
<td>0.3</td>
<td>0.7</td>
</tr>
</tbody>
</table>

7.2.1 The N-body simulations

Since our goal is to thoroughly investigate the scale dependence of halo bias at large scales in the presence of initial non-Gaussian conditions for halo masses as close as possible to \( M_* \), our simulations need on one hand to be run on a big enough volume such that we can push for large scales and on the other hand they need to have high resolution to be able to trust the sensitivity on such low mass ranges.

\(^1\) We are of course aware of the fact that such a value for \( f_{\text{NL}} \) local is excluded by current observations. However, here the focus is on the amplitude of the scale dependent bias proportional to \( f_{\text{NL}} \), rather than the amplitude of \( f_{\text{NL}} \) itself, therefore it is a good strategy to enhance as much as possible this effect in the non Gaussian simulation so as to have a clean signal.
We achieve this goal by running the 8 sets of simulations outlined in Table 6. These simulations were run on the Baobab cluster at the University of Geneva and on the Odin cluster at the Max Planck Institute in Garching. The cosmology is a flat ΛCDM model with Ω_m = 0.3, h = 0.7, n_s = 0.967 and varying σ_8 as shown on Table 6. The transfer function was obtained from the Boltzmann code CLASS ([161]). The initial particle displacements were implemented at z_i = 99 using the public code 2LPTic ([250, 98]) for realizations with Gaussian initial conditions and its modified version ([207]) for non-Gaussian initial conditions of the local type. The simulations were evolved using the public code Gadget2 ([99]). In particular, for the first 5 sets of simulation with volume 8 Gpc^3/h^3, the mass of each particle is M_{part} = 1.8 × 10^{11} M_⊙, so that a typical low mass halo with 50 particles has a mass of M_{min} = 9.2 × 10^{12} M_⊙. For the cosmology with σ_8 = 0.85, we have that v = δ_{sc} / σ(M_*) = 1 for M_* = 5.6 × 10^{12} M_⊙.

We use two different algorithms for finding DM halos: a spherical overdensity (SO) algorithm, called Amiga Halo Finder (AHF) ([251, 100]) and a Friends-of-Friends (FOF) technique developed in the Rockstar code ([252]), for which we use a linking length of b = 0.28. All the results shown refer to redshift z = 0. We leave the investigation of this effect at higher redshifts for future work. Here and henceforth, error bars will represent the standard deviation of the mean calculated from the different realizations

\[ \sigma_{\text{mean}} = \frac{1}{N} \sqrt{\frac{1}{N} \sum_{i=1}^{N} (x_i - \mu)^2} \]  

(270)

where \(x_i\) is the value for the i-th realization, \(N\) is the number of realizations and \(\mu\) is the mean among the realizations.

### Halo mass function

We compute the halo mass function for each set of simulations for the two algorithms by counting halos in 30 logarithmically spaced mass bins. We compare the measurements to fits optimized for SO halos ([253], hereafter Tinker2010) and FOF halos ([59], hereafter S&T99). In Figure 21 we show the results for the Gaussian simulations with \(σ_8 = 0.85\) where we include both box sizes to check convergence at low mass.

In Figure 22, we compare directly the two different halo finders for the case of the 2Gpc/h box. The halo mass function in the presence of primordial non Gaussianity is typically modified with respect to the case of Gaussian initial condition in the high peak limit, with a deviation which depends on the sign and amplitude of \(f_{NL}\) (see [254, 255] and [197] for an analytic prediction). To check this behavior, we plot in Figure 23 the ratio of the non Gaussian mass functions over the Gaussian mass functions.
7.2 NON GAUSSIAN BIAS FROM N-BODY SIMULATIONS

7.2.3 Linear bias

At large enough scales, the halo bias is scale independent, even though dependent on the mass and redshift of the halo population considered, in the case of Gaussian initial conditions. This constant value can be measured by taking the ratio of the halo and matter power spectra in two different ways

\[ b_{hm}^G = \frac{P_{hm}^G}{P_{mm}^G} \]  \hspace{1cm} (271)

\[ b_{hh}^G = \sqrt{\frac{P_{hh}^G}{P_{mm}^G}} \]  \hspace{1cm} (272)

where in \( b_{hh}^G \) one needs to subtract for the shot noise, which in this case we take to be constant and equal to the inverse of the number density \( 1/\bar{n} \), assuming that DM halos are a Poisson sampled.

To measure these power spectra, we extract dark matter and halo fluctuation fields \( \delta_m(k) \) and \( \delta_h(k) \) by interpolating particles (dark matter and halo centers) on a three dimensional grid of size 512.

We decide to split the halo catalogs in three mass bins with equal number of halos, these bins are displayed in Table 7 along with linear bias values that we extract by taking ratios as in Eqs. (271) and (272) and averaging over the values corresponding to the interval ranging from the second to the tenth mode, that is, \( k \in [0.004, 0.03] \), since at higher wavenumbers higher order biases start to be important (such as \( b_2 \)). In Figure 24 (shown as Figure 21).

Figure 21: Halo mass function for the Gaussian simulations with \( \sigma_8 = 0.85 \) for both box sizes for the SO algorithm (left) and FOF(right) and corresponding fits in dotted lines. On the lower panel, we show the relative difference between the fit and the measurement.
Figure 22: Comparison of SO and FoF halo finders for Gaussian simulations with $\sigma_8 = 0.85$ for both box sizes (upper box) and their relative difference (lower box).

we display the ratios for SO halos and FoF halos. We also display the central mass value of the bin, which we compute by weighting the mass with the halo mass function,

$$\bar{M} = \frac{\int dm \, m \, dn/dm}{\int dm \, dn/dm}. \quad (273)$$

<table>
<thead>
<tr>
<th>SO halos 2 Gpc/h</th>
<th>FoF halos 2 Gpc/h</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mass $M$</td>
<td>$b_{mh}$ $c_{mh}$ $b_{hh}$ $c_{hh}$ $b_{1\text{lin}}$</td>
</tr>
<tr>
<td>9.2 – 14</td>
<td>11.1 0.02 1.06 0.05 1.04</td>
</tr>
<tr>
<td>14 – 27</td>
<td>18.9 0.01 1.18 0.01 1.15</td>
</tr>
<tr>
<td>27 – 3000</td>
<td>82.9 0.01 1.63 0.02 1.60</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SO halos 1 Gpc/h</th>
<th>FoF halos 1 Gpc/h</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mass $M$</td>
<td>$b_{mh}$ $b_{hh}$ $b_{1\text{lin}}$</td>
</tr>
<tr>
<td>1.1 – 1.8</td>
<td>1.13 0.14 1.14</td>
</tr>
<tr>
<td>1.8 – 3.9</td>
<td>18.9 1.28 1.27</td>
</tr>
<tr>
<td>3.9 – 1000</td>
<td>80.5 1.70 1.70</td>
</tr>
</tbody>
</table>

Table 7: Measured values for linear bias at redshift $z = 0$, where mass ranges are expressed in units of $10^{12}M_\odot$ and refer to the 2Gpc/h box sets (left) and the 1Gpc/h box sets (right).
7.2.3.1 Scale dependent bias

Local non-Gaussianity gives rise to a scale dependent bias at large scales. At lowest order, there are a number of other effects to consider. Let us introduce the quantity we want to measure, using the notation of ([256]), as

$$\frac{P_{\text{NG}}(k, f_{\text{NL}})}{P_{\text{mm}}(k, 0)} = b_{\text{hm}}^G + \Delta b_x(k, f_{\text{NL}}) + \Delta b_i(f_{\text{NL}}) + \beta_m(k, f_{\text{NL}}) + O(b_G^2, f_{\text{NL}}^2). \quad (274)$$

We are taking the ratio between the halo-matter cross power spectrum in non-Gaussian simulations over the matter auto power spectrum in Gaussian simulations. This ratio is the sum of the linear Eulerian Gaussian bias $b_{\text{hm}}^G$, the scale independent correction that arises from the change in the mean number density of halos in the presence of PNG (see [51, 240]), the change in the power spectrum of matter given by PNG ([257, 258, 259, 249])

$$\beta_m(k, f_{\text{NL}}) = \frac{P_{\text{mm}}(k, f_{\text{NL}}) - P_{\text{mm}}(k, 0)}{P_{\text{mm}}(k, 0)} \quad (275)$$

and finally the scale dependent shift,

$$\Delta b_x(k, f_{\text{NL}}) = 2f_{\text{NL}} \frac{b_X^{\text{NG}}(k)}{M(k)}. \quad (276)$$

where $M(k) \propto k^2$ at large scales, so that $\Delta b_x(k, f_{\text{NL}}) \propto 1/k^2$ at large scales, a scaling which is absent in all the other terms of Eq. (274). The term $O(b_G^2, f_{\text{NL}}^2)$ indicates that the expression is valid at first order both in bias and PNG effects. Notice that, for

Figure 23: Halo mass function (upper boxes) and relative difference (lower boxes) for the Gaussian and non Gaussian simulations with $\sigma_8 = 0.85$ for both box sizes for the SO algorithm (left panel) and FoF(right panel).
these high values we chose for $f_{\text{NL}} = \pm 250$, second order effects scaling as $f_{\text{NL}}^2$ may be important. However, we can exploit the fact that we have run simulations with both negative and positive $f_{\text{NL}}$, and identify terms proportional to $f_{\text{NL}}^2$ by

$$2 \frac{P_{\text{NG}}(k, f_{\text{NL}}^2)}{P_{\min}(k, 0)} = \frac{P_{\text{NG}}(k, f_{\text{NL}}^+)}{P_{\min}(k, 0)} + \frac{P_{\text{NG}}(k, f_{\text{NL}}^-)}{P_{\min}(k, 0)}$$  \hspace{1cm} (277)

and subtract Eq. (277) from Eq. (274).

Since our final goal is to test the relation in Eq. (269), we distinguish two different prescriptions for predicting this scale dependent signal, one which we call “Dalal et al.(DB)” prediction and the second “peak-background split (PBS)” prediction.

$$b_{\text{DB}}^{\text{NG}} = \delta_{\text{sc}} (b_{\text{Eul}}^1(M) - 1)$$  \hspace{1cm} (278)

$$b_{\text{PBS}}^{\text{NG}} = \frac{\partial \ln \bar{n}_h}{\partial \ln \sigma_8}.$$  \hspace{1cm} (279)

where in the first relation we are subtracting one from the measured Eulerian bias to get the Lagrangian quantity, using the first order relation (see e.g. Appendix §11.1)

$$b_{\text{Eul}}^1 = 1 + b_{\text{Lag}}^1.$$  \hspace{1cm} (280)

Measuring $b_{\text{DB}}^{\text{NG}}$ therefore is rather simple: we have already estimated the linear bias in the previous section, Table 7, and we use the value $\delta_{\text{sc}} = 1.687$. We note that, in this analysis, as a reference value for the linear bias we will use the cross value $b_{\text{hm}}$ only, as

Figure 24: Linear bias for SO halos (left) and FoF halos (right) at redshift $z = 0$ for the Gaussian simulations with $\sigma_8 = 0.85$ for three mass bins, where we are using the 2Gpc/h box sets. We have corrected for the shot-noise in $P_{\text{hh}}$. 

138
the halo auto power spectrum $P_{hh}$ is hampered by shot-noise which may deviate from the constant value $1/\bar{n}$ we consider here. More discussion about this can be found in [260] and in Appendix 12.

The measurement of $b_{NG}^{PBS}$ requires to compute numerically the derivative of the halo mass function with respect to the normalisation amplitude $\sigma_8$. Using the 4 realizations of the 3 sets with Gaussian initial conditions with different amplitude $\sigma_8 = 0.83, 0.85, 0.87$ we can perform this task rather precisely, as shown in Figure 25. Going into more details, we compute this derivative in the following way

$$b_{NG}^{PBS}(M) = \frac{1}{4} \sum_{i=1}^{4} \frac{0.85}{2\bar{n}_i(M, \sigma_8 = 0.85)} \frac{\bar{n}_i(M, \sigma_8 = 0.87) - \bar{n}_i(M, \sigma_8 = 0.83)}{0.02}$$

and thus obtain $b_{NG}^{PBS}$ as a function of mass. To get a prediction for the three mass bins we are considering here, we weight the values of $b_{NG}^{PBS}$ within each bin with the halo mass function, we display the values for $b_{NG}^{PBS}$ and $b_{NG}^{DB}$ in Table 8 for the 2 Gpc/h box and the 1 Gpc/h box. To visualize better these results, we plot all these values in Figure 26 as a function of the central mass of the bin, calculated as in Eq. (273), and we also

Figure 25: Logarithmic derivative of the halo mass function as a function of $\sigma_8$ for both halo finder algorithms (upper), with relative difference (lower) for the 2Gpc/h box sets. Vertical dashed lines delimit the three mass bins we chose for the subsequent analysis. The black dotted lines are the results for the Tinker2010 and S&T99 fits.
plot the fits of Tinker2010, for SO halos, and S&T99, for FoF halos, where we use the DB prediction, \( b_{NG}^{fit} = \delta_{sc} b_{1}^{fit} \).

Figure 26: Non gaussian bias amplitude \( b_{NG} \) as a function of mass for both prescriptions and all the mass where we use the SO halo finder (left panel) and the FoF halo finder (right panel). Corresponding fits are marked in dotted black line (Tinker2010) and dashed-dotted black line (S&T99).

<table>
<thead>
<tr>
<th>SO halos</th>
<th>2 Gpc/h</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mass</td>
<td>( M )</td>
</tr>
<tr>
<td>9.2 – 14</td>
<td>11.1</td>
</tr>
<tr>
<td>14 – 27</td>
<td>18.9</td>
</tr>
<tr>
<td>27 – 3000</td>
<td>82.9</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>FoF halos</th>
<th>2 Gpc/h</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mass</td>
<td>( M )</td>
</tr>
<tr>
<td>9.2 – 14</td>
<td>11.1</td>
</tr>
<tr>
<td>14 – 27</td>
<td>18.9</td>
</tr>
<tr>
<td>27 – 3000</td>
<td>80.5</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SO halos</th>
<th>1 Gpc/h</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mass</td>
<td>( M )</td>
</tr>
<tr>
<td>1.1 – 1.8</td>
<td>1.43</td>
</tr>
<tr>
<td>1.8 – 3.9</td>
<td>2.62</td>
</tr>
<tr>
<td>3.9 – 1000</td>
<td>19.1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>FoF halos</th>
<th>1 Gpc/h</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mass</td>
<td>( M )</td>
</tr>
<tr>
<td>1.1 – 1.8</td>
<td>1.43</td>
</tr>
<tr>
<td>1.8 – 3.9</td>
<td>2.62</td>
</tr>
<tr>
<td>3.9 – 1000</td>
<td>18.3</td>
</tr>
</tbody>
</table>

Table 8: Measured values for \( b_{NG} \) for both prescriptions and both halo finder algorithms, where mass ranges are expressed in units of \( 10^{12} M_{\odot} \).
7.2 NON GAUSSIAN BIAS FROM N-BODY SIMULATIONS

Figure 27: Non gaussian bias as a function of the wavenumber $k$, $\Delta b_{\text{hm}}(k, f_{\text{NL}})$ measured for the three mass bins (where bin 1 is the least massive) at redshift $z = 0$ in the 2Gpc/h box sets of non Gaussian simulations with $f_{\text{NL}} = 250$. halos are identified with the SO algorithm (left panel) and FoF algorithm (right panel). The two prescriptions $b_{\text{DB}}^{\text{NG}}$ and $b_{\text{PBS}}^{\text{NG}}$ are shown in dotted and solid black line, respectively.

Figure 28: Same as Figure 27 for $f_{\text{NL}} = -250$. 

141
7.2.4 Final results and discussion

It is now timely to show the results for $\Delta \kappa^{\text{hm}} (k, f_{\text{NL}})$ in Figure 27 and 28 for positive and negative $f_{\text{NL}}$, respectively. As expected from Eq. (276), at large scales, $k \lesssim 0.1$ h/Mpc, the scale dependence becomes noticeable and we have verified that indeed the scaling is proportional to $1/k^2$. For all the mass bins and both halo finders, the general prediction $b_{\text{NG}}^{\text{PBS}}$ of Eq. (279) fits perfectly the data. On the other hand, the prediction $b_{\text{NG}}^{\text{DB}}$ of Eq. (278), which is assumed to be valid for universal mass functions, tends to underestimate the measurement, especially in the case of the SO halo finder and for the lowest mass bins. In this limit case, the disagreement is about 3-$\sigma$ away from the estimated measurement. Similar results were found in [227], where an overestimate around 40% was found on $f_{\text{NL}}$ in order to match the measurements, when the prescription $b_{\text{NG}}^{\text{DB}}$ was used.

Our results for FoF halos seem instead in tension with previous results [227, 249], which found that $b_{\text{NG}}^{\text{DB}}$ caused an underestimation of $f_{\text{NL}}$ and similar result where found in [261]. This disagreement is probably due to our choice of the FoF code: even though “Rockstar” uses a FoF technique to find halo centers, the estimation of the halo mass is done using a SO method: this corresponds to an “hybrid” algorithm that seems to perform quite well. For the highest mass bin, which ranges from $2.7 \times 10^{13} M_\odot$ to $3.0 \times 10^{15} M_\odot$, results are in good agreement with the measurement for both prescriptions.

These results raise some concern on the validity of the universality assumption for SO halos. The reason for such a disagreement for low mass halos may be due to the fact that, in this regime, the spherical approximation fails and therefore any algorithm which bases its technique on this assumption can not be trusted at percent level precision.

Nevertheless, results for the more general peak-background prescription, $b_{\text{NG}}^{\text{PBS}}$, are encouraging, as they seem to be independently performing well for any halo finder and for a wide range of masses. As we outlined in §7.2.3.1, estimating the logarithmic derivative of the halo mass function w.r.t. $\sigma_8$ requires the only measure of the halo mass function for Gaussian simulations with a different $\sigma_8$, which is not much more complicated and time consuming than estimating the linear bias, required for the $b_{\text{NG}}^{\text{DB}}$ prediction. Moreover, given that the prescription $b_{\text{NG}}^{\text{DB}}$ systematically underestimates the signal, in this setting, employing $b_{\text{NG}}^{\text{PBS}}$ for Fisher forecast should give better predictions on the precision with which we can measure $f_{\text{NL}}$. We leave this check for future work.

7.3 SQUEEZING THE HALO BISPECTRUM

Having discussed the effect of PNG on the two-point function of DM halos, we now turn on higher order clustering statistics such as the galaxy 3-point correlation function, which has recently been measured in [262, 263, 264, 265, 266]. This quantity clearly contains additional information on PNG, as it is a natural observable to constrain different PNG shapes, while it also provides a consistency check for the constraints obtained with the power spectrum [257, 242, 244, 238, 267, 268, 269, 270, 271, 272, 273, 274]. However,
modelling the scale and shape dependence of the galaxy bispectrum is very challenging (see e.g. [275, 276, 245, 217, 277, 278] for recent attempts in the context of PNG). Any non-linearity in the description of the galaxy number over-density induces non-Gaussianity in its distribution. One important source of such non-linearity arises from biasing. Furthermore, the nonlinear bias of LSS tracers generates scale-dependence and stochasticity which complicate the interpretation of the measurements.

To illustrate the complications brought by the scale-dependence of halo bias, we will focus on the squeezed limit of the halo-matter cross-bispectrum $B_{hmm}$ in the presence of a local-type PNG. Moreover, we will present measurements at the initial conditions only, to avoid the contamination induced by nonlinear clustering. The squeezed limit is particularly convenient because expressions greatly simplify, and the calculation becomes very tractable. Surprisingly, we find that is not trivial to get a good agreement with the data even in this simple limit, which can thus be used to discriminate between different models of halo bias.

Using the ESP model and the formalism of integrated perturbation theory [103], we derive expressions for the halo bispectrum in Lagrangian space $B_{hhh}$ and the bispectrum involving a halo and two matter over-density modes $B_{hmm}$, including the contribution of PNG. We compare our theoretical predictions to measurements of $B_{hmm}$ in the initial conditions of a series of N-body simulation with a large local-type PNG. In the squeezed limit, the signature of PNG is clean and very sensitive to the second- and third-order biasing. We find the measurements to be in reasonably good agreement with the ESP, while a simplistic local bias model fails to even qualitatively describe the simulations. For comparison, we also use a model in which the bias parameters are explicitly derived from a peak-background split, as described in Appendix 14. This model works better than local biasing, but does not capture the correct behaviour of the bispectrum when the wavemode corresponding to the matter fluctuation field is squeezed.

7.3.1 Halo bispectra in the presence of primordial non-Gaussianity

The main quantity of interest is the bispectrum $B_{hhh}$, i.e. the ensemble average of three halo number overdensity fields, since it is directly related to the bispectrum of galaxy number counts that can be extracted from galaxy survey data. However, there are some difficulties with this statistics: its measurement in simulations can be very noisy; it is affected by stochasticity; and we will see that, in the presence of PNG, there is no clear way to compute it from a Lagrangian bias expansion. Therefore, we will instead focus on the bispectrum $B_{hmm}$ involving one halo number overdensity and two DM overdensity fields, which suffers much less from those problems. We will perform the calculation in Lagrangian space where the Lagrangian halo biases are established. We will consider specifically the effect of a local-type PNG on these bispectra and illustrate its sensitivity to the biasing model. For convenience, we decompose $B_{hmm} = B_{hmm}^G + \Delta B_{hmm}^{NG}$ into a contribution generated by Gaussian initial conditions and by PNG.
In order to organise the calculation of correlation functions, we take advantage of the connection that exists between the iPT and the peak approach. In the spirit of iPT, Eq. (79) can be generalised to Eulerian space upon defining a multi-point propagator for biased tracers \[103],

\[
\left\langle \frac{\delta^n \delta_{\text{ESP}}(k, z)}{\delta \delta_L(k_1, z_1) \cdots \delta \delta_L(k_n, z_n)} \right\rangle = (2\pi)^3 \delta^{3-3n} \delta_L(k - k_{1...n}) \Gamma_{\text{ESP}}^{(n)}(k_1, \ldots, k_n; z),
\]

(282)

In Lagrangian space, i.e. in the limit \(z \to \infty\), they match exactly our renormalised ESP bias functions,

\[
\Gamma_{\text{ESP}}^{(n)}(k_1, \ldots, k_n; z \to \infty) \to c_L^{(n)}(k_1, \ldots, k_n; z_\ast).
\]

(283)

Moreover, they are similar to the multi-point propagators \(\Gamma_{\text{m}}^{(n)}\) of the matter distribution employed in [279]. Therefore, similar diagrammatic and counting rules apply.

The propagators \(\Gamma_{\text{m}}^{(n)}\) and \(\Gamma_{\text{ESP}}^{(n)}\) can be used to calculate \(N\)-point correlation functions of matter fields or biased tracers at any redshift. For instance, [280] derived expression for the bispectrum of the matter field in the presence of PNG. Similarly, [281] computed the halo bispectrum \(B_{\text{hhh}}\) within the iPT framework, using the multi-point propagator introduced in [214]. Here, we followed the same strategy to derive both \(B_{\text{hhh}}\) and \(B_{\text{hmm}}\) at 1-loop. We also checked that, in Lagrangian space, the results agree with a calculation based on ESP perturbative expansion.

In all our calculations, we keep track of terms proportional to the bispectrum \(B_L\) and trispectrum \(T_L\) of the linear density field since we are interested in the effect of PNG. For the sake of completeness, the full expressions of \(B_{\text{hmm}}\) and \(B_{\text{hhh}}\) at 1-loop are presented in Appendix 13. We will now summarise the relevant theoretical expressions.

7.3.1.1 Squeezed bispectra in the initial conditions in the presence of PNG

As mentioned previously, the focus of this work is on the bispectra \(B_{\text{hhh}}\) and \(B_{\text{hmm}}\). We compare the theoretical prediction for the squeezed limit bispectra against that measured on N-body simulation at the initial redshift \(z_i \sim O(100)\). Given that the measurement is done at high redshifts, we are sensitive to the local PNG without being too concerned about the modelling of non-Gaussianities induced by the gravitational evolution. Following Eq. (283), we will hereafter approximate \(\Gamma_{\text{ESP}}^{(n)}\) by \(c_L^{(n)}\), yet include the small nonlinearities in the matter through the usual PT kernels, \(\Gamma_{\text{m}}^{(n)} = F_{\text{m}}^{(n)}\). We can ignore the non-linear evolution of matter in the multi-point propagator \(\Gamma_{\text{ESP}}^{(n)}\) since it is suppressed with respect to the non-linear biasing by powers of the growth factor, which is small at high redshifts, as we will see below.

Let \(B_{\text{hmm}}\) and \(B_{\text{hmm}}\) be the cross-bispectra when the squeezed Fourier mode \(k_1 = k_l\) corresponds to the halo field \(\delta_h(k, z)\) and matter field \(\delta_m(k, z)\), respectively. Retaining
only the terms with the strongest divergent behaviour in the squeezed limit, the contribution induced by PNG reads

\[
\lim_{k_1 \to 0} \Delta B_{\text{hmnn}}^{\text{NG}}(k_1, k_s, k_s; z_i) \approx \left( \frac{D(z_i)}{D(z_s)} \right)^2 c_1^L(k_s) B_L(k_1, k_s, k_s) \\
+ \left( \frac{D(z_i)}{D(z_s)} \right)^3 P_L(k_s) F_m^{(2)}(k_s, k_1) \int \frac{d^3q}{(2\pi)^3} c_2^L(q, -q) B_L(k_1, q, q) \\
+ \frac{1}{2} \left( \frac{D(z_i)}{D(z_s)} \right)^2 \int \frac{d^3q}{(2\pi)^3} c_2^L(q, -q) T_L(k_s, -k_s, q, -q),
\]

where \( k_s \equiv k_2 \simeq -k_3 \) is the short mode and \( k_1 \equiv k_1 \) is the long mode. Moreover, the linear matter power spectrum \( P_L \), bispectrum \( B_L \) and trispectrum \( T_L \) are evaluated at the redshift \( z_s \) of halo collapse, like the Lagrangian bias coefficients \( c_1^L \). The powers of growth factors account for the redshift evolution of \( B_{\text{hmnn}} \). Since the initial conditions of the simulations are typically laid down at redshift \( z_i \sim O(100) \), and the collapse redshift of halo is in the range \( 0 \leq z_s \leq 2 \), each factor of \( D(z_i)/D(z_s) \) represents a two orders of magnitude suppression. This corresponds to the intuition that, at \( z_i \), the non-linear evolution of matter is small. We have kept next to leading order terms in this small parameter since, for very small squeezed Fourier mode, the stronger divergence of the trispectrum can compensate for this suppression. We call Eq. (284) the “halo squeezed” bispectrum.

Analogously, for the halo matter matter bispectrum, when we take the limit of one of the Fourier modes corresponding to a matter overdensity field to zero, we obtain

\[
\lim_{k_1 \to 0} \Delta B_{\text{mnmm}}^N(k_1, k_s, k_s; z_i) \approx \left( \frac{D(z_i)}{D(z_s)} \right)^2 c_1^L(k_s) B_L(k_1, k_s, k_s) \\
+ \frac{1}{2} \left( \frac{D(z_i)}{D(z_s)} \right)^3 P_L(k_s) c_2^L(-k_s, k_1) \int \frac{d^3q}{(2\pi)^3} F_m^{(2)}(q, -q) B_L(k_1, q, q) \\
+ \left( \frac{D(z_i)}{D(z_s)} \right)^3 \int \frac{d^3q}{(2\pi)^3} c_2^L(q, k_s - q) F_m^{(2)}(q, k_s - q) P_L(|k_s - q|) B_L(k_1, q, q) \\
+ \frac{1}{2} \left( \frac{D(z_i)}{D(z_s)} \right)^2 P_L(k_s) \int \frac{d^3q}{(2\pi)^3} c_2^L(-k_s, q, -q) B_L(k_1, q, -q) \\
+ \frac{1}{2} \left( \frac{D(z_i)}{D(z_s)} \right)^2 \int \frac{d^3q}{(2\pi)^3} c_2^L(q, k_s - q) T_L(-k_s, k_1, q, k_s - q).
\]

In this particular limit, the third order bias contribution is not suppressed relative to the tree level contribution. We call Eq. (285) the “matter squeezed” bispectrum.

Although we have stopped at one loop, the loop expansion does not in principle correspond to an expansion in a small parameter. However, higher loops will necessarily involve either higher powers of \( \Phi \) (such as two insertions of the primordial bispectrum, or a non-zero primordial 5-point function), or higher order contributions arising from

7.3 SQUEEZING THE HALO BISPECTRUM

145
the non-linear evolution of matter. The former are suppressed since the potential perturbations are $\Phi \sim 10^{-5}$, but they may become important at extremely large scales as they diverge with higher powers of $1/k_l$. The latter will be suppressed by factors of $D(z_i)$ corresponding to the smallness of nonlinearities in the evolution of matter at $z \sim z_i$.

Figure 29: The halo-matter-matter bispectrum, where the squeezed Fourier mode corresponds to the halo overdensity. The primordial non-Gaussianity is parameterised by $f_{\text{NL}} = 250$, $g_{\text{NL}} = 0$ and $\tau_{\text{NL}} = (6f_{\text{NL}}/5)^2$. The triangular configuration is chosen to have two sides equal to $k_s$ and the third equal to $k_l$. The bispectrum is shown for halos of mass $M = 1.8 \times 10^{14} M_{\odot} h^{-1}$ at redshift $z_i = 0$. On the left we plot the bispectra as a function of $k_l$, fixing $k_s = 0.38 \text{ Mpc}^{-1} h$ while on the left we fix $k_s = 0.006 \text{ Mpc}^{-1} h$. Different lines correspond to the dominant contributions given in Eq. (286).

7.3.1.2 The specific case of a local type PNG

Eqs. (284) and (285) are straightforwardly specialised to local PNG using the expressions for the bispectrum and trispectrum. For the halo squeezed case, we get

$$\lim_{k_l \to 0} \Delta B_{\text{hmm}}^{NG}(k_1, k_s, k_s; z_i) \approx \frac{4f_{\text{NL}}}{M(k_1) D^3(z_s)} \int \frac{d^3q}{(2\pi)^3} c_L^2(q, -q) P_L(q), \quad (286)$$

$$146$$
whereas, for the matter squeezed bispectrum, we obtain

\[
\lim_{k_l \to 0} \Delta B_{\text{NG}}^{\text{hm}}(k_l, k_s, k_s; z_i) \approx \frac{4 f_{NL}}{M(k_l)} \frac{D^2(z_i)}{D^3(z_s)} P_L(k_s) P_L(k_l) c_L^T(k_s) \\
+ \frac{2 f_{NL}}{M(k_l)} \frac{D^2(z_i)}{D^3(z_s)} P_L(k_s) P_L(k_l) \int \frac{d^3q}{(2\pi)^3} c_L^S(-k_s, q, -q) P_L(q) \\
+ \frac{27 g_{NL} + 25 \tau_{NL}}{9 M(k_l)} \frac{D^2(z_i)}{D^3(z_s)} P_L(k_l) \int \frac{d^3q}{(2\pi)^3} c_L^2(q, k_s - q) \\
\times \left[ \frac{M(|k_s - q|)}{M(k_s) M(q)} P_L(|k_s - q|) P_L(k_s) + \frac{M(q)}{M(k_s) M(|k_s - q|)} P_L(|k_s - q|) P_L(k_s) \right],
\]

(287)

where we have only retained the dominant terms (neglecting those which involve \(F^{(2)}\) since they are suppressed by a factor of \(D(z_i)/D(z_s)\)) together with the trispectrum contribution.

In Figs. 29 and 30 we plot the different contributions to Eqs. (286) and (286). Here and henceforth, we always divide the bispectrum by a factor of \((2\pi)^6\). For the halo squeezed bispectrum, we indeed see that the trispectrum loop dominates at larger scales due to its stronger divergence. At those large scales, we have also explicitly checked that the additional loop corrections are indeed suppressed. Higher order primordial correlation functions will have even stronger divergences, but these have to overcome a suppression of higher powers of \(\Phi\). Therefore, we do not expect them to be significant on the scales probed by surveys, even in the presence of large PNG. For the matter squeezed bispectrum, the loop involving the bispectrum is of the same order as the tree level. However, we expect higher order loops to be suppressed by the non-linearity of the DM at \(z = z_i\).
Figure 30: The halo-matter-matter bispectrum, where the squeezed Fourier mode corresponds to the matter overdensity. The bispectra are calculated for the same non-Gaussian parameters and in the same configuration as in Fig 29. The bispectra are shown for halos at $z_*=0$. The top panel corresponds to the halos of mass $M = 2.2 \times 10^{13} M_\odot h^{-1}$ while ones at the bottom are for halos of mass $M = 1.8 \times 10^{14} M_\odot h^{-1}$. Different lines correspond to the dominant contributions given in Eq. (287).
Let us now turn to the halo bispectrum. An analogous calculation to that performed above gives

\[
\lim_{k_i \to 0} \Delta B_{\text{halo}}^{NG}(k_1, k_2, k_3; z_i) \approx \frac{1}{D(z_s)} \frac{4f_{NL}}{\mathcal{M}(k_1)} c_1^4(k_1) P_L(k_1) \left( c_1^4(k_s) \right)^2 P_L(k_s) \\
+ \frac{1}{D(z_s)} \frac{4f_{NL}}{\mathcal{M}(k_1)} c_1^4(k_1) P_L(k_1) c_1^4(k_s) P_L(k_s) \int \frac{d^3 q}{(2\pi)^3} \left[ c_2^4(q, k_s - q) \right]^2 P_L(|k_s - q|) P_L(q) \\
+ \frac{1}{D^2(z_s)} \frac{4f_{NL}}{\mathcal{M}(k_1)} c_1^4(k_1) P_L(k_1) c_1^2(k_s) c_1^2(k_l) P_L(k_s) P_L(k_l) \int \frac{d^3 q}{(2\pi)^3} c_2^2(q, -q) P_L(q) \\
+ \frac{1}{D^2(z_s)} \frac{50f_{NL}}{\mathcal{M}^2(k_1)} P_L(k_1) P_L(k_s) \left[ c_1^4(k_s) \right]^2 \int \frac{d^3 q}{(2\pi)^3} c_2^2(q, -k_s - q) \\
\times \left[ \frac{\mathcal{M}(|k_s + q|)}{\mathcal{M}(k_1)} P_L(k_s) P_L(q) + \frac{\mathcal{M}(q)}{\mathcal{M}(k_1)} \mathcal{M}(q) P_L(k_s) P_L(|k_s + q|) \\
+ \frac{\mathcal{M}(k_s)}{\mathcal{M}(q) \mathcal{M}(|k_s + q|)} P_L(q) P_L(|k_s + q|) \right].
\] (288)
Let us make a few comments regarding this expression. Firstly, unlike $\Delta B_{hhhh}^{NG}$, one-loop terms are not suppressed by any small parameter in Eq. (288). This is apparent in Fig. 31, in which we separately plot the contributions of the tree level, the loops involving the primordial bispectrum and those involving the primordial trispectrum. One-loop terms involving a primordial bispectrum are clearly large compared to the tree level. This is compatible with the result of [281], who found that the situation is even worse at two loops. This suggests that there may be no perturbative expansion in this case. A possible cause of this problem is the fact that the bias parameters have not been appropriately defined in the presence of PNG, i.e. they have not been appropriately “renormalised”. In the Gaussian case, the renormalisation of the ESP bias coefficients $c_{ijkqlm}$ is naturally taken care of by the orthonormal polynomials. However, these polynomials are orthonormal only w.r.t. Gaussian weights such as Normal or chi-square distributions. In the presence of PNG, the linear bias $c_1^L(k)$ will for instance receive contributions from integrals involving the skewness. We leave a thorough treatment of this problem for future work.

Secondly, the term in the forth line turns out to be proportional to the usual peak-background split non-Gaussian bias [51]. In the peak formalism, the non-Gaussian bias is obtained from a one-loop integration of the primordial bispectrum and the second order bias, like in iPT. As shown in [83], the relation

$$\int \frac{d^3q}{(2\pi)^3} c_L^2(k, \sigma) P_L(q) = \left( \frac{\partial \ln \bar{n}}{\partial \ln \sigma_8} \right). (289)$$

holds exactly for a constant deterministic barrier, while it appears that it is not satisfied in the ESP approach with moving barrier [6]. This issue will be addressed in future work. Here, we note that, even in the case of a moving barrier, one still expects Eq. (289) to hold reasonably well for halos with a mass larger than a few $M^*(z^*)$. All the measurements presented here satisfy this condition.

Thirdly, the terms in the first three lines of Eq. (288) nearly add up to

$$\lim_{k_l \to 0} \Delta B_{hhhh}^{NG} = \frac{1}{D(z^*)} \frac{4f_{NL}^2}{M(k_1)} c_1^L(k_1) P_L(k_1) P_h(k_3). \quad (290)$$

where $P_h(k)$ is the power spectrum of proto-halo centres at one loop. However, the term in the second line has the wrong coefficient for this simplification to take place (it should come with a factor of $1/2$). There has been a claim in the literature [282] that a result similar to Eq. (290) holds for the matter overdensity bispectrum. However, it is straightforward to realise that, at one loop in standard perturbation theory, the algebra for computing the matter bispectrum is the same as that used here. Therefore, the result would be the same upon replacing the bias coefficients by the DM non-linear evolution kernels. Hence, at the one loop level in standard perturbation theory, the non-Gaussian part of the result of [282] does not hold. This is somewhat unsurprising since, unlike the Gaussian part of their result (which we agree with), there is no symmetry argument to relate the bispectrum to the power spectrum [133, 282, 283].
Finally, the discrete nature of the proto-halo centres induces shot-noise corrections. We have ignored them here since our main focus is on the cross-bispectrum $B_{hhmm}$, which involves only one halo field. In $B_{hshm}$, these shot noise corrections could be modelled from first principle using peak theory, along the lines of e.g. [68].

7.3.1.3 Other theoretical approaches

To emphasise the importance of Lagrangian $k$-dependent bias contributions, we will also compare our measurements to a local bias approach and to the S12i model.

The local bias model In the local bias approach, $\Delta B_{hmm}$ and $\Delta B_{mhm}$ are still given by Eqs. (286) and (287), yet we turn off the scale-dependent terms in the Lagrangian bias functions $c_n^L(k_1, \ldots, k_n)$. Therefore, we have

$$c_n^L(k_1, \ldots, k_n) \equiv b_n$$

in the local bias model, where $b_n = b_{n00}$ is computed from the ESP approach.

The S12i model In addition to the ESP and local bias approach, we will also consider a model for the halo-matter bispectrum in which the NG bias parameters are explicitly obtained from a peak-background split (PBS). In this model, the signatures of PNG are encoded both in the NG matter bispectrum and in the NG bias parameters predicted by the PBS. Previous studies [209, 245] have already derived the PBS bias parameters and used them to predict the halo bispectrum in the presence of PNG. Here, we will follow the derivation of [207]. Although our final prediction eventually agrees with the standard PNG bias formula, we will refer to this model as S12i, to signify that it was inspired by the general derivation given in [207]. We will now summarise the main ingredients of this model. Details can be found in Appendix 14.

For the halo-squeezed case, we adopt the tree-level-only bispectrum

$$
\lim_{k_l\to 0} \Delta B_{hmm}^N(k_l, k_s, k_s; z_i) = \left( \frac{D(z_i)}{D(z_\ast)} \right)^2 \left[ b_1^{(1)}(z_\ast) + b_1^{(2)}(k_l, z_\ast) \right] \tilde{W}_R(k_l) B_{L}(k_s, k_s, k_l) 
+ \left( \frac{D(z_i)}{D(z_\ast)} \right)^3 b_1^{(2)}(k_l, z_\ast) \tilde{W}_R(k_l) B_{mG}^C(k_s, k_s, k_l; z_\ast),
$$

where $b_1^{(1)}$ is the usual Gaussian PBS bias parameter, $b_1^{(2)}$ is the NG bias parameter and $B_{mG}^C$ is the matter bispectrum induced by gravitational nonlinearities, and given by Eq. (406) at tree-level. We will use the scale-independent bias parameter measured from the Lagrangian cross power spectrum between halo and matter as an estimate for $b_1^{(1)}$.

In [207], the NG bias parameter $b_1^{(2)}$ is derived from the excursion set theory in a general setting (see the review in Appendix 14). As shown in [207], under the assumption of Markovianity of the excursion set walk and the universality of the mass function,
the general expression for $b_1^{(2)}$ given in Eq. (402) reduces to the well-known formula Eq. (403) [49, 51, 50] in the low-$k$ approximation. Although Eq. (402) is quite general, it is technically more difficult to evaluate than Eq. (403) because an accurate measurement of the numerical mass function is required. In particular, as discussed in Appendix 14, for halos resolved with few particles (group 1 in our case), it is numerically more accurate to compute $b_1^{(2)}$ using Eq. (403) instead. Therefore we shall evaluate $b_1^{(2)}$ using Eq. (403), while $b_1^{(1)}$ is obtained from the Gaussian simulations.

In the matter-squeezed case, in addition to the tree level bispectrum, we also include the 1-loop correction proportional to $b_3$

$$\lim_{k_l \to 0} \Delta B_{\text{NG}}^{\text{hmh}}(k_l, k_s, k_s; z_i) = \left( \frac{D(z_i)}{D(z_s)} \right)^2 \left[ b_1^{(1)}(z_s) + b_1^{(2)}(k_s, z_s) \right] \tilde{W}_R(k_s) B_L(k_s, k_l, k_s)$$

$$+ \left( \frac{D(z_i)}{D(z_s)} \right)^3 b_1^{(2)}(k_s, z_s) \tilde{W}_R(k_s) B^G_m(k_s, k_l, k_s; z_s)$$

$$+ \left( \frac{D(z_i)}{D(z_s)} \right)^2 \frac{b_3^{\text{ESP}}(z_s)}{2} \tilde{W}_R(k_s) P_L(k_s) \int \frac{d^3q}{(2\pi)^3} \tilde{W}_R(q) \tilde{W}_R(|k_l - q|) \times B_L(-k_l, q, k_l - q).$$

The reason for including this $b_3$-loop correction is because in the matter-squeezed case, the halo field can be quite nonlinear and hence the high order bias parameters can be important; we find that the analogous term in the peak model calculations is significant in the matter-squeezed case. However, the value of $b_3$ sensitively depends on the prescriptions used to compute it. In Eq. (410), we will use the ESP result, i.e. $b_3^{\text{ESP}} = b_{300}$. We have checked that when $b_3$ is computed using Mo & White (MW) [284] and Sheth & Tormen (ST) [285] bias parameters, the $b_3$-loop often worsens the agreement with the simulation results, while when the ESP result is used for $b_3$ we find the agreement often improved compared with tree-level-only results. It is worth stressing that this is one of the few examples where we find the results are sensitive to $b_3$ and hence able to differentiate different schemes used to compute it (see also the measurements of $b_3$ using cross-correlations or the separate Universe simulations [286, 287]).

### 7.3.2 Comparison to numerical simulations

In this section, we confront our model predictions based on the ESP approach to the numerical simulation results, and contrast the ESP predictions with those obtained with the local bias and $S_{12}$i models. We stress that none of the models considered here has remaining free parameters that can be fitted to the bispectrum data.
Table 9: Halo samples used in this analysis. The mean mass $\bar{M}$ and the cross bias parameter measured from the Lagrangian cross power spectrum $b_c^*$ are shown.

<table>
<thead>
<tr>
<th>Mass range ($10^{12} M_\odot h^{-1}$)</th>
<th>Bin 1: 3.8 − 9.2</th>
<th>Bin 2: 9.2 − 92</th>
<th>Bin 3: 92 − 920</th>
</tr>
</thead>
<tbody>
<tr>
<td>$z_*=2$</td>
<td>$M = 5.53$</td>
<td>$M = 15.9$</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>$b_c^* = 3.0$</td>
<td>$b_c^* = 4.6$</td>
<td>—</td>
</tr>
<tr>
<td>$z_*=1$</td>
<td>$M = 5.68$</td>
<td>$M = 20.0$</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>$b_c^* = 1.6$</td>
<td>$b_c^* = 2.5$</td>
<td>—</td>
</tr>
<tr>
<td>$z_*=0$</td>
<td>—</td>
<td>$M = 23.6$</td>
<td>$M = 193$</td>
</tr>
<tr>
<td></td>
<td>$b_c^* = 1.3$</td>
<td>$b_c^* = 2.5$</td>
<td></td>
</tr>
</tbody>
</table>

7.3.2.1 N-body simulations and halo catalogues

We use a series of cosmological simulations evolving 1536$^3$ particles in cubic box of size $L_{\text{box}} = 2000 \text{ Mpc} h^{-1}$. These simulations were run on the Baobab cluster at the University of Geneva. The cosmology is a flat $\Lambda$CDM model with $\Omega_m = 0.3$, $\sigma_8 = 0.85$, and $n_s = 0.967$. The transfer function was obtained from the Boltzmann code CLASS [161]. The initial particle displacements were implemented at $z_i = 99$ using a modified version of the public code 2LPTic [98, 207]. We shall use two different types of initial conditions for each of the four sets of realisations: Gaussian and PNG with $f_{\text{NL}} = 250$. We note that this value is much higher than the latest Planck constraint on $f_{\text{NL}}$, which reads $2.5 \pm 5.0$ for temperature data alone and $0.8 \pm 5.0$ when combining the temperature with the polarisation data [56]. We use a large value of $f_{\text{NL}}$ to highlight the impact of the local PNG on the clustering statistics of DM halos more easily. The simulations were evolved using the public code Gadget2 [99], and the DM halos identified with the spherical overdensity (SO) code AHF [251, 100]. We follow [96] and adopt a threshold of $\Delta = 200$ times the mean matter density of the Universe. In what follows, all the results are the average over four realisations, while the error bars represent the 1σ fluctuations among the realisations.

To construct the proto-halo catalogues, we trace the DM particles that belong to a virialized halo at redshift $z_*$ back to the initial redshift $z_i$. Their centre-of-mass position furnishes an estimate of the position of the proto-halo centre. We consider three different values of $z_* = 0, 1$ and 2, and split the halo catalogues into three different mass bins. The properties of the halo catalogues used in this analysis is shown in Table 9.

The proto-halo centres and the initial DM particles are interpolated onto a regular cubical grid using the Cloud-in-Cell (CIC) algorithm to generate the fields $\delta_h(x, z_i)$ and $\delta_m(x, z_i)$. The cross-bispectra $B_{\text{hmm}}$ and $B_{\text{mhm}}$ are computed following [257]. We consider isosceles triangular configurations with one long mode and two short modes, $(k_l, k_s, k_s)$, up to a maximum wavenumber equal to $120k_F$, where $k_F = 2\pi / L_{\text{box}}$ is the fundamental mode of the simulations.
facing the challenge of precision measurements of the amplitude of the non-Gaussian bias

7.3.2.2 Halo mass function and local bias parameters

The halo-matter cross-bispectra will be compared to those predicted by the ESP, local bias and the S12i model suitably averaged over halo mass. Namely, we will display

\[ B_{XYZ}(\bar{M}, k_l, k_s, z_i) \equiv \frac{1}{\mathcal{J}_{M_{\text{min}}}^{M_{\text{max}}}} \int_{M_{\text{min}}}^{M_{\text{max}}} \frac{dM}{\bar{n}_h(M)} \int_{M_{\text{min}}}^{M_{\text{max}}} dM \bar{n}_h(M) B_{XYZ}(k_l, k_s, z_i) , \]

where \( M_{\text{min}} \) and \( M_{\text{max}} \) are the minimum and maximum halo mass of a given bin. For the excursion set peaks and local bias predictions, we will use the ESP mass function of Eq. (66) to perform the mass averaging.

Fig. 32 shows the logarithmic mass function of our simulated SO halos at redshifts \( z_0, 1 \) and 2. The error bars show the scatter among the realisations. The curves represent the ESP theoretical predictions based on the square-root barrier \( B = \delta_{sc} + \beta \sigma_0 \) with lognormal scatter \( \beta \) described in Sec. §2.1. While our ESP, parameter-free prediction is reasonably good at redshift \( z_0 = 0 \), it underestimates the abundance of
massive halos at higher redshift. Note, however, that we have used the same mean \( \langle \beta \rangle \) and variance \( \text{Var}(\beta) \) at all redshift, even though these were inferred from halos which virialized at \( z_* = 0 \) only (see [96] for details). It is plausible that the mean and variance in the linear collapse threshold depend on redshift. In particular, an increase in the variance with redshift would amplify the Eddington bias (i.e. more low mass halos would be scattered into the high mass tail) and, therefore, improve the agreement with simulations. Notwithstanding, we will not explore this issue any further here, and use the barrier shape at \( z_* = 0 \) throughout.

For the S12i model, we use Eq. (403) with the Gaussian scale-independent bias directly measured in the simulations, so that the tree-level results are already averaged over the mass range of a halo bin. Note, however, that the \( b_3 \)-loop is evaluated at the mean mass of the halo bin.

We have not checked the extent to which the ESP predictions for the usual scale-independent, or local bias parameters \( b_n \equiv b_{n,00} \) agree with the simulations. Previous studies based on 1-point cross-correlation measurements [82, 84] and the separate Universe approach [287] have shown that the ESP predictions up to \( b_{300} \) are accurate at the \( \lesssim 10\% \) percent level for massive halos. However, it is pretty clear that the density peak approximation eventually breaks down at low mass [60]. This is also reflected in the measurement of \( \chi_1 \) performed by [1] which shows that, while \( \chi_1 \) is still negative for \( M \sim 10^{13} M_\odot h^{-1} \), it does not assume the value \( \chi_1 = -3/(2\sigma_1^2) \) predicted by the peak constraint.

### 7.3.2.3 Halo-matter cross-bispectra

We begin with a consistency check and plot, in Fig. 33, the measured halo-matter-matter bispectrum in the initial conditions (\( z_i = 99 \)) of the Gaussian simulations for the three collapse redshifts \( z_* = 0, 1, 2 \). The halos are chosen from the second mass bin in our simulations with the mass range and mean mass given in Table 9. The left and right panels show the case where the squeezed mode corresponds to the halo (\( B^G_{hmm} \)) and the matter overdensity (\( B^G_{mhm} \)), respectively. We can understand the figure using the tree level halo-matter-matter cross-bispectrum, which consists of two terms:

\[
\langle \delta_h(k_1, z_i)\delta_m(k_2, z_i)\delta_m(k_3, z_i) \rangle_G \approx \left( \frac{D(z_i)}{D(z_*)} \right)^2 c_L^2(k_2, k_3) P_L(k_2) P_L(k_3) + \left( \frac{D(z_i)}{D(z_*)} \right)^3 c_L^3(k_1) P_L(k_1) \left[ F_m^{(2)}(k_1, k_2) P_L(k_2) + F_m^{(2)}(k_1, k_3) P_L(k_3) \right]
\]

where the prime denotes the fact that we have neglected the factors of \( 2\pi \) and the Dirac delta function. We have ignored the intrinsic non-linearity in the gravitational evolution of the halo number over-density (see Eq. (283)), which is always subdominant in the initial conditions. In Fig. 33, we plot the ESP tree-level bispectrum. This bispectrum does not generally vanish away from the squeezed limit, i.e. for \( k_s \lesssim 0.1 \text{Mpc}^{-1} h \). As
$k_s$ increases however, the halo-squeezed or matter-squeezed bispectra rapidly decreases because of the power spectrum suppression. In the matter-squeezed case, the $c_{12}^2$-term is dominant over the $c_{12}^1$-term because of one less power of the growth factor. In the halo-squeezed case, say $k_2, k_3 \gg k_1$, because we generally have $P_L(k_2), P_L(k_3) \ll P_L(k_1)$, the second term is boosted by a factor $P_L(k_1)/P_L(k_2)$. Thus as the short mode $k_s$ increases, the $c_{12}^2$-term dominates at first, but the $c_{12}^1$-term takes over when the configuration is sufficiently squeezed. Overall, given large scatter of the data, the Lagrangian, halo-mass cross-bispectrum cannot easily distinguish between different models.

We will now compare the non-Gaussian contributions to the cross-bispectra $\Delta B_{\text{hmm}}^{\text{NG}}$ and $\Delta B_{\text{mhm}}^{\text{NG}}$ measured in the simulations to those predicted by the ESP, the local Lagrangian bias and the $S_{12i}$ model. Figs. 34 and 35 display the results for the halo-squeezed and matter-squeezed case, respectively. To ensure that the halo mass $M$ is significantly larger than $M_\star (z_\ast)$ and, thus, comply with the spherical collapse approximation, we consider the high mass bins (bin 2 and bin 3) at redshift $z_\ast = 0$. At $z_\ast = 1$ and 2 however, $M_\ast \lesssim 10^{12} M_\odot / h$ such that we use the two lowest mass bins (bin 1 and bin 2), whereas we discard bin 3 which, owing to the low number density of these massive halos, does not provide good statistics. We plot our results as a function of the wavenumber $k_s$ of the short mode while we keep the long mode fixed at $k_l = 2k_F$.

For the halo-squeezed case shown in Fig. 34, our results indicate that, while there is a noticeable discrepancy between the prediction of the local bias model and simulations, the predictions of both the ESP approach and the $S_{12i}$ model are a better fit to simulations. In this case, the dominant loop contribution is proportional to the second order bias. For matter squeezed case shown in Fig. 35, the predictions from the $S_{12i}$ model and the local bias model are qualitatively similar and neither of the two is in a good agreement with our simulations. In this case, the dominant loop contribution which is as large as the tree-level comes from the $b_3$-loop. In both models this contribution is included, assuming the third order bias $b_3$ to be constant. The fact that the $S_{12i}$ model works well in the halo-squeezed case, while it fails in the matter-squeezed limit suggests that the scale-dependence of $b_3$ can not be neglected.
Figure 33: The halo-matter-matter bispectrum for three redshifts $z_* = 0, 1, 2$ for Gaussian initial conditions. The halos are chosen from bin2 as defined in Table 9. The plots on the left show the bispectra when the squeezed Fourier mode corresponds to the halo overdensity while on the right we show the matter-squeezed case. The solid line is the theoretical prediction of the ESP model including only the tree-level Gaussian contribution while the points are simulation measurements. The triangular configuration is chosen such that two sides are equal to $k_s$ and the third equal to $k_l$. The long-wavelength mode is fixed at $k_l = 0.006 \, \text{Mpc}^{-1} h$ while varying the short mode $k_s$. 

\[ 7.3 \text{ SQUEEZING THE HALO BISPECTRUM} \]
Figure 34: Halo-matter-matter bispectrum, where the squeezed Fourier mode corresponds to the halo overdensity. The three rows correspond to redshifts $z^* = 0, 1, 2$. The mass range and mean masses of the three mass bins denoted as bin1, bin2 and bin3 are given in Table 9. The solid line is the ESP model prediction while the dashed line is the S12i prediction and the dashed-dotted line is the prediction of the local model. The triangular configuration is the same as in Fig. 33.
Figure 35: Halo-matter-matter bispectrum, where the squeezed Fourier mode corresponds to the matter overdensity. The mass bins and the redshifts are the same as in Fig. 34. The solid line is the ESP model prediction while the dashed line is the S12i model prediction and the dashed-dotted line is the prediction of the local model. The triangular configuration is the same as in Fig. 33.
Notwithstanding, although the ESP model furnishes the best fit to all the numerical data, they appear to overestimate strongly the measurements for the larger mass bin at $z_*=2$. The discrepancy is more pronounced for $\Delta B_{\text{mhm}}^{\text{NG}}$ than for $\Delta B_{\text{hmh}}^{\text{NG}}$. Even though we have not explored this issue in details, we suspect this might be due to the fact that, in our approximation, the bias coefficients are calculated from the Gaussian probability density whereas they should, in fact, be computed from the non-Gaussian PDFs. These $O(f_{\text{NL}}^2)$ corrections are expected to increase with the halo mass $M$ and the value of $f_{\text{NL}}$. Furthermore, they have a sign opposite to the first-order non-Gaussian contribution and, e.g., would lower the theoretical predictions for $f_{\text{NL}} > 0$.

7.3.3 Conclusion

In this section we have focused on halo bias and its impact on the halo-matter cross bispectrum, which involves one halo and two matter overdensity fields. Unlike the halo auto bispectrum, which is noisier and relatively difficult to model, the cross bispectrum offers a fairly clean probe of the halo bias. We have studied its squeezed limit in the presence of primordial non-Gaussianity (PNG) of the local type, which induces a characteristic $1/k^2$ scale-dependence. Our fiducial halo bias model is the excursion set peak (ESP) approach. In the squeezed limit, the ESP cross bispectrum can be easily computed using the integrated perturbation theory (iPT). We have also considered a simple local bias scheme, and a model ($S_{12i}$) in which the NG bias factors are explicitly computed from a peak-background split. In all cases, the model parameters are constrained with statistics other than the cross bispectrum. We have carried out the calculations and the measurements at the initial redshift of the simulations, in order to mitigate the contamination induced by the nonlinear gravitational evolution. This can be thought of as a good approximation for the Lagrangian space.

Interestingly, even the simple squeezed limit is not trivial. We have found that the results strongly depend on the scale-dependence of the quadratic and third-order bias functions. In the halo-squeezed limit (i.e. when the long mode corresponds to the halo fluctuation field), both the ESP and the $S_{12i}$ models are in reasonable agreement with the simulations (see Fig. 34). For the matter-squeezed case however, the ESP model fares significantly better than the $S_{12i}$ model (see Fig 35). This, however, does not call the peak-background split into question. After all, all the ESP bias factors can be derived from a peak-background split, as shown in [65, 66, 67]. Rather, this suggests that the $S_{12i}$ bias prescription suffers from inconsistencies. Finally, the local bias model, which is equivalent to keeping the scale-independent terms in the ESP bias functions, dramatically fails at reproducing the numerical results. While the dominant loop contribution to the halo-squeezed case is proportional to the second-order bias, the matter-squeezed limit is controlled by the contribution from the third-order bias. Hence, each of them furnishes a test of the modelling of the second and third order halo bias.
respectively. It should be noted that these loop corrections are as large as the tree level. Notwithstanding, this does not necessarily indicate the breaking of the perturbative expansion because the higher-order loops turn out to be suppressed (see Sec. §7.3.1.1).

There are many directions in which this work can be extended. First of all, our measurements are not what will be actually measured in real surveys. Future work will take into account the nonlinear gravitational evolution, which was omitted here for sake of clarity. Still, we do not expect our conclusions to change noticeably because gravitational nonlinearities cannot generate a signal in the squeezed limit, as is apparent from Fig. 33. In addition, we have found that, within the iPT, it is unclear whether a perturbative expansion in terms of small bias parameters holds, suggested that this formalism could be improved. This will involve the renormalisation of the bias coefficients in the presence of PNG along, e.g., the lines of [288] (a recent treatment of this issue with a different formalism). To perform a similar analysis in Eulerian space, which would match more closely what is measured in galaxy surveys, a self-consistent calculation of the halo bispectrum $B_{hhh}$ must be performed. Namely, this should include the effects of stochasticity induced, for example, by halo exclusion (see [260]). Furthermore, it would certainly be interesting to explore other triangular shapes, and consider PNG of the equilateral or orthogonal type. These shapes leave no imprint on the scale dependent bias, so that the galaxy bispectrum likely is our best hope for observing them in the future. Finally, the potential of the galaxy bispectrum for constraining PNG was investigated in previous works [242, 244, 243]. In these studies however, the biasing relation is modelled using simple local model. It would be interesting to revisit their analysis using more sophisticated bias models such as the ESP approach.
The fine understanding of the large scale structure of our Universe is a central theme in cosmology. Our knowledge in the field is nowadays (and will be even more in the future) dominated by data. Clustering statistics of the large scale structure provide a wealth of information both on the initial conditions of cosmic structure formation and on its subsequent gravitational evolution. Extracting robust information from galaxy correlations presents several challenges, of which galaxy bias is one of the most significant. The relation between the observed galaxies and the underlying mass distribution generally is non-linear and scale-dependent. Therefore, it is essential to consider extensions to the simple local bias model, in which the bias parameters are assumed to be constant. Since galaxies are expected to reside in DM halos, a first step for such studies is to investigate halo biasing, as halos can be more easily modeled theoretically. A theoretical study of halo biasing goes necessarily through the modeling of strongly non-linear physics: from the collapse process of overdense regions of the mass distribution forming virialized objects to the formation of structures on large scales from gravitational instability of small primordial perturbations, dealing with such challenges demands a hybrid approach which combines analytic and numerical tools to build solid predictions to support observations.

In this thesis, we have tried to tackle a few of these challenges: we have investigated in details the predictions of the excursion set peaks model, a biasing scheme that starts from the assumption that peaks of the initial Gaussian DM field are the preferential environments for the formation of virialized halos at later time. In this framework, all bias parameters describe the change of the tracer number density with respect to the perturbation of a particular quantity around the local background. The well known linear bias is the response to a change in the mean density, but we can define as well bias parameters related to the response of the tracer number density to a perturbation of the local curvature of the matter density field. The resulting scheme is composed by an effective expansion of non-linear and scale dependent bias parameters that can be used for a better understanding of the N-point correlation function of tracers. We have tested a few of these scale dependent bias parameters against numerical N-body simulations as well as in non standard cosmologies such as cosmologies with massive neutrinos. We have also tackled the issue of the existence of a statistical halo velocity bias which does not decay over time, a prediction which is matter of debate and far from being settled.

Understanding the clustering properties of the late Universe allows us also to glimpse the physics of the very early moments of history of the Universe. What mechanism produced the primordial perturbations we observe in the CMB and that provided the seed for the gravitational clustering and the formation of structures? As of today, we have no
answer to this interesting question. The most promising path seems to constrain signatures of primordial non-Gaussianity, namely if the random seeds of the initial perturbations are to be described with higher than 2-point statistics. Primordial non-Gaussianity is known to leave an imprint in the formation of structures, in particular it can locally modulate the amplitude of matter fluctuations generating a scale dependent bias whose signal is amplified at large scales. In this thesis, we have studied this effect in detail, providing new forecasts on the precision with which galaxy surveys, such as Euclid, will be able to constrain such signatures and analyzing the expected amplitude of this effect in numerical N-body simulations by measuring the power spectrum and bispectrum of halos and DM.

Most of the questions that have been addressed in this thesis remain unanswered. As cosmologists, we are well aware that our Universe is complicated and that a series of experimental and theoretical barriers preclude its understanding. Nevertheless, we are also aware that every little step counts and that great intuitions in the past have come in times when everything seemed incomprehensible (or completely understood). The work of a theoretical cosmologist is undoubtedly interesting: all the fields of theoretical physics, from relativity, to quantum mechanics, statistical mechanics, particle physics and more are needed for a complete study of the entire Universe. Moreover, now more than ever before, a theoretical cosmologist needs to be in contact with observed data, as we miss any direct evidence of the two great mysteries of our time: dark matter and dark energy. A great wealth of new information is expected to be extracted from forthcoming galaxy redshift surveys and a new stage of CMB probes which search for primordial gravitational waves signals. The challenge is therefore to search for new hints in these data that can allow us to solve those mysteries and unveil the dark side of the Universe.


Bibliography


Bibliography


Bibliography


Bibliography


Bibliography


[204] R. de Putter and O. Doré, Designing an Inflation Galaxy Survey: how to measure $\sigma(f_{NL}) \sim 1$ using scale-dependent galaxy bias, arXiv:1412.3854.
Bibliography


180


182
Bibliography


Bibliography


184


Bibliography


Part III

APPENDIX
THE ESP MODEL: USEFUL FUNCTIONS

8.1 THE CURVATURE FUNCTION OF DENSITY PEAKS

One can define in general the integrals

\[ G_\kappa^{(\alpha)}(\gamma_1, \gamma_1^1) \equiv \int_0^\infty \mathrm{d}x x^\kappa f_\alpha(x) P_G(x - \gamma_1^1; 1 - \gamma_1^2) \]

and the 0-th order is the curvature function of density peaks \[47\]

\[ G_0^{(\alpha)}(x) = \frac{1}{\alpha^4} \left\{ \frac{e^{-\frac{5\alpha^2}{2}}}{\sqrt{10\pi}} \left( \alpha x^2 - \frac{16}{5} \right) \right. \]

\[ + \frac{e^{-\frac{5\alpha^2}{8}}}{\sqrt{10\pi}} \left( \frac{31}{2} \alpha x^2 + \frac{16}{5} \right) + \frac{\sqrt{\alpha}}{2} (\alpha x^3 - 3\alpha x) \]

\[ \times \left[ \text{Erf} \left( \sqrt{\frac{5\alpha}{2} x} \right) + \text{Erf} \left( \sqrt{\frac{5\alpha}{2} x} \right) \right] \}. \]

Note that \[65\] introduced the extra variable \( \alpha \) in order to get a closed form expression for their 2-point peak correlation, while \[66\] showed that \( \alpha \neq 1 \) can be interpreted as a long-wavelength perturbation in \( I_2(x) \).

8.2 BIVARIATE \( \chi^2 \) DISTRIBUTIONS

We take the following expression for the bivariate \( \chi^2 \)-distribution

\[ \chi_2^2(x, y; \epsilon) = \frac{(xy)^{k/2-1}}{2^k \Gamma^2(k/2)} \left( 1 - \epsilon^2 \right)^{-k/2} e^{-\frac{x+y}{2\left(1-\epsilon^2\right)}} \]

\[ \times {}_0F_1 \left( \frac{k}{2}; \frac{e^2 xy}{4(1-\epsilon^2)^2} \right), \]

189
where $x$ and $y$ are distributed as $\chi^2$-variables with $k$ d.o.f., $\epsilon^2 \leq 1$ is their correlation and $_0F_1$ is a confluent hypergeometric function. On using the fact that modified Bessel functions of the first kind can be written as $I_\alpha(x) = i^{-\alpha}J_\alpha(ix)$, where

$$J_\alpha(x) = \frac{(x/2)^\alpha}{\Gamma(\alpha + 1)}\, _0F_1(\alpha + 1; -\frac{x^2}{4}) ,$$ (299)

the bivariate $\chi^2$-distribution can be reorganized into the product

$$\chi_k^2(x, y; \epsilon) = \chi_k^2(x)\chi_k^2(y|x; \epsilon) ,$$ (300)

where

$$\chi_k^2(y|x; \epsilon) = e^{-\frac{y + x^2\epsilon}{2(1-\epsilon^2)}}\left(\frac{y}{\epsilon^2x}\right)^{\alpha/2}I_{\alpha/2}\left(\frac{\epsilon\sqrt{xy}}{1-\epsilon^2}\right) ,$$ (301)

and $\alpha = k/2 - 1$. This conditional distribution takes a form similar to that of a non-central $\chi^2$-distribution $\chi_k^2(x; \lambda)$, where $\lambda$ is the non-centrality parameter. Fig. 36 displays $\chi_k^2(y|x; \epsilon)$ for several values of $x$, assuming $k = 3$ and $5$. Note that $\chi_k^2(y|x = k; \epsilon)$ is different from $\chi_k^2(y)$. 

190
Figure 36: Conditional chi-squared distribution $\chi^2_k(y|x;\epsilon)$ for 3 and 5 degrees of freedom. Results are shown for several values of $x$ and a fixed cross-correlation coefficient $\epsilon = 0.7$. The dashed (green) curve represents the unconditional distribution $\chi^2_k(y)$.

Using the series expansion of $\chi^2_k(x;\lambda)$ in terms of Laguerre polynomials, we arrive at

$$\chi^2_k(y|x;\epsilon) = \frac{e^{-y/(1-\epsilon^2)}}{2^k(1-\epsilon^2)^{k+1}} \left( \frac{y}{2} \right)^x \times \sum_{j=0}^{\infty} \frac{(-\epsilon^2)^j}{\Gamma\left(\frac{1}{2}k+j\right)} \left( \frac{x^2}{2} \right)^j L_j^{(\alpha)} \left[ \frac{y}{2(1-\epsilon^2)} \right].$$

This series expansion is used to obtain the right-hand side of Eq. (87).
In this Appendix, we give explicit expressions for the Lagrangian perturbative bias expansion appropriate to excursion set peaks, $\delta^L_{\text{ESP}}(x)$, and for the third-order Lagrangian ESP bias function, $c^L_3(k_1, k_2, k_3)$, which is relevant to the evaluation of the matter-squeezed cross-bispectrum $\Delta B_{\text{mhm}}$. The general methodology can be found in [66, 67]. We follow common practice and write the perturbative bias expansion in terms of the field $\delta$ and its derivative, rather than the normalised variables introduced in Sec. $\S$??.
\[ 
\delta^{\text{ESP}}_{\text{esp}}(x) = b_{100} \delta_R(x) - b_{010} \nabla^2 \delta_R(x) - b_{001} \frac{d \delta_R}{dR}(x) \\
+ \frac{1}{2} b_{200} \delta_R^2(x) + \frac{1}{2} b_{020} \left[ \nabla^2 \delta_R(x) \right]^2 + \frac{1}{2} b_{002} \left( \frac{d \delta_R}{dR} \right)^2(x) \\
- b_{110} \delta_R(x) \nabla^2 \delta_R(x) - b_{101} \delta_R(x) \frac{d \delta_R}{dR}(x) + b_{011} \nabla^2 \delta(x) \frac{d \delta_R}{dR}(x) \\
+ \lambda_1 \left( \nabla \delta_R \right)^2(x) + \frac{3}{2} \omega_{10} \left[ \partial_{ij} \delta_R - \frac{1}{3} \delta_{ij} \nabla^2 \delta_R \right]^2(x) \\
+ \frac{1}{3!} b_{300} \delta_R^3(x) - \frac{1}{3!} b_{030} \left[ \nabla^2 \delta_R(x) \right]^3 - \frac{1}{3!} b_{003} \left( \frac{d \delta_R}{dR} \right)^3(x) \\
- \frac{1}{2} b_{210} \delta_R^2(x) \nabla^2 \delta_R(x) + \frac{1}{2} b_{120} \delta_R(x) \left[ \nabla^2 \delta_R(x) \right]^2 \\
- \frac{1}{2} b_{201} \delta_R^2(x) \frac{d \delta_R}{dR}(x) + \frac{1}{2} b_{021} \left[ \nabla^2 \delta_R(x) \right]^2 \frac{d \delta_R}{dR}(x) \\
+ \frac{1}{2} b_{102} \delta_R(x) \left( \frac{d \delta_R}{dR} \right)^2(x) - b_{012} \nabla^2 \delta_R(x) \left( \frac{d \delta_R}{dR} \right)^2(x) \\
+ b_{111} \delta_R(x) \nabla^2 \delta_R(x) \frac{d \delta_R}{dR}(x) + c_{100100} \delta_R(x) \left[ \nabla \delta_R(x) \right]^2 \\
- c_{010100} \left[ \nabla \delta_R(x) \right]^2 \nabla^2 \delta_R(x) - c_{001100} \left[ \nabla \delta_R(x) \right]^2 \frac{d \delta_R}{dR}(x) \\
+ \frac{3}{2} c_{100010} \left[ \partial_{ij} \delta_R - \frac{1}{3} \delta_{ij} \nabla^2 \delta_R \right]^2(x) \frac{d \delta_R}{dR}(x) - \frac{3}{2} c_{010010} \left[ \partial_{ij} \delta_R - \frac{1}{3} \delta_{ij} \nabla^2 \delta_R \right]^2(x) \nabla^2 \delta_R(x) \\
- \frac{3}{2} c_{001010} \left[ \partial_{ij} \delta_R - \frac{1}{3} \delta_{ij} \nabla^2 \delta_R \right]^2(x) \frac{d \delta_R}{dR}(x) + \frac{45}{2 \sqrt{7}} \omega_{01} \left[ \partial_{ij} \delta_R - \frac{1}{3} \delta_{ij} \nabla^2 \delta_R \right]^3(x). 
\]

A few comments are in order. Firstly, the Lagrangian perturbative expansion generally is a series in orthonormal polynomials [67]. Here, we have only written the term with highest power for simplicity, with the implicit rule that all zero-lag correlators should be discarded in the evaluation of \( \langle \delta^{\text{esp}}_{\text{ESP}}(x_1) \delta^{\text{esp}}_{\text{ESP}}(x_2) \rangle \). Secondly, we have taken advantage of the fact that the bias coefficients \( c_{ijklm} \) sometimes simplify. In particular, \( c_{ijklm} = b_{ijkl}, c_{000000} = \chi_q \) and \( c_{0000lm} = \omega_{lm} \), where \( b_{ijkl}, \chi_q \) and \( \omega_{lm} \) are given in Sec. \( \S2.2 \). Finally, note that the last term of this expression is proportional to the third invariant \( J_3 = \langle 0 | \mathcal{L}(\zeta_t^3) \rangle \), where \( \zeta_t \) is the traceless part of the Hessian. The corresponding bias factor \( \omega_{01} \) is given by Eq. \( \langle 76 \rangle \) with \( \ell = 0 \) and \( m = 1 \), i.e.

\[ 
\sigma^2 \omega_{01} = \frac{5}{3 \sqrt{7}} \langle J_3 | pk \rangle ,
\]

where the ensemble average is performed at the locations of density peaks.
The first- and second-order bias functions \( c^L_1(k) \) and \( c^L_2(k_1, k_2) \) are given in Sec. §2.2.1. For sake of completeness, the third-order Lagrangian EŠP bias function is

\[
c^L_3(k_1, k_2, k_3) = \begin{align*}
&\left\{ b_{900} + b_{030} k_1^2 k_2^2 k_3^2 - b_{003} \frac{d \ln \tilde{W}_R}{d R}(k_1) \frac{d \ln \tilde{W}_R}{d R}(k_2) \frac{d \ln \tilde{W}_R}{d R}(k_3) \right. \\
&\quad + b_{210} \left( k_1^2 + 2 \text{ cyc.} \right) - b_{201} \left[ \frac{d \ln \tilde{W}_R}{d R}(k_1) + 2 \text{ cyc.} \right] - b_{021} \left[ k_1^2 k_2^2 \frac{d \ln \tilde{W}_R}{d R}(k_2) \frac{d \ln \tilde{W}_R}{d R}(k_3) + 2 \text{ cyc.} \right] \\
&\quad + b_{012} \left[ k_1^2 \frac{d \ln \tilde{W}_R}{d R}(k_2) \frac{d \ln \tilde{W}_R}{d R}(k_3) + 2 \text{ cyc.} \right] + b_{102} \left[ \frac{d \ln \tilde{W}_R}{d R}(k_2) \frac{d \ln \tilde{W}_R}{d R}(k_3) + 2 \text{ cyc.} \right] \\
&\quad + b_{120} \left( k_1^2 k_2^2 + 2 \text{ cyc.} \right) - b_{111} \left[ k_1^2 \frac{d \ln \tilde{W}_R}{d R}(k_2) + 5 \text{ perm.} \right] - 2c_{100100} \left( k_2 \cdot k_3 + 2 \text{ cyc.} \right) \\
&\quad - 2c_{010100} \left[ k_1^2 (k_2 \cdot k_3) + 2 \text{ cyc.} \right] - 2c_{001100} \left[ k_1 \cdot k_2 \frac{d \ln \tilde{W}_R}{d R}(k_3) + 2 \text{ cyc.} \right] \\
&\quad + c_{100010} \left[ (3(k_1 \cdot k_2)^2 - k_1^2 k_2^2) + 2 \text{ cyc.} \right] + c_{010010} \left[ k_1^2 (3(k_2 \cdot k_3)^2 - k_2^2 k_3^2) + 2 \text{ cyc.} \right] \\
&\quad - c_{001010} \left[ (3(k_1 \cdot k_2)^2 - k_1^2 k_2^2) \frac{d \ln \tilde{W}_R}{d R}(k_3) + 2 \text{ cyc.} \right] \\
&\quad - 5 \cdot 3^3 \omega_0 \left[ (k_1 \cdot k_2) (k_2 \cdot k_3) (k_3 \cdot k_1) - \frac{1}{3} \left( (k_1 \cdot k_2) k_3^2 + 2 \text{ cyc.} \right) + \frac{2}{9} k_1^2 k_2^2 k_3^2 \right] \\
&\quad \times \tilde{W}_R(k_1) \tilde{W}_R(k_2) \tilde{W}_R(k_3).
\end{align*}
\]

In the ESP implementation of [82], tophat filters appear whenever the indices \( i \) or \( k \) are non-zero, whereas Gaussian filters arise whenever the remaining \( j, q, l \) and \( m \) are \( \geq 1 \).
10

THE HALO BOLTZMANN EQUATION: DETAILS AND CALCULATIONS

10.1 CONDITIONAL PROBABILITIES

In this Appendix we present the techniques we use to compute conditional probabilities of initially Gaussian fields which develop non-Gaussian fluctuations as time evolves. This method is inspired by the path integral approach and a similar application can be found in [289].

Let us suppose there are \( n \) stochastic variables \( X_i \), with \( i = 1, \ldots, n \). We define the connected two-point correlators as

\[
\mu_{ij} = \langle X_i X_j \rangle_c
\]

and the connected three-point correlators as

\[
\mu_{ijk} = \langle X_i X_j X_k \rangle_c.
\]

Thanks to Bayes’ theorem, we can define the conditional probability that \( X_1 \) assumes a value given that the other variables assume defined values as

\[
P(X_1|X_2, \ldots, X_n) = \frac{P(X_1, X_2, \ldots, X_n)}{P(X_2, \ldots, X_n)},
\]

where

\[
P(X_2, \ldots, X_n) = \int dX_1 P(X_1, X_2, \ldots, X_n).
\]

The basic quantity to compute is therefore \( P(X_1, X_2, \ldots, X_n) \). Following Ref. [?], it can be expressed in the following way

\[
P(X_1, X_2, \ldots, X_n) = \int \mathcal{D}\lambda \, e^{i\lambda_i X_i} e^{-\frac{1}{2} \lambda_i \lambda_j \mu_{ij}} e^{\left(\frac{-i}{6} \lambda_i \lambda_j \lambda_k \mu_{ijk} \right)},
\]

where

\[
\mathcal{D}\lambda \equiv \int_{-\infty}^{\infty} \frac{d\lambda_1}{2\pi} \ldots \int_{-\infty}^{\infty} \frac{d\lambda_n}{2\pi}.
\]
Using the fact that

$$\lambda e^{i\lambda X} = -i\partial_X e^{i\lambda X}, \quad (312)$$

perturbing for small three-point correlators and defining with subscript $g$ the Gaussian counterparts, we find

$$P(X_1|X_2, \ldots, X_n) = P_g(X_1|X_2, \ldots, X_n)$$

$$- \frac{\mu_{ijk}}{6P_g(X_2, \ldots, X_n)} \partial_i \partial_j \partial_k P_g(X_1, X_2, \ldots, X_n)$$

$$+ \frac{P_g(X_1, X_2, \ldots, X_n)}{6P_g(X_2, \ldots, X_n)} \mu_{ijk} \int dX_1 \partial_i \partial_j \partial_k P_g(X_1, X_2, \ldots, X_n). \quad (313)$$

The Gaussian probability can be expressed as

$$P_g(X_1, X_2, \ldots, X_n) = \frac{1}{(2\pi)^n/2} \frac{1}{(\text{Det } C)^{1/2}} e^{-\frac{1}{2}X_i C_{ij}^{-1}X_j}, \quad (314)$$

where $C$ is the covariance matrix. We now write

$$\mu_{mqr} \partial_m \partial_q \partial_r = \mu_{111} \delta_1^3$$

$$+ 3\hat{\mu}_{1q} \partial_1 \partial_q \partial_r$$

$$+ 3\hat{\mu}_{1r} \partial_1 \partial_1 \partial_r$$

$$+ \hat{\mu}_{mqr} \partial_m \partial_q \partial_r, \quad (315)$$

198
where the hat in the correlator \( \hat{\mu}_{mqr} \) means that we are taking indices different from 1. Using equation (313) we get

\[
\langle X_1|X_2, \ldots , X_n \rangle = \int dX_1 X_1 P(X_1|X_2, \ldots , X_n)
\]

\[
= \langle X_1|X_2, \ldots , X_n \rangle_{g}
\]

\[
- \frac{\mu_{111}}{6P_g(X_2, \ldots , X_n)} \int dX_1 X_1 \partial_1^3 P_g(X_1, X_2, \ldots , X_n)
\]

\[
- \frac{3\hat{\mu}_{1qr}}{6P_g(X_2, \ldots , X_n)} \int dX_1 X_1 \partial_1 \partial_q \partial_r P_g(X_1, X_2, \ldots , X_n)
\]

\[
- \frac{3\hat{\mu}_{11r}}{6P_g(X_2, \ldots , X_n)} \int dX_1 X_1 \partial_1^2 \partial_r P_g(X_1, X_2, \ldots , X_n)
\]

\[
- \frac{\hat{\mu}_{mqr}}{6P_g(X_2, \ldots , X_n)} \int dX_1 X_1 \partial_m \partial_q \partial_r P_g(X_1, X_2, \ldots , X_n)
\]

\[
+ \frac{\langle X_1|X_2, \ldots , X_n \rangle_{g}}{6P_g(X_2, \ldots , X_n)} \mu_{111} \int dX_1 \partial_1^3 P_g(X_1, X_2, \ldots , X_n)
\]

\[
+ \frac{\langle X_1|X_2, \ldots , X_n \rangle_{g}}{6P_g(X_2, \ldots , X_n)} 3\hat{\mu}_{1qr} \int dX_1 \partial_1 \partial_q \partial_r P_g(X_1, X_2, \ldots , X_n)
\]

\[
+ \frac{\langle X_1|X_2, \ldots , X_n \rangle_{g}}{6P_g(X_2, \ldots , X_n)} 3\hat{\mu}_{11r} \int dX_1 \partial_1^2 \partial_r P_g(X_1, X_2, \ldots , X_n)
\]

\[
+ \frac{\langle X_1|X_2, \ldots , X_n \rangle_{g}}{6P_g(X_2, \ldots , X_n)} \hat{\mu}_{mqr} \int dX_1 \partial_m \partial_q \partial_r P_g(X_1, X_2, \ldots , X_n),
\]

that can be simplified in

\[
\langle X_1|X_2, \ldots , X_n \rangle = \int dX_1 X_1 P(X_1|X_2, \ldots , X_n)
\]

\[
= \langle X_1|X_2, \ldots , X_n \rangle_{g}
\]

\[
+ 0
\]

\[
+ \frac{\beta_{1qr}}{2P_g(X_2, \ldots , X_n)} \partial_q \partial_r P_g(X_2, \ldots , X_n)
\]

\[
+ 0
\]

\[
- \frac{\beta_{mqr}}{6P_g(X_2, \ldots , X_n)} \partial_m \partial_q \partial_r P_g(X_2, \ldots , X_n)
\]

\[
+ 0
\]

\[
+ 0
\]

\[
+ \frac{\langle X_1|X_2, \ldots , X_n \rangle_{g}}{6P_g(X_2, \ldots , X_n)} \beta_{mqr} \partial_m \partial_q \partial_r P_g(X_2, \ldots , X_n).
\]
Similarly,

\[
\langle X_1^2 | X_2, \ldots, X_n \rangle = \int dX_1 X_1^2 P(X_1|X_2,\ldots,X_n)
\]
\[
= \langle X_1^2 | X_2, \ldots, X_n \rangle_g
\]
\[
+ \frac{\hat{\mu}_{1qr}}{P_g(X_2,\ldots,X_n)} \partial_q \partial_r \int dX_1 X_1 P_g(X_1,X_2,\ldots,X_n)
\]
\[
- \frac{\hat{\mu}_{11r}}{P_g(X_2,\ldots,X_n)} \partial_r P_g(X_2,\ldots,X_n)
\]
\[
- \frac{\hat{\mu}_{mqr}}{6P_g(X_2,\ldots,X_n)} \partial_m \partial_q \partial_r \int dX_1 X_1^2 P_g(X_1,X_2,\ldots,X_n)
\]
\[
+ \langle X_1^2 | X_2, \ldots, X_n \rangle_g \hat{\mu}_{mqr} \partial_m \partial_q \partial_r P_g(X_2,\ldots,X_n).
\] (318)

10.2 CUMULANTS AT SECOND-ORDER

In Appendix 10.1 we have showed how to calculate the average of a stochastic variable which is conditioned by a number of other variables, taking into account its non-Gaussian nature.

In this Appendix we apply such a technique to compute the average of \( \vec{\nabla} \Phi \) on the peak up to second-order. This amounts to using (317), which requires the knowledge of the the cumulants \( \mu_{mqr} = \langle X_m X_q X_r \rangle \), where the \( X_i \) range among the seven Gaussian variables \( \delta, \vec{\nabla}_i \Phi, \) and \( \vec{\nabla}_i \delta \) (\( i = 1, 2, 3 \)). Here we are omitting the dependence on \( r \) and \( t \) as the variables are all evaluated at the same point. Because of rotational invariance, the only non vanishing cumulants are

\[
\langle \delta(\vec{r},t) \delta(\vec{r},t) \rangle, \quad \langle \delta(\vec{r},t) \vec{\nabla}_i \Phi(\vec{r},t) \vec{\nabla}_j \Phi(\vec{r},t) \rangle, \quad \langle \delta(\vec{r},t) \vec{\nabla}_i \delta(\vec{r},t) \vec{\nabla}_j \delta(\vec{r},t) \rangle.
\]

(319)
We show the explicit computation of one of them, the others are calculated in a similar fashion. Let us take for example \( \langle \delta \nabla_i \Phi \nabla_j \delta \rangle \). Going to Fourier space we have

\[
\langle \delta \nabla_i \Phi \nabla_j \delta \rangle = \int \frac{d^3k_1}{(2\pi)^3} \frac{d^3k_2}{(2\pi)^3} \frac{d^3k_3}{(2\pi)^3} \frac{d^3q}{(2\pi)^3} \left[ \langle \delta(q) \delta(k_1 - q) \delta(k_2) \delta(k_3) \rangle \right] F_2(q, k_1 - q) + \\
+ \text{cyc.} \left( i \alpha \frac{k_2}{k_2^2} \right) \left( -ik_3 \right) e^{-i(k_1 + k_2 + k_3) \cdot r}
\]

\[
= \alpha \int \frac{d^3k_1}{(2\pi)^3} \frac{d^3k_2}{(2\pi)^3} \frac{d^3k_3}{(2\pi)^3} P(k_1) P(k_2) P(k_3) \frac{k_2}{k_2^2} \left( \frac{k_2, k_3}{k_2^2} \right) + \\
- \alpha \int \frac{d^3k_1}{(2\pi)^3} \frac{d^3k_2}{(2\pi)^3} \frac{d^3k_3}{(2\pi)^3} P(k_1) P(k_3) \frac{k_2}{k_2^2} \left( \frac{k_2, (k_1 + k_2)}{k_2^2} \right),
\]

(320)

where we have used the usual standard perturbation theory kernel

\[
F_2(q_1, q_2) = \frac{5}{7} + \frac{1}{2} \left( \frac{q_1 \cdot q_2}{q_1 q_2} \right) + \frac{2}{7} \left( \frac{q_1 \cdot q_2}{q_1 q_2} \right)^2
\]

(321)

and the Poisson equation

\[
\nabla^2 \Phi = \alpha \delta.
\]

(322)

Let us define \( A \) by

\[
\langle \delta \nabla_i \Phi \nabla_j \delta \rangle = A \delta_{ij}.
\]

(323)

By taking the trace we get

\[
A = -\alpha \frac{17}{21} \int \frac{dk_1}{2\pi^2} \frac{dk_2}{2\pi^2} k_1^2 k_2^2 P(k_1) P(k_2) \left[ 1 + \frac{(k_1 + k_2) \cdot k_2}{|k_1 + k_2|^2} \right],
\]

(324)

where \( \mu = \cos \theta \), being \( \theta \) the angle between \( k_1 \) and \( k_2 \). We can now integrate over \( \mu \) and obtain

\[
A = -\alpha \frac{17}{21} \zeta_0^4.
\]

(325)

A similar calculation of the other cumulants gives

\[
\langle \delta \nabla_i \Phi \nabla_j \delta \rangle = \frac{34}{7} \zeta_0^4,
\]

(326)

\[
\langle \delta \nabla_i \Phi \nabla_j \Phi \rangle = \frac{\alpha^2}{3} K_1 \delta_{ij},
\]

(327)

\[
\langle \delta \nabla_i \Phi \nabla_j \Phi \rangle = -\alpha \frac{17}{21} \zeta_0^4 \delta_{ij},
\]

(328)

\[
\langle \nabla_i \Phi \nabla_j \Phi \nabla_i \Phi \rangle = -\frac{\alpha}{2} \langle \delta \nabla_i \Phi \nabla_j \Phi \rangle = -\frac{\alpha^3}{6} K_1 \delta_{ij},
\]

(329)

\[
\langle \nabla_i \delta \nabla_j \Phi \nabla_i \Phi \rangle = -\frac{\alpha^2}{3} K_2 \delta_{ij}.
\]

(330)
Here we have defined the quantities

\[\kappa_1 = \int \frac{dk_1dk_2}{4\pi^4} P(k_1)P(k_2) \times \]
\[\times \left[ \frac{5}{168 k_1^3 k_2^3} \begin{bmatrix} 2k_1 k_2 (k_1^2 + k_2^2) (3k_1^4 - 14k_1^2 k_2^2 + 3k_2^4) + 3(k_1^2 - k_2^2)^4 \left( \log \frac{|k_1 - k_2|}{k_1 + k_2} \right) \end{bmatrix} \right]. \]

\[\kappa_2 = \int \frac{dk_1dk_2}{4\pi^4} P(k_1)P(k_2) \left[ \frac{7}{3} k_1^2 k_2^2 + \right. \]
\[+ \left( k_1^2 + k_2^2 \right) \frac{5}{336 k_1^3 k_2^3} \begin{bmatrix} 2k_1 k_2 (k_1^2 + k_2^2) (3k_1^4 - 14k_1^2 k_2^2 + 3k_2^4) + 3(k_1^2 - k_2^2)^4 \left( \log \frac{|k_1 - k_2|}{k_1 + k_2} \right) \end{bmatrix} \right]. \]  

(331)

These quantities are finite: the integrand is finite for \( k_1 \to k_2 \), and for \( k_1 \) or \( k_2 \to 0 \) the power spectra lead to a finite result.

We are now ready to compute the average of \( \tilde{\nabla} \Phi \) up to second-order. We use (317) which contains the Gaussian probabilities, encoded by the covariance matrix for the variables \( (\delta, \tilde{\nabla} \Phi, \tilde{\nabla} \delta) \) (we understand the symmetric components of the covariance matrix)

\[ C = \begin{pmatrix} \sigma_0^2 & 0 & 0 \\ \frac{\alpha^2}{3} \sigma_{-1}^2 \mathbb{I}_3 & -\frac{\alpha}{3} \sigma_0^2 \mathbb{I}_3 \\ \frac{1}{3} \sigma_1^2 \mathbb{I}_3 \end{pmatrix}. \]  

(332)

By using (323) for all the cumulants we find

\[ \langle \tilde{\nabla} \Phi \rangle_{pk} = -\alpha \frac{\sigma_0^2}{\sigma_1^4} \tilde{\nabla} \delta + \frac{\delta \tilde{\nabla} \delta}{\sigma_0^2 \sigma_1^4} \left( \langle \delta \tilde{\nabla} \Phi \cdot \tilde{\nabla} \delta \rangle + \alpha \langle \delta \tilde{\nabla} \delta \cdot \tilde{\nabla} \delta \rangle \frac{\sigma_0^2}{\sigma_1^2} \right) \]
\[= -\alpha \frac{\sigma_0^2}{\sigma_1^4} \tilde{\nabla} \delta + \frac{3}{21} \sigma_0^2 \sigma_1^4 \delta \tilde{\nabla} \delta. \]  

(333)

In a similar fashion, we can compute the covariance

\[ \langle \nabla_i \Phi \cdot \nabla_j \Phi \rangle_{pk} = \frac{1}{3} \delta_{ij} \left[ \alpha^2 \left( \sigma_{-1}^2 - \frac{\sigma_0^4}{\sigma_1^4} \right) \right. \]
\[ + \left. \langle \delta \tilde{\nabla} \Phi \cdot \tilde{\nabla} \delta \rangle \frac{\delta_{pk}}{\sigma_0^2} + \alpha^2 \left( \langle \delta \tilde{\nabla} \delta \cdot \tilde{\nabla} \delta \rangle \frac{\sigma_0^2}{\sigma_1^2} + 2\alpha \langle \delta \tilde{\nabla} \Phi \cdot \tilde{\nabla} \delta \rangle \frac{\delta_{pk}}{\sigma_1^2} \right) \right] = \]
\[ = \alpha^2 \frac{1}{3} \delta_{ij} \left[ \left( \sigma_{-1}^2 - \frac{\sigma_0^4}{\sigma_1^4} \right) + \delta_{pk} \left( \kappa_1 \frac{1}{\sigma_0^2} - \frac{20 \sigma_0^4}{21 \sigma_1^4} \right) \right]. \]  

(334)
In this section we report the intermediate results needed to compute the conditioned expected value \( \langle \delta(\vec{r}', t) \big| \vec{\psi}(\vec{r}, t), p_k \rangle \) appearing in (157).

The formula needed to compute this conditioned probability is again (317). In this case, differently from what discussed in the previous section, also \( \delta(\vec{r}') \) (or, shortly, \( \delta' \)) appears among the variables and hence in the covariance matrix. In order to write down the new covariance matrix corresponding to (332) we order the variables as \( \{ \delta(\vec{r}'), \delta(\vec{r}), \vec{\nabla}\Phi(\vec{r}), \vec{\nabla}\delta(\vec{r}) \} \) and we denote \( \vec{\ell} = (\vec{r}' - \vec{r}) \), so that the covariance matrix reads

\[
C = \begin{pmatrix}
\sigma_0^2 & \xi(\ell) & \eta(\ell)\vec{\ell} & -\frac{d\xi(\ell)}{d\ell} \vec{\ell} \\
\sigma_0^2 & 0 & 0 & 0 \\
\frac{1}{3} \sigma_0^2 \mathbb{I}_3 & 1 & \frac{1}{3} \sigma_0^2 \mathbb{I}_3 \\
\frac{1}{3} \sigma_0^2 \mathbb{I}_3 & 1 \sigma_1^2 \mathbb{I}_3 & \frac{1}{3} \sigma_0^2 \mathbb{I}_3 \\
\end{pmatrix},
\]

where we have defined

\[
\xi(r) = \int \frac{d^3k}{(2\pi)^3} P(k) j_0(kr),
\]

\[
\eta(r) = \int \frac{d^3k}{(2\pi)^3} P(k) j_1(kr) \frac{1}{kr},
\]

and \( j_n(x) \) are the spherical Bessel functions. The following formulae for their integrals over the solid angle will be needed:

\[
\int d^3r \frac{r_i r_j}{r^3} \frac{1}{r} \frac{d\xi}{dr} = -\frac{4\pi}{3} \sigma_0^2 \delta_{ij},
\]

\[
\int d^3r \frac{r_i r_j}{r^3} \eta(r) = \frac{4\pi}{3} \sigma_1^2 \delta_{ij}.
\]
We can now write down the final result after the application of \((317)\). We mark in blue the terms that vanish with the subsequent integration \(\int d^3\ell \. K_1\). We also denote the cumulants with a compact notation, so that for example \(\mu_{\bar{\nu} \bar{\delta} \bar{\psi}} = \langle \delta(\bar{r'}) \bar{\psi} \bar{\delta}(\bar{r}) \rangle\).

\[
\langle \delta(\bar{r'},t) \bar{\psi}(\bar{r},t), \delta, \bar{\psi}, \bar{\delta} \rangle = \frac{\delta_1}{c_0^4} \xi(\ell) - \frac{3}{c_0^4 - c_0^2 \ell^2} \left( c_0^2 \eta(\ell) + \frac{\ell}{c_0^4} \frac{1}{\ell} \frac{d \xi}{d \ell} \right) \vec{\ell} \cdot \bar{\delta}_1 + \frac{1}{2} \mu_{\bar{\nu} \bar{\delta} \bar{\psi}} \bar{\nu}_{\bar{\psi}} \frac{\delta_1}{c_0^4 - c_0^2 \ell^2} \left( \frac{\delta_1}{c_0^4} \right)
\]

\[
+ 3 \bar{\nu}_{\bar{\gamma}} \bar{\nu}_{\bar{\delta}} + \left( \frac{\partial_1}{c_0^4 - c_0^2 \ell^2} \right) \left( \frac{\partial_1}{c_0^4 - c_0^2 \ell^2} \right) \frac{\delta_1}{c_0^4 - c_0^2 \ell^2} + \frac{1}{2} \mu_{\bar{\nu} \bar{\psi} \bar{\delta}} \left( \frac{\delta_1}{c_0^4 - c_0^2 \ell^2} \right)
\]

\[
+ \frac{1}{2} \mu_{\bar{\nu} \bar{\psi} \bar{\delta}} \left( \frac{\delta_1}{c_0^4 - c_0^2 \ell^2} \right)
\]

After performing the integral prescribed in \((157)\), we obtain the result \((158)\).

In \((158)\) there are two cumulants containing \(\delta(\bar{r'})\), \(\mu_{\bar{\nu} \bar{\psi} \bar{\delta}}\) and \(\bar{\mu}_{\bar{\nu} \bar{\psi} \bar{\delta}}\). We report here the result of the integrals

\[
\int d\ell \. \bar{\ell} \cdot \langle \delta(\bar{r'}) \bar{\delta}(\bar{r}) \rangle = \frac{1}{2} \bar{K} + \frac{10}{21} c_0^3 \ell^2
\]

\[
\int d\ell \. \bar{\ell} \cdot \langle \bar{\delta}(\bar{r'}) \bar{\delta}(\bar{r}) \rangle = \frac{17}{21} c_0^4
\]

where \(\bar{K} \) is defined in \((331)\).
Here we outline the details of the calculation of non local biases that we introduce in Chapter §4.1. We solve Eq. (170) order by order in perturbation theory up to third order.

### 11.1 First-order

At first-order from Eq. (170) we get

\[
\delta_h^{(1)}(x, \tau) - \delta^{(1)}(x, \tau) = 0, \tag{342}
\]

or

\[
\delta_h^{(1)}(x, \tau) = \delta^{(1)}(x, \tau_i) + \delta^{(1)}(x, \tau) - \delta^{(1)}(x, \tau_i), \tag{343}
\]

where \(\tau_i\) is some initial time. We assume that the initial bias expansion is local and depends only on the linear DM density contrast through the (Lagrangian) bias coefficients

\[
\delta_h(x, \tau_i) = \sum_\ell \frac{b_L^\ell(\tau_i)}{\ell!} \left( \delta^{(1)}(x, \tau_i) \right)^\ell
\approx \sum_\ell \frac{b_L^\ell(\tau)}{\ell!} \left( \delta^{(1)}(x, \tau) \right)^\ell, \tag{344}
\]

where \(b_L^\ell(\tau) = b_L^\ell(\tau_i)(a(\tau_i)/a(\tau))^\ell\). Using \(\delta_h^{(1)}(x, \tau_i) = b_L^1(\tau)\delta(x, \tau)\), we obtain the standard result

\[
\delta_h^{(1)}(x, \tau) \approx \left( 1 + b_L^1(\tau) \right) \delta^{(1)}(x, \tau). \tag{345}
\]
11.1.2 Second-order

At second-order we may use the first-order result to write Eq. (170) in the form

\[ \delta_h^{(2)}(x, \tau) - \delta^{(2)}(x, \tau) + \nabla \cdot \left[ \left( \delta^{(1)}_h(x, \tau) - \delta^{(1)}(x, \tau) \right) \mathbf{v}^{(1)}(x, \tau) \right] = 0, \]

which is solved by

\[
\begin{align*}
\delta_h^{(2)}(x, \tau) &= \delta_h^{(2)}(x, \tau) + \delta^{(2)}(x, \tau) - \int_\tau^\tau \eta \, b_1^2(\eta) \nabla \cdot \left[ \delta^{(1)}(x, \eta) \mathbf{v}^{(1)}(x, \eta) \right] \\
&\approx \frac{1}{2} b_2^2(\tau) \left( \delta^{(1)}(x, \tau) \right)^2 + \delta^{(2)}(x, \tau) - \int_{\tau}^{\tau} \eta \, b_1^2(\eta) \nabla \cdot \left[ \delta^{(1)}(x, \eta) \mathbf{v}^{(1)}(x, \eta) \right] \\
&+ \int_{\tau}^{\tau} \eta \, b_1^2(\eta) \delta^{(1)}(x, \eta) \delta^{(1)}(x, \eta) \\
&= \frac{1}{2} b_2^2(\tau) \left( \delta^{(1)}(x, \tau) \right)^2 + \delta^{(2)}(x, \tau) - \frac{\tau}{2} b_1^2(\tau) \nabla \delta^{(1)}(x, \tau) \cdot \mathbf{v}^{(1)}(x, \tau) \\
&+ b_1^2(\tau) \left( \delta^{(1)}(x, \tau) \right)^2,
\end{align*}
\]

or

\[
\begin{align*}
\delta_h^{(2)}(x, \tau) &= -\frac{1}{H} b_1^2(\tau) \nabla \delta^{(1)}(x, \tau) \cdot \mathbf{v}^{(1)}(x, \tau) + \frac{1}{2} b_2^2(\tau) \left( \delta^{(1)}(x, \tau) \right)^2 \\
&+ \delta^{(2)}(x, \tau) + b_1^2(\tau) \left( \delta^{(1)}(x, \tau) \right)^2.
\end{align*}
\]

To perform the time integrals, we have used the scalings provided by the first-order quantities (in matter-domination)

\[
\begin{align*}
\mathbf{v}_i^{(1)}(x, \tau) &= -\frac{\tau}{3} \partial_i \phi(x) = \frac{2}{3H} \partial_i \phi(x), \\
\delta^{(1)}(x, \tau) &= \frac{\tau^2}{6} \nabla^2 \phi(x) = \frac{2}{3H^2} \nabla^2 \phi(x),
\end{align*}
\]

where \( \phi(x) \) is the initial condition for the gravitational potential \( \Phi(x, \tau) \). In order to elaborate further the second-order halo density contrast, we remind the reader that \[138, 290, 291\]

\[
\begin{align*}
\delta^{(2)}(x, \tau) &= \frac{\tau^4}{2 \cdot 126} \left[ 5(\nabla^2 \phi(x))^2 + 2 \partial_x \partial_y \phi(x) \partial^x \partial^y \phi(x) + 7 \partial^x \phi(x) \nabla^2 \partial^y \phi(x) \right] \\
&= \frac{5}{7} \left( \delta^{(1)}(x, \tau) \right)^2 + \frac{8}{63H^4} \partial_x \partial_y \phi(x) \partial^x \partial^y \phi(x) - \frac{1}{H} \nabla \delta^{(1)}(x, \tau) \cdot \mathbf{v}^{(1)}(x, \tau)
\end{align*}
\]
and define the non-local bias operator
\[ s_{ij}(x, \tau) = \frac{2}{3H^2} \partial_i \partial_j \Phi(x, \tau) - \frac{1}{3} \delta_{ij} \delta(x, \tau), \] (350)
to get
\[ \partial_k \partial_\rho \varphi(x) \partial^k \partial^\rho \varphi(x) = \frac{9H^4}{4} \left( s^{(1)}(x, \tau) \right)^2 + \frac{3H^4}{4} \left( \delta^{(1)}(x, \tau) \right)^2. \] (351)
From this expression we deduce
\[ \delta^{(2)}(x, \tau) = \frac{17}{21} \left( \delta^{(1)}(x, \tau) \right)^2 + \frac{2}{7} \left( s^{(1)}(x, \tau) \right)^2 - \frac{1}{H} \nabla \delta^{(1)}(x, \tau) \cdot \mathbf{v}^{(1)}(x, \tau). \] (352)
We finally arrive at
\[ \delta^{(2)}_h(x, \tau) \simeq \left( 1 + b_1^T(\tau) \right) \delta^{(2)}(x, \tau) + \left( \frac{1}{2} b_2^T(\tau) + \frac{4}{21} b_1^T(\tau) \right) \left( \delta^{(1)}(x, \tau) \right)^2 - \frac{2}{7} b_1^T(\tau) \left( s^{(1)}(x, \tau) \right)^2. \] (353)
This result reproduces exactly the one derived in Refs. [63, 109] and shows that at second–order in perturbation theory the halo overdensity is not a local function of the underlying matter overdensity.

### 11.1.3 Third-order

We now proceed to the original part of the computation at third-order. The equation to solve is
\[ \dot{\delta}^{(3)}_h(x, \tau) - \delta^{(3)}(x, \tau) + \nabla \cdot \left[ \left( \delta^{(1)}_h(x, \tau) - \delta^{(1)}(x, \tau) \right) \mathbf{v}^{(2)}(x, \tau) \right] + \nabla \cdot \left[ \left( \delta^{(2)}_h(x, \tau) - \delta^{(2)}(x, \tau) \right) \mathbf{v}^{(1)}(x, \tau) \right] = 0, \] (354)
where [290, 291]
\[ \mathbf{v}^{(2)}_i(x, \tau) = \frac{\tau^3}{18} \left[ \begin{array}{c} -\partial_i \partial_j \varphi(x) \partial^j \varphi(x) - \frac{3}{7} \partial_i \left( (\nabla^2 \varphi(x))^2 - \partial_\rho \varphi(x) \partial^k \partial^\rho \varphi(x) \right) \\
- \partial_j \nabla^2 \varphi(x) \partial^i \varphi(x) - \partial_i \partial_j \varphi(x) \partial^j \partial^i \varphi(x) - \frac{3}{7} \left( (\nabla^2 \varphi(x))^2 - \partial_\rho \varphi(x) \partial^k \partial^\rho \varphi(x) \right) \end{array} \right] \] (355)
Eq. (354) gives

\[
\delta_h^{(3)}(x, \tau) = \delta_h^{(3)}(x, \tau) + \int^\tau d\eta \, b_1^3(\eta) \vec{\nabla} \cdot \left[ \delta^{(1)}(x, \eta) \vec{v}^{(2)}(x, \eta) \right] \\
- \int^\tau d\eta \, b_1^2(\eta) \vec{\nabla} \cdot \left[ \delta^{(2)}(x, \eta) \vec{v}^{(1)}(x, \eta) \right] \\
- \int^\tau d\eta \, \vec{\nabla} \cdot \left[ \left( \frac{1}{2} b_2^1(\eta) + \frac{4}{21} b_1^3(\eta) \right) \left( \delta^{(1)}(x, \eta) \right)^2 - \frac{2}{7} b_1^1(\tau) \left( \delta^{(1)}(x, \tau) \right)^2 \right] \vec{v}^{(1)}(x, \eta),
\]

(356)

or

\[
\delta_h^{(3)}(x, \tau) = \frac{1}{3!} b_3^3(\tau) \left( \delta^{(1)}(x, \tau) \right)^3 + \delta^{(3)}(x, \tau) - \int^\tau d\eta \, b_1^2(\eta) \vec{\nabla} \delta^{(1)}(x, \eta) \cdot \vec{v}^{(2)}(x, \eta) \\
- \int^\tau d\eta \, b_1^2(\eta) \delta^{(1)}(x, \eta) \vec{\nabla} \cdot \vec{v}^{(2)}(x, \eta) - \int^\tau d\eta \, b_1^2(\eta) \vec{\nabla} \delta^{(2)}(x, \eta) \cdot \vec{v}^{(1)}(x, \eta) \\
- \int^\tau d\eta \, \vec{\nabla} \cdot \left( \frac{1}{2} b_2^1(\eta) + \frac{4}{21} b_1^3(\eta) \right) \left( \delta^{(1)}(x, \eta) \right)^2 \vec{v}^{(1)}(x, \eta) \\
+ \frac{2}{7} \int^\tau d\eta \, b_1^2(\eta) \vec{\nabla} \left( \delta^{(1)}(x, \eta) \right)^2 \cdot \vec{v}^{(1)}(x, \eta) + \frac{2}{7} \int^\tau d\eta \, b_1^2(\eta) \left( \delta^{(1)}(x, \eta) \right)^2 \vec{v}^{(1)}(x, \eta).
\]

(357)

We can integrate over time to obtain

\[
\delta_h^{(3)}(x, \tau) = \frac{1}{3!} b_3^3(\tau) \left( \delta^{(1)}(x, \tau) \right)^3 + \delta^{(3)}(x, \tau) - \frac{\tau}{4} b_1^2(\tau) \vec{\nabla} \delta^{(1)}(x, \tau) \cdot \vec{v}^{(2)}(x, \tau) \\
- \frac{\tau}{4} b_1^2(\tau) \delta^{(1)}(x, \tau) \vec{\nabla} \cdot \vec{v}^{(2)}(x, \tau) - \frac{\tau}{4} b_1^2(\tau) \vec{\nabla} \delta^{(2)}(x, \tau) \cdot \vec{v}^{(1)}(x, \tau) \\
- \frac{\tau}{4} b_1^2(\tau) \delta^{(2)}(x, \tau) \vec{\nabla} \cdot \vec{v}^{(1)}(x, \tau) \\
- \left( \frac{\tau}{4} b_2^1(\tau) + \frac{\tau}{21} b_1^3(\tau) \right) \left\{ \vec{\nabla} \left( \delta^{(1)}(x, \tau) \right)^2 \right\} \cdot \vec{v}^{(1)}(x, \tau) + \left( \delta^{(1)}(x, \tau) \right)^2 \vec{\nabla} \cdot \vec{v}^{(1)}(x, \tau) \\
+ \frac{\tau}{14} b_1^2(\tau) \vec{\nabla} \left( \delta^{(1)}(x, \tau) \right)^2 \cdot \vec{v}^{(1)}(x, \tau) + \frac{\tau}{14} b_1^2(\tau) \left( \delta^{(1)}(x, \tau) \right)^2 \vec{\nabla} \cdot \vec{v}^{(1)}(x, \tau).
\]

(358)

Now, the mass conservation equation for the DM at third-order reads

\[
\delta^{(3)}(x, \tau) = \vec{\nabla} \cdot \vec{v}^{(3)}(x, \tau) = - \vec{\nabla} \delta^{(2)}(x, \tau) \cdot \vec{v}^{(1)}(x, \tau) - \vec{\nabla} \delta^{(1)}(x, \tau) \cdot \vec{v}^{(2)}(x, \tau) \\
- \delta^{(2)}(x, \tau) \vec{\nabla} \cdot \vec{v}^{(1)}(x, \tau) - \delta^{(1)}(x, \tau) \vec{\nabla} \cdot \vec{v}^{(2)}(x, \tau).
\]

(359)

Since \(\delta^{(3)}(x, \tau)\) scales like \(a^3\), we can rewrite it as
\[ 3\mathcal{H}\delta^{(3)}(x, \tau) + \vec{\nabla} \cdot \mathbf{v}^{(3)}(x, \tau) = - \left[ \vec{\nabla} \delta^{(2)}(x, \tau) \cdot \mathbf{v}^{(1)}(x, \tau) + \vec{\nabla} \delta^{(1)}(x, \tau) \cdot \mathbf{v}^{(2)}(x, \tau) \right. \\
+ \delta^{(2)}(x, \tau) \vec{\nabla} \cdot \mathbf{v}^{(1)}(x, \tau) + \delta^{(1)}(x, \tau) \vec{\nabla} \cdot \mathbf{v}^{(2)}(x, \tau) \right]. \] (360)

Eq. (358) then becomes

\[ \delta^{(3)}_h(x, \tau) = \frac{1}{3!} b_3^1(\tau) \left( \delta^{(1)}(x, \tau) \right)^3 + \left( \frac{3}{2} b_1^1(\tau) + 1 \right) \delta^{(3)}(x, \tau) + \frac{1}{2 \mathcal{H}} b_1^1(\tau) \theta^{(3)}(x, \tau) \\
- \left( \frac{\tau}{4} b_1^1(\tau) + \frac{\tau}{21} b_1^1(\tau) \right) \left\{ \vec{\nabla} \left[ \left( \delta^{(1)}(x, \tau) \right)^2 \cdot \mathbf{v}^{(1)}(x, \tau) + \left( \delta^{(1)}(x, \tau) \right)^2 \vec{\nabla} \cdot \mathbf{v}^{(1)}(x, \tau) \right] \right\} \\
+ \frac{\tau}{14} b_1^1(\tau) \vec{\nabla} \left( s^{(1)}(x, \tau) \right)^2 \cdot \mathbf{v}^{(1)}(x, \tau) + \frac{\tau}{14} b_1^1(\tau) \left( s^{(1)}(x, \tau) \right)^2 \vec{\nabla} \cdot \mathbf{v}^{(1)}(x, \tau), \] (361)

where \( \theta(x, \tau) = \vec{\nabla} \cdot \mathbf{v}(x, \tau) \) satisfies at any order in perturbation theory the DM momentum equation

\[ \dot{\theta}(x, \tau) + \mathcal{H} \theta(x, \tau) + \partial^i v_i(x, \tau) \partial^j v_j(x, \tau) + \mathbf{v}(x, \tau) \cdot \vec{\nabla} \theta(x, \tau) = -\frac{3}{2} \mathcal{H}^2 \theta(x, \tau). \] (362)

Following Refs. [64, 143], we introduce another non-local coefficient

\[ t_{ij}(x, \tau) = -\frac{1}{\mathcal{H}} \left( \partial_i v_j(x, \tau) - \frac{1}{3} \delta_{ij} \theta(x, \tau) \right) - s_{ij}(x, \tau). \] (363)

It is traceless and vanishes at first-order in perturbation theory as

\[ s_{ij}^{(1)}(x, \tau) = -\frac{1}{\mathcal{H}} \partial_i v_j^{(1)}(x, \tau) - \frac{1}{3} \delta_{ij} \delta^{(1)}(x, \tau), \] (364)

and therefore

\[ t_{ij}^{(1)}(x, \tau) = \frac{1}{3} \delta_{ij} \left( \frac{1}{\mathcal{H}} \theta^{(1)}(x, \tau) + \delta^{(1)}(x, \tau) \right) = 0. \] (365)

This implies that \( t^2(x, \tau) = t_{ij}(x, \tau) t_{ij}(x, \tau) \) is fourth-order and we can neglect it here and henceforth. In particular, up to third-order,

\[ \partial^i v_i(x, \tau) \partial^j v_j(x, \tau) = -2 \mathcal{H}^2 t(x, \tau) \cdot s(x, \tau) + \frac{1}{3} \theta^2(x, \tau) + \mathcal{H}^2 s^2(x, \tau). \] (366)
The equation for $\theta(x, \tau)$ then becomes

$$
\dot{\theta}(x, \tau) + \mathcal{H}\theta(x, \tau) - 2\mathcal{H}^2 t(x, \tau) \cdot s(x, \tau) + \frac{1}{3} \theta^2(x, \tau) + \mathcal{H}^2 s^2(x, \tau)
+ v(x, \tau) \cdot \vec{\nabla}\theta(x, \tau) = -\frac{3}{2} \mathcal{H}^2 \delta(x, \tau).
$$

(367)

Since

$$
\vec{\nabla} \cdot \mathbf{v}^{(2)}(x, \tau) = \frac{\tau^3}{18} \left[ -\partial_x \nabla^2 \varphi(x) \partial_t \varphi(x) - \frac{4}{7} \partial_x \partial_t \varphi(x) \partial_x \partial_t \varphi(x) - \frac{3}{7} (\nabla^2 \varphi(x))^2 \right] 
= \vec{\nabla} \delta^{(1)}(x, \tau) \cdot \mathbf{v}^{(1)}(x, \tau) - \frac{2}{7} \mathcal{H} \left( s^{(1)}(x, \tau) \right)^2 - \frac{4}{21} \left( \delta^{(1)}(x, \tau) \right)^2,
$$

we have

$$
-\frac{1}{\mathcal{H}} \vec{\nabla} \cdot \mathbf{v}^{(2)}(x, \tau) - \delta^{(2)}(x, \tau) = \frac{2}{7} \left( s^{(1)}(x, \tau) \right)^2 - \frac{4}{21} \left( \delta^{(1)}(x, \tau) \right)^2,
$$

(369)

and it is convenient to define two new non-local bias operators [64]

$$
\eta(x, \tau) = -\frac{\theta(x, \tau)}{\mathcal{H}} - \delta(x, \tau)
$$

(370)

and

$$
\psi(x, \tau) = \eta(x, \tau) - \frac{2}{7} s^2(x, \tau) + \frac{4}{21} \delta^2(x, \tau).
$$

(371)

By construction $\eta(x, \tau)$ is a second-order quantity and $\psi(x, \tau)$ is third-order quantity. Eq. (361) becomes then

$$
\delta^{(3)}_h(x, \tau) = \frac{1}{3!} b_2 \left( \delta^{(1)}(x, \tau) \right)^3 + \left( \frac{3}{2} b_1 \right) \delta^{(3)}(x, \tau)
+ \frac{1}{2} b_1 \left( \frac{1}{2} b_2 + \frac{2}{21} b_1 \right) \left[ \vec{\nabla} \left( \delta^{(1)}(x, \tau) \right)^2 \right] \cdot \mathbf{v}^{(1)}(x, \tau)
+ \frac{1}{\mathcal{H}} \left( \frac{1}{2} b_1 \right) \vec{\nabla} \left( s^{(1)}(x, \tau) \right)^2 \cdot \mathbf{v}^{(1)}(x, \tau)
+ \frac{1}{\mathcal{H}} b_1 \left( s^{(1)}(x, \tau) \right)^2 \vec{\nabla} \cdot \mathbf{v}^{(1)}(x, \tau).
$$

(372)

Using Eq. (352) we can rewrite it as
\[ \delta_h^{(3)}(x, \tau) = \frac{1}{3!} b_1^1(\tau) \left( \delta^{(1)}(x, \tau) \right)^3 + (1 + b_1^1(\tau)) \delta^{(3)}(x, \tau) \]
\[ + \frac{1}{2} b_1^1(\tau) \left( -\phi^{(3)}(x, \tau) - \frac{4}{21} \delta^{(1)}(x, \tau) \cdot \delta^{(2)}(x, \tau) + \frac{8}{21} \delta^{(1)}(x, \tau) \delta^{(2)}(x, \tau) \right) \]
\[ - \frac{1}{\mathcal{H}} \left( \frac{1}{2} b_2^1(\tau) + \frac{2}{21} b_1^1(\tau) \right) \delta^{(1)}(x, \tau) \left\{ 2 \nabla \cdot \delta^{(1)}(x, \tau) \cdot \nabla \delta^{(1)}(x, \tau) - \mathcal{H} \left( \delta^{(1)}(x, \tau) \right)^2 \right\} \]
\[ + \frac{1}{\mathcal{H}} \frac{b_1^1(\tau)}{\mathcal{H}} \left\{ \left( \frac{1}{2} b_2^1(\tau) \right) \cdot \nabla \left( \delta^{(1)}(x, \tau) \right)^2 \cdot \nabla \cdot \delta^{(1)}(x, \tau) \right\} + \frac{1}{\mathcal{H}} \frac{b_1^1(\tau)}{\mathcal{H}} \left( \delta^{(1)}(x, \tau) \right)^2 \cdot \nabla \cdot \delta^{(1)}(x, \tau). \tag{373} \]

or
\[ \delta_h^{(3)}(x, \tau) = (1 + b_1^1(\tau)) \delta^{(3)}(x, \tau) + \frac{1}{2} \left( \frac{1}{2} b_2^1(\tau) + \frac{4}{21} b_1^1(\tau) \right) \delta^{(1)}(x, \tau) \delta^{(2)}(x, \tau) \]
\[ + \frac{1}{2} b_1^1(\tau) \left( -\phi^{(3)}(x, \tau) - \frac{4}{21} \delta^{(1)}(x, \tau) \cdot \delta^{(2)}(x, \tau) \right) \]
\[ - \frac{4}{7} \left( \frac{1}{2} b_2^1(\tau) + \frac{29}{54} b_1^1(\tau) \right) \delta^{(1)}(x, \tau) \left( \delta^{(1)}(x, \tau) \right)^2 \cdot \frac{1}{\mathcal{H}} b_1^1(\tau) \nabla \cdot \delta^{(1)}(x, \tau). \tag{374} \]

Let us concentrate on the last term of the expression (374). We can operate a series of manipulations on it
\[ \nabla \cdot \left( \delta^{(1)}(x, \tau) \right)^2 \cdot \nabla \delta^{(1)}(x, \tau) = 2 s_{ij}^{(1)}(x, \tau) \delta^{(1)}_{ijk} s_{ij}^{(1)}(x, \tau) \]
\[ = - \frac{2}{\mathcal{H}} s_{ij}^{(1)}(x, \tau) v_{jk}^{(1)}(x, \tau) \delta^{(1)}_{ij}(x, \tau) \]
\[ = - \frac{2}{\mathcal{H}} s_{ij}^{(1)}(x, \tau) \left[ \delta_{ij} \left( v_{ik}^{(1)}(x, \tau) \right) - \partial_i v_{jk}^{(1)}(x, \tau) - \partial_j v_{ik}^{(1)}(x, \tau) \right] \]
\[ = - \frac{2}{\mathcal{H}} s_{ij}^{(1)}(x, \tau) \left[ \delta_{ij} \left( v_{ik}^{(1)}(x, \tau) \right) - \partial_i v_{jk}^{(1)}(x, \tau) - \partial_j v_{ik}^{(1)}(x, \tau) \right] \]
\[ + 2 \mathcal{H} s_{ij}^{(1)}(x, \tau) \left( s_{ik}^{(1)}(x, \tau) + \frac{1}{3} \delta_{ik}^{(1)}(x, \tau) \right) \left( s_{jk}^{(1)}(x, \tau) + \frac{1}{3} \delta_{jk}^{(1)}(x, \tau) \right) \]
\[ = - \frac{2}{\mathcal{H}} s_{ij}^{(1)}(x, \tau) \left( -\frac{3 \mathcal{H}^2}{2} s_{ij}^{(2)}(x, \tau) \right) \]
\[ + 2 \mathcal{H} s_{ij}^{(1)}(x, \tau) \left( s_{ik}^{(1)}(x, \tau) s_{jk}^{(1)}(x, \tau) + \frac{2}{3} \delta_{ij}^{(1)}(x, \tau) \delta^{(1)}(x, \tau) \right) \]
\[ = 3 \mathcal{H} s_{ij}^{(1)}(x, \tau) s_{ij}^{(2)}(x, \tau) \]
\[ + 5 s_{ij}^{(1)}(x, \tau) \mathcal{H} \left( 1 - \frac{3 \mathcal{H}^2}{2} s_{ij}^{(2)}(x, \tau) \right) \]
\[ = -2 \mathcal{H} s_{ij}^{(1)}(x, \tau) s_{ij}^{(2)}(x, \tau) - 5 \mathcal{H} s_{ij}^{(1)}(x, \tau) s_{ij}^{(2)}(x, \tau) \]
\[ + 2 \mathcal{H} s_{ij}^{(1)}(x, \tau) \left( s_{ij}^{(1)}(x, \tau) \right)^2 \delta^{(1)}(x, \tau). \tag{375} \]
Therefore
\[
\frac{1}{7H} b_1^i(\tau) \overline{\nabla} \left( s_i^{(1)}(x, \tau) \right)^2 \cdot v^{(1)}(x, \tau) = -\frac{2}{7} b_1^i(\tau) s_i^{(1)}(x, \tau) s_i^{(2)}(x, \tau) - \frac{5}{2} b_1^i(\tau) s_i^{(1)}(x, \tau) t_i^{(2)}(x, \tau) \\
+ \frac{2}{7} b_1^i(\tau) \left( s_i^{(1)}(x, \tau) \right)^3 + \frac{4}{21} b_1^i(\tau) \left( s_i^{(1)}(x, \tau) \right)^2 \delta^{(1)}(x, \tau),
\]
(376)
so that
\[
\delta_h^{(3)}(x, \tau) = (1 + b_1^i(\tau)) \delta^{(3)}(x, \tau) \\
+ 2 \left( \frac{1}{2} b_2^i(\tau) + \frac{4}{21} b_1^i(\tau) \right) \delta^{(1)}(x, \tau) \delta^{(2)}(x, \tau) \\
+ \left[ \frac{1}{3!} b_3^i(\tau) - \frac{13}{21} \left( \frac{1}{2} b_2^i(\tau) + \frac{2}{21} b_1^i(\tau) \right) \right] \delta^{(1)}(x, \tau) \delta^{(2)}(x, \tau) \\
- \frac{1}{2} b_1^i(\tau) \left( -\psi(x, \tau) - \frac{8}{7} s^{(1)}(x, \tau) \cdot s^{(2)}(x, \tau) \right) \\
- \frac{4}{7} \left( \frac{1}{2} b_2^i(\tau) + \frac{2}{81} b_1^i(\tau) \right) \delta^{(1)}(x, \tau) \left( s^{(1)}(x, \tau) \right)^2 \\
- \frac{5}{7} b_1^i(\tau) s_i^{(1)}(x, \tau) t_i^{(2)}(x, \tau) + \frac{2}{7} b_1^i(\tau) \left( s_i^{(1)}(x, \tau) \right)^3.
\]
(377)
Combining the results (345), (353), and (377), we finally get
\[
\delta_h(x, \tau) = (1 + b_1^i(\tau)) \delta(x, \tau) + \left( \frac{1}{2} b_2^i(\tau) + \frac{4}{21} b_1^i(\tau) \right) \delta^2(x, \tau) - \frac{2}{7} b_1^i(\tau) s^2(x, \tau) \\
+ \left[ \frac{1}{3!} b_3^i(\tau) - \frac{13}{21} \left( \frac{1}{2} b_2^i(\tau) + \frac{2}{21} b_1^i(\tau) \right) \right] \delta^3(x, \tau) \\
- \frac{1}{2} b_1^i(\tau) \psi - \frac{4}{7} \left( \frac{1}{2} b_2^i(\tau) + \frac{2}{81} b_1^i(\tau) \right) \delta(x, \tau) s^2(x, \tau) \\
- \frac{5}{7} b_1^i(\tau) s(x, \tau) \cdot t(x, \tau) + \frac{2}{7} b_1^i(\tau) s^3(x, \tau).
\]
(378)

### 11.2 The Halo-Mass Correlator

The typical integral we need to calculate is
\[
\mathcal{I}_{ij}(k) = \int \frac{d^3 q}{(2\pi)^3} P_i(q) P_j(|\vec{k} - \vec{q}|) F_2(\vec{q}, \vec{k} - \vec{q}),
\]
(379)
where
\[ F_2(\vec{q}_1, \vec{q}_2) = \frac{5}{7} + \frac{1}{2} \frac{\vec{q}_1 \cdot \vec{q}_2}{q_1 q_2} \left( \frac{q_1}{q_2} + \frac{q_2}{q_1} \right) + \frac{2}{7} \left( \frac{\vec{q}_1 \cdot \vec{q}_2}{q_1 q_2} \right)^2. \]  

(380)

Furthermore

\[ S(\vec{q}_1, \vec{q}_2) = \left( \frac{\vec{q}_1 \cdot \vec{q}_2}{q_1 q_2} \right)^2 \frac{1}{3}. \]  

(381)

We now define the following variables

\[ \mu = \frac{\vec{k} \cdot \vec{q}}{k q}, \quad r = \frac{q}{k} \]  

(382)

such that

\[ \int \frac{d^3 q}{(2\pi)^3} = \frac{k^3}{(2\pi)^2} \int_0^\infty r^2 dr \int_{-1}^1 d\mu, \]

\[ \eta k = |\vec{k} - \vec{q}| = \sqrt{q^2 + k^2 - 2\vec{k} \cdot \vec{q}} = k \sqrt{r^2 + 1 - 2r\mu}, \]

\[ F_2(\vec{q}, \vec{k} - \vec{q}) = \frac{5}{7} + \frac{1}{2} \left[ (\vec{k} \cdot \vec{q} - q^2) \left( \frac{1}{q^2} + \frac{1}{k^2 \eta^2} \right) \right] + \frac{2}{7} \left( \frac{\vec{k} \cdot \vec{q} - q^2}{q^2 k^2 \eta^2} \right)^2 \]

\[ = \frac{5}{7} + \frac{1}{2} \left( \mu r - r^2 \right) \frac{2r^2 + 1 - 2r\mu}{r^2 \eta^2} + \frac{2}{7} \frac{(\mu r - r^2)^2}{r^2 \eta^2} \]

\[ = \frac{3r + 7\mu - 10\mu^2 r}{14r(1 + r^2 - 2r\mu)} \]

\[ S(\vec{q}, \vec{k} - \vec{q}) = \frac{(\mu r - r^2)^2}{r^2 \eta^2} \frac{1}{3}. \]  

(383)

We finally obtain

\[ I_{ij}(k) = \frac{k^3}{56\pi^2} \int_0^\infty dr \int_{-1}^1 d\mu \left( k \sqrt{r^2 + 1 - 2r\mu} \right) \frac{r(3r + 7\mu - 10\mu^2 r)}{(1 + r^2 - 2r\mu)}. \]  

(384)

Notice that, in order to avoid the IR divergence when \( |\vec{k} - \vec{q}| \) goes to zero, by exploring the symmetry of the integral, we can rewrite it as

\[ I_{ij}(k) = \int \frac{d^3 q}{(2\pi)^3} P_i(q) P_j(|\vec{k} - \vec{q}|) F_2(\vec{q}, \vec{k} - \vec{q}) \Theta \left( |\vec{k} - \vec{q}| - q \right) + (i \leftrightarrow j) \]
= \frac{k^3}{56\pi^2} \int_0^\infty dr \int_{-1}^1 d\mu \, P_i(k r) P_j (kr) \left( k \sqrt{r^2 + 1 - 2r\mu} \right) \frac{r (3r + 7\mu - 10\mu^2 r)}{(1 + r^2 - 2r\mu)} \Theta (1 - 2r\mu) + (i \leftrightarrow j),
\tag{385}
\end{equation}

where $\Theta(x)$ is the Heaviside step function. Similarly

$$\int \frac{d^3q}{(2\pi)^3} P_i(q) P_j(|\vec{k} - \vec{q}|) F_2(q, \vec{k} - \vec{q}) S(q, \vec{k} - \vec{q}) = \frac{k^3}{168\pi^2} \int_0^\infty dr \int_{-1}^1 d\mu \, P_i(k r) P_j (kr) \left( k \sqrt{r^2 + 1 - 2r\mu} \right)$$

$$\times \frac{(3r + 7\mu - 10\mu^2 r)(1 - 3\mu^2 + 4\mu r - 2r^2)}{r(1 + r^2 - 2r\mu)^2}.$$

\(\tag{386}\)
12.1 CONVERGENCE OF THE LOGARITHMIC DERIVATIVE

To make sure that our computation of the logarithmic derivative $b_{NG}^{PBS}$ is solid, we check that the measurement is consistent for different box sizes, namely with 1Gpc/h box simulations. In this case we choose a smaller box in order to see if $b_{NG}^{PBS}$ crosses zero and turns negative, in this way changing the sign of the final effect in the scale dependent bias. In Figure 37 we show this check. In Table 8 we show the values $b_{NG}$ for this smaller box simulations for both prescriptions, where the three mass bins have changed accordingly, keeping the condition of each having the same number halos.

Figure 37: Logarithmic derivative of the halo mass function as a function of $\sigma_8$ for both halo finder algorithms for both 2Gpc/h and 1Gpc/h box sets. Here we show results for the first method explained in the text. The solid horizontal black line corresponds to $b_{NG} = 0$ to highlight the zero-crossing of the logarithmic derivative at mass $M \approx 10^{13} M_\odot$. 
In this appendix, we make a brief analysis of the noise correction in our measurement of the halo-halo power spectrum $P_{hh}$. For this purpose, let us define the stochasticity matrix as (see [226])

$$
\sigma_n^2 = \langle (\delta_h - b_{mh}\delta_m)^2 \rangle \tag{387}
$$

$$
= \hat{P}_{hh} - 2b_{mh}\hat{P}_{mh} + b_{mh}^2 \hat{P}_{mm} \tag{388}
$$

where the hat indicates quantities measured from simulations. We plot the stochasticity matrix $\sigma_n^2$ as a function of the wavenumber $k$ in Figure 38. Following [68] we also plot in Figure 39 the relative difference between the halo-matter and the halo-halo linear bias, which we defined in Eqs. (271) and (272)

$$
\frac{\Delta b}{b} - 1 = \frac{b_{hh}}{b_{mh}} - 1 \tag{389}
$$
Figure 39: Relative difference between halo-matter and the halo-halo linear bias as a function of $k$ for the three mass bins and both halo finder algorithms for the 2Gpc/h box sets for the Gaussian simulations.
In this appendix we show the general expression of the halo bispectrum \( B_{\text{hhh}} \) and the halo matter matter bispectrum \( B_{\text{hmh}} \) in the iPT formalism up to one loop. Our expressions hold in Lagrangian space, but the algebra in Eulerian space is analogous. The calculation of \( B_{\text{hhh}} \) has been done in \([281]\), and our results agree with theirs. We first write the bispectrum for Gaussian initial conditions. The halo bispectrum is

\[
B_{\text{hhh}}^G(k_1, k_2, k_3; z_i) = c_f^2(k_1) c_f^2(k_2) c_f^2(k_3) P_L(k_1) P_L(k_2) + 2 \text{ perm.}
\]

\[
+ \frac{1}{2} c_f^2(k_1) P_L(k_1) \int \frac{dq}{(2\pi)^3} c_f^2(k_1, k_2 - q, q) c_f^2(k_2 - q, q) P_L(q) P_L(|k_2 - q|) + 5 \text{ perm.}
\]

\[
+ \int \frac{dq}{(2\pi)^3} c_f^2((q, q + k_1) c_f^2(q + k_1, k_2 - q) c_f^2(q, k_2 - q) P_L(q) P_L(|k_1 + q|)) P_L(|k_2 - q|).
\]

An analogous expression holds for \( B_{\text{hmh}} \) though some of the permutations are no longer symmetric

\[
B_{\text{hmh}}^G(k_1, k_2, k_3; z_i) = \left( \frac{D(z_i)}{D(z_s)} \right)^2 c_f^2(k_2, k_3) P_L(k_2) P_L(k_3)
\]

\[
+ \frac{1}{2} \left( \frac{D(z_i)}{D(z_s)} \right)^3 c_f^2(k_1) c_f^2(k_2, k_3 - q, q) P_L(q) P_L(|k_3 - q|) + (k_2 \leftrightarrow k_3)
\]

\[
+ \frac{1}{2} \left( \frac{D(z_i)}{D(z_s)} \right)^3 P_L(k_2) \int \frac{dq}{(2\pi)^3} F_m^{(2)}(k_2, k_3 - q, q) P_L(q) P_L(|k_3 - q|) + (k_2 \leftrightarrow k_3)
\]

\[
+ \frac{1}{2} \left( \frac{D(z_i)}{D(z_s)} \right)^4 c_f^2(k_1) P_L(k_1) \int \frac{dq}{(2\pi)^3} F_m^{(2)}(k_1, k_2 - q, q) P_L(q) P_L(|k_3 - q|) + (k_2 \leftrightarrow k_3)
\]

\[
+ \frac{1}{2} \left( \frac{D(z_i)}{D(z_s)} \right)^4 \int \frac{dq}{(2\pi)^3} c_f^2(q, q + k_1) F_m^{(2)}(q + k_1, k_2 - q) P_L(q) P_L(k_2 - q)
\]

\[
\times P_L(q) P_L(|k_1 + q|) P_L(|k_2 - q|).
\]
The halo bispectrum in the presence of non-Gaussian initial conditions is modified by

$$
\Delta B_{\text{hmhm}}^{NG}(k_1, k_2, k_3; z_i) = c_1^2(k_1) c_1^2(k_2) c_1^2(k_3) B_L(k_1, k_2, k_3)
$$

$$
+ c_1^2(k_1) \int \frac{d^3q}{(2\pi)^3} c_2^2(q, k_2 - q) c_2^2(k_1 + q, k_2 - q) P_L(|q|, |k_2 - q|) B_L(k_1, q, |k_1 + q|) + 2 \text{ perm.}
$$

$$
+ \frac{1}{2} c_1^2(k_1) c_1^2(k_2) P_L(k_1) \int \frac{d^3q}{(2\pi)^3} c_2^2(k_1, q, k_2 - q) B_L(k_2, q, |k_2 - q|) + (k_1 \leftrightarrow k_2) + 2 \text{ perm.}
$$

$$
+ \frac{1}{2} P_L(k_1) c_1^2(k_1) \left[ c_2^2(k_1, k_2) \int \frac{d^3q}{(2\pi)^3} c_2^2(q, k_2 - q) B_L(k_2, q, |k_2 - q|) + (k_2 \leftrightarrow k_3) \right] + 2 \text{ perm.}
$$

$$
+ \frac{1}{2} c_1^2(k_1) c_1^2(k_2) \int \frac{d^3q}{(2\pi)^3} c_2^2(q, k_3 - q) T_L(k_1, k_2, q, k_3 - q) + 2 \text{ perm.},
$$

(392)

An analogous expression holds for $\Delta B_{\text{hmhm}}$

$$
\Delta B_{\text{hmhm}}^{NG}(k_1, k_2, k_3; z_i) = \left( \frac{D(z_i)}{D(z)} \right)^2 c_1^2(k_1) B_L(k_1, k_2, k_3)
$$

$$
+ \left( \frac{D(z_i)}{D(z)} \right)^4 \frac{1}{2} P_L(k_1) c_1^2(k_1) \left[ F_m^{(2)}(k_1, k_2) \int \frac{d^3q}{(2\pi)^3} F_m^{(2)}(q, k_2 - q) B_L(k_2, q, |k_2 - q|) + (k_2 \leftrightarrow k_3) \right]
$$

$$
+ \left( \frac{D(z_i)}{D(z)} \right)^3 \frac{1}{2} \left[ P_L(k_3) c_1^2(k_2, k_3) \int \frac{d^3q}{(2\pi)^3} F_m^{(2)}(q, k_2 - q) B_L(k_2, q, |k_2 - q|) + (k_2 \leftrightarrow k_3) \right]
$$

$$
+ \left( \frac{D(z_i)}{D(z)} \right)^3 \frac{1}{2} \left[ P_L(k_3) c_1^2(k_2, k_3) \int \frac{d^3q}{(2\pi)^3} c_2^2(q, k_1 - q) B_L(k_1, q, |k_1 - q|) + (k_2 \leftrightarrow k_3) \right]
$$

$$
+ \left( \frac{D(z_i)}{D(z)} \right)^4 \frac{1}{2} c_1^2(k_1) \left[ \int \frac{d^3q}{(2\pi)^3} c_2^2(q, k_1 - q) F_m^{(2)}(k_2 + q, k_1 - q) P_L(|q|, |k_2 + q|) + (k_2 \leftrightarrow k_3) \right]
$$

$$
+ \left( \frac{D(z_i)}{D(z)} \right)^3 \frac{1}{2} \left[ \int \frac{d^3q}{(2\pi)^3} F_m^{(2)}(q, k_2 - q) c_2^2(k_3 + q, k_2 - q) P_L(|k_2 - q|) B_L(k_3, q, |k_3 + q|) + (k_2 \leftrightarrow k_3) \right]
$$

$$
+ \left( \frac{D(z_i)}{D(z)} \right)^4 \frac{1}{2} \left[ c_1^2(k_1) P_L(k_1) \left[ \int \frac{d^3q}{(2\pi)^3} F_m^{(3)}(k_1, q, k_2 - q) B_L(k_2, q, |k_2 - q|) + (k_2 \leftrightarrow k_3) \right] \right]
$$

$$
+ \left( \frac{D(z_i)}{D(z)} \right)^3 \frac{1}{2} \left[ c_1^2(k_1) \left[ \int \frac{d^3q}{(2\pi)^3} F_m^{(2)}(q, k_3 - q) T_L(k_1, k_2, q, k_3 - q) + (k_2 \leftrightarrow k_3) \right] \right]
$$

$$
+ \left( \frac{D(z_i)}{D(z)} \right)^2 \frac{1}{2} \left[ \frac{d^3q}{(2\pi)^3} c_2^2(q, k_1 - q) T_L(k_3, k_2, q, k_1 - q) \right].
$$

(393)
In this section, we present a model for the halo bispectrum in which the NG bias parameters are explicitly derived from a peak-background split. We will first review the derivation of the NG bias parameters following [207], before discussing the bispectrum prescription used in the analysis. Because our model is inspired by [207], we will refer to it as S12i.

### 14.1 Excursion Set Bias from a Peak-Background Split

Ref. [207] took advantage of the peak-background split to derive the PNG bias parameters within the excursion set theory. For the Gaussian case, the biases can be obtained upon considering the modulation of local density fluctuations by a long wavelength perturbation $\delta_l$. The effect of $\delta_l$ can be implemented through a position-dependent offset in the collapse threshold [45, 284, 59].

In the local PNG model, the short mode is given by

$$\Phi_s(q) = \phi_s(q) + K_{q}^{S}[\phi, \phi],$$  \hspace{1cm} (394)

with the coupling term

$$K_{q}^{S}[\phi, \phi] = f_{NL} \int \frac{d^3 p_1}{(2\pi)^3} \frac{d^3 p_2}{(2\pi)^3} \delta_D(q - p_{12}) \left[ \phi_s(p_1)\phi_s(p_2) + 2\phi_l(p_1)\phi_s(p_2) \right], \hspace{1cm} (395)$$

where $\phi_s$ and $\phi_l$ denote the Gaussian short and long modes.

We focus on the coupling $\phi_l\phi_s$, which induces a modulation of the small-scale cumulants, such as the variance and skewness. This was used in Ref. [207] to derive the bias in the presence of PNG.

We begin by summarising the general rules for computing the $n^{th}$ cumulant, $\langle \Phi^n \rangle_{\phi_l}$, in the case of a PNG of the local quadratic type. Here, the subscript $\phi_l$ indicates that the long mode is held fixed in the expectation value. The rules are as follows:

- Write down $n$ points representing the short modes.

The long modes can only arise from the $\phi_l\phi_s$ coupling, and, hence, can be thought of as arising from the short modes.
To each of the short mode, we can attach another short mode through the $\phi_s\phi_s$ coupling in Eq. (395). If no short mode is attached, then it is simply a random Gaussian wavemode. If another short mode is attached, then it contributes a factor of $f_{NL}$.

• Connect the short modes together with the power spectrum, so that the resulting diagram is connected. Note that we do not need to care about the long modes because they are fixed in the ensemble average $\langle \ldots \rangle_{\phi_l}$.

• Finally, we can also choose to attach a long mode to each Gaussian short mode. If so, then the coupling contributes a factor of $f_{NL}$.

From the above diagrammatic rule, we see that the topology of the diagrams is entirely determined by the short modes, while the number of long modes is determined by the number of free short modes at the end. We also note that powers of $f_{NL}$ increase for higher order cumulants because we need more short mode couplings to form a connected diagram.

The leading PNG correction that is modulated by the long mode arises from the variance and reads

$$
\langle \delta_s^2(x) \rangle_{\phi_l} = \int \frac{d^3q_1}{(2\pi)^3} \frac{d^3q_2}{(2\pi)^3} e^{i\mathbf{q}_1 \cdot \mathbf{x}} \alpha(q_1) \alpha(q_2) \langle \Phi_s(q_1) \Phi_s(q_2) \rangle_{\phi_l}
$$

$$
= 4f_{NL} \int \frac{d^3q}{(2\pi)^3} \alpha(q) P_{\phi}(q) \varphi_q(x),
$$

with $\alpha$ and $\varphi$ defined as

$$
\alpha(k,z) = M(k)D(z),
$$

$$
\varphi_q(x) = \int \frac{d^3p}{(2\pi)^3} e^{i\mathbf{p} \cdot \mathbf{x}} \varphi_q(\mathbf{p}) \sum_{j=1}^{\infty} D C_j(x) \frac{\partial \Pi(\delta, \sigma^2_M, \phi_l)}{\partial C_j(x)} \bigg|_{\phi_l=0} \phi_l(k),
$$

The density contrast of the halo fluctuation field can then be expanded in terms of the long mode by means of a functional derivative. In particular, the linear halo bias term is given by

$$
\frac{1}{N} \int_{-\infty}^{\delta_c} d\delta_s \int \frac{d^3k}{(2\pi)^3} \sum_{j=1}^{\infty} \frac{\partial C_j(x)}{\partial \phi_l(k)} \frac{\partial \Pi(\delta, \sigma^2_M, \phi_l)}{\partial C_j(x)} \bigg|_{\phi_l=0} \phi_l(k),
$$

where $C_j$ denotes the $j$th cumulant, $\langle \delta_s \rangle_{\phi_l}$, $\sigma_M$ is the rms variance of matter fluctuation on the halo mass scale $M$, and the normalisation factor $N$ is given by the first crossing distribution in the absence of the long mode $\phi_l$,

$$
N = \partial_M \int_{-\infty}^{\delta_c} d\delta_s \Pi(\delta, \sigma^2_M, 0).
$$
In this formalism, the linear halo bias already depends on all the cumulants. For \( j = 1 \), we recover the standard, Gaussian peak-background split bias \( b_1^{(1)} \). The leading PNG correction arises from the variance and the corresponding NG bias correction can be written

\[
b_1^{(2)}(k) = \frac{\partial \sigma_M^2[I(k), \mathcal{F}]}{\alpha(k) \mathcal{F}},
\]

(400)

where \( \mathcal{F} \) is the first-crossing distribution and the integral \( I \) is defined as

\[
I(k) = 4f_{\text{NL}} \int \frac{d^3p}{(2\pi)^3} \alpha(p) \alpha(|k-p|)K_p^2(k-p, -k)P_\phi(|k-p|).
\]

(401)

In terms of the mass function, we can write \( b_1^{(2)} \) as

\[
b_1^{(2)} = \frac{\partial M}{\partial \ln M} \left[ I_{\text{dn}} - \frac{\partial \sigma_M^2}{\partial \ln M} \right] \alpha(k) \frac{dn}{d\ln M}.
\]

(402)

As emphasised in [207], under the assumption of Markovian random walks and a universal mass function, Eq. (400) can be written in the well-known form [49, 51, 50]

\[
b_1^{(2)}(z^*) = \frac{3f_{\text{NL}} \Omega_m H_0^2 \delta_c k^2 T(k)}{D(z^*)} b_1^{(1)}(z^*)
\]

(403)

upon taking the low-\( k \) limit of \( I \).

The more general prediction Eq. (402) can be computed numerically using the mass function measured in an \( N \)-body simulation. In this work, we have used halos with at least 20 particles. However, the low mass end of this halo mass function significantly underestimates the true mass function even though the halo clustering properties, which are the focus of our analysis, are well reproduced. We have checked that, when the numerical mass function is accurately determined, Eq. (402) often predicts the halo scale-dependent PNG bias more accurately than Eq. (403). Nevertheless, we use Eq. (403) for the computations in the main text, which is appropriate to the \( M \gtrsim M_\ast \) halos we consider.

### 14.2 Analytic Prediction for the Halo-Matter Bispectrum

We write the halo density as

\[
\delta_h(k) = \left[ b_1^{(1)} + b_1^{(2)}(k) \right] \delta_m^R(k) + \frac{b_2}{2} \star \delta_m^R \star \delta_m^R(k),
\]

(404)

where \( b_2 \) includes not only the local bias \( b_2 \), but also nonlocal terms [90]. For the model presented here, \( \delta_m^R \) is the nonlinear matter density field smoothed with a top-hat filter. In
the effective window function of halos was found to be more extended than a top hat.

We can write the tree-level cross bispectrum as

\[
\langle \delta_m(\mathbf{k}_1, z_i)\delta_m(\mathbf{k}_2, z_i)\delta_h(\mathbf{k}_3, z_i) \rangle' = \left( \frac{D(z_i)}{\hat{D}(z_i)} \right)^2 \left[ b_1^{(1)}(z_+) + b_1^{(2)}(k_3, z_+) \right] \tilde{W}_R(k_3) B_L(k_1, k_2, k_3) \\
+ \left( \frac{D(z_i)}{\hat{D}(z_i)} \right)^3 b_1^{(2)}(k_3, z_+) \tilde{W}_R(k_3) B_m^G(k_1, k_2, k_3; z_+) \\
+ B_{mnh}^G(k_1, k_2, k_3; z_i)
\]

where \( B_{mnh}^G \) is the bispectrum in the Gaussian case, which we will assume is given by

the Gaussian simulation, and \( B_m^G \) is the DM bispectrum in the case of Gaussian initial

conditions, which is

\[
B_m^G(k_1, k_2, k_3; z_+) = 2 F_2(k_1, k_2) P_L(k_1) P_L(k_2) + 2 \text{ cyc.}
\]

at tree-level.

The second order bias receives a contribution from PNG, \( b_2^{(2)} \), which can be computed
under the same assumptions that lead to Eq. (403). It is given by [207]

\[
b_2^{(2)}(k_1, k_2) = \frac{4 f_{NL}^2 \hat{f} \delta_i b_1^{(1)}}{a(k_1) a(k_2)}.
\]

However, we found that this contribution is at least two orders of magnitude smaller
than the term proportional to \( b_1^{(1)}(z_+) B_m^G(k_1, k_2, k_3, z_+) \), and we thus shall neglect it here.

Note also that, although Eq. (403) for \( b_1^{(2)} \) is obtained in the low-\( k \) approximation (the
low-\( k \) limit of Eq. (401)), we have checked that the effects are negligible in the final
results even for the matter-squeeze cases in which the wavenumbers are not small.

Finally, as discussed in the main text, for the matter-squeeze case, the loop contribu-
tion due to the third order bias is the dominant one, and reads

\[
B_{b_{\text{loop}}} = \langle \delta_m(\mathbf{k}_1)\delta_m(\mathbf{k}_2) \rangle' \left[ \frac{\delta^{R} \ast \delta^{R} \ast \delta^{R}}{m^3} \right] \tilde{W}_R(k_1) P_L(k_1) \\
= \frac{b_3^{(1)}}{2} \tilde{W}_R(k_1) P_L(k_1) \int \frac{d^3q}{(2\pi)^3} \tilde{W}_R(q) \tilde{W}_R(|k_2 - q|) B_L(-k_2, q, k_2 - q) + (k_1 \leftrightarrow k_2)
\]

Before computing the bispectrum terms, it is instructive to check the scale-independent
bias parameters obtained from different bias schemes. In Fig. 40, we compare three
prescriptions for the PBS Gaussian bias parameters: MW [284], ST [285], and the scale-
independent ESP bias. The MW and ST bias are derived from the Press-Schechter mass
function [46] and the Sheth-Tormen mass function [59], respectively. For the ESP bias,
we use \( b_{00}^{(1)} \equiv b_{n00} \). We have assumed \( z_+ = 0 \). As can be seen, the ST results often fall
between those of MW and ESP. The difference between these prescriptions increases with
the order of the bias parameter. Consequently, $B_{b_3\text{loop}}$ turns out to be fairly sensitive to the exact value of $b_3^{(1)}$. We also note that $b_3^{(1)}$ varies rapidly in the high peak regime, which implies that a small error can lead to large differences in the prediction. This might explain why, for large halo mass, including the $b_3$-loop rarely improves the agreement with the simulations. When the bias parameters are computed using MW and ST prescriptions, the inclusion of the $b_3^{(1)}$-loop often worsens the agreement with the simulations relative to the tree-level-only prediction. The ST results perform marginally better than the MW ones. On the other hand, the ESP results often lead to a better agreement with the numerical data. This is likely due to the fact that our ESP implementation is designed to reproduce the clustering of SO halos, which are the ones analysed here.

Finally we summarise the cross bispectrum model used in the main text. For the halo squeezed case, we adopt the tree-level only bispectrum

$$
\lim_{k_i \to 0} \Delta B_{hmm}^{NG}(k_i, k_s, k_s; z_i) = \left( \frac{D(z_i)}{D(z_*)} \right)^2 \left[ b_1^{(1)}(z_*) + b_1^{(2)}(k_l, z_*) \right] \tilde{W}_R(k_1) B_L(k_s, k_s, k_1) + \left( \frac{D(z_i)}{D(z_*)} \right)^3 b_1^{(2)}(k_l, z_*) \tilde{W}_R(k_1) B_m^G(k_s, k_s, k_l; z_*) , \quad (409)
$$
THE S121 MODEL: THEORETICAL CONSIDERATIONS

while in the matter-squeezed case we include also the $b_3$-loop

$$\lim_{k_l \to 0} \Delta B_{\text{mhm}}^{NG}(k_l, k_s, k_s; z_i) = \left( \frac{D(z_i)}{D(z_s)} \right)^2 \left[ b_1^{(1)}(z_s) + b_1^{(2)}(k_s, z_s) \right] \tilde{W}_R(k_s) B_L(k_s, k_l, k_l) \quad (410)$$

$$+ \left( \frac{D(z_i)}{D(z_s)} \right)^3 b_1^{(2)}(k_s, z_s) \tilde{W}_R(k_s) B_m^{C}(k_s, k_l, k_s; z_s)$$

$$+ \left( \frac{D(z_i)}{D(z_s)} \right)^2 \frac{b_3^{\text{ESP}}(z_s)}{2} \tilde{W}_R(k_s) P_L(k_s) \int \frac{d^3q}{(2\pi)^3} \tilde{W}_R(q) \tilde{W}_R(|k_l - q|)$$

$$\times B_L(-k_l, q, k_l - q). \quad (411)$$

We note that we will use $b_3^{\text{ESP}}$ in the loop calculations.