Università degli Studi di Padova

Dipartimento di Fisica e Astronomia "Galileo Galilei"

Master Degree in Astrophysics and Cosmology

Final Dissertation

Extracting Cosmological Information at Non-Linear Scales using Machine Learning

Thesis Supervisor:
Prof. Michele Liguori

Candidate:
Suprio Dubey

Thesis Co-Supervisors:
Dr. Andrea Ravenni
Dr. Gabriel Jung

Academic Year 2023/2024

To my mother

Abstract

Estimation of precise cosmological parameters has been a longstanding challenge in modern cosmology. Forthcoming cosmological surveys such as DESI, Euclid, SPHEREX, and the Roman space telescope will allow us to observe a large fraction of the sky and provide a very detailed map of the large-scale structure of the universe. The vast amount of data collected from these surveys will make small scales accessible with high precision. The small, non-linear scales will help shed more light on unresolved topics in cosmology, such as explaining cosmological tensions, examining the early universe, studying neutrinos, and so on.

In this work, we address such questions by applying a set of different estimators to the Quijote-png halo catalogues and Quijote-massive neutrinos halo catalogues at non-linear cosmological scales, up to $k_{\rm max} = 0.5\,h\,{\rm Mpc}^{-1}$ to find the best approach to constrain Primordial non-Gaussianity (PNG) which gives insight into the early universe and total neutrino mass M_ν. We use various summary statistics including the power spectrum, bispectrum, halo mass function, marked power spectrum, and marked modal bispectrum to train deep neural networks (NN) and perform likelihood-free inference of cosmological, PNG, and $M\nu$ parameters. We also look into a thorough comparison of summary statistics to determine their optimal combination in terms of PNG and $M\nu$ sensitivity.

.

Acknowledgement

I would like to extend my sincere gratitude to my supervisor, Professor Michele Liguori, and my co-supervisors, Dr. Andrea Ravenni and Dr. Gabriel Jung, for giving me the opportunity to work on this project. I am deeply thankful for their insightful discussions and constant support, which have significantly broadened my knowledge. I also appreciate the valuable discussions with my examiner, Professor Alessandro Renzi.

I would also like to express my heartfelt thanks to my friends Anant, Chelsea, Siddhesh, Ankita, Helena, and Mayank, who stood by me through difficult times, kept me motivated, and supported me in every way possible. Additionally, I am grateful to my flatmates and other friends who made my time in Padova an enjoyable experience. Lastly, I extend my deepest gratitude to my mother, whose unwavering support and encouragement have inspired me to pursue my dreams.

Contents

1 Introduction **1**

2 Introduction to Cosmology **3**
2.1 The Center of the Universe . . . 3
2.2 Friedmann Equation . . . 4
2.2.1 Space-time Curvature . . . 4
2.2.2 Stress-energy Tensor . . . 6
2.3 The Cosmological Constant . . . 7
2.4 Cosmic Distances . . . 8
2.5 Cosmic Accounting: The Universe's Energy Budget . . . 10
2.6 Thermal History of the Universe . . . 11
2.7 The Early Universe . . . 12
2.7.1 The Horizon Problem . . . 13
2.7.2 The Flatness Problem . . . 14
2.7.3 Dynamics of Inflation . . . 14
2.7.4 Inflation-Induced Cosmological Perturbations . . . 19
2.7.5 The Power Spectrum . . . 27
2.7.6 From Quantum to Cosmological Fluctuations . . . 29
2.8 Primordial Non-Gaussianity . . . 31
2.9 Neutrino Mass . . . 33

3 Perturbation Theory and the Formation of Large-Scale Structures **36**
3.1 Euler Perturbation Theory . . . 37
3.1.1 The Vlasov Equation . . . 37
3.1.2 Eulerian Dynamics . . . 38
3.1.3 Linear Solution . . . 39
3.1.4 Eulerian Non-Linear Perturbation Theory . . . 41
3.1.5 General Solutions in Einstein-de Sitter Cosmology . . . 41
3.2 Lagrangian Perturbation theory . . . 44
3.2.1 Lagrangian Dynamics . . . 44
3.2.2 Linear Solution . . . 45
3.2.3 Second-order solution . . . 46
3.3 N-body Simulation . . . 47

4 Statistical Analysis of Random Cosmic Field **49**
4.1 Random fields . . . 49
4.2 Two-point correlation function and power spectrum . . . 51
4.2.1 Perturbative Expansion: up to one-loop . . . 52
4.3 Three-point correlation function and bispectrum . . . 53
4.3.1 One-loop matter bispectrum . . . 55
4.3.2 Marked Statistics . . . 56

5 Methodology and Data Analysis **59**
5.1 Quijote Simulation . . . 60
5.2 Summary Statistics Estimators . . . 60
5.2.1 Power Spectrum . . . 60
5.2.2 Bispectrum . . . 61
5.2.3 Halo Mass Function . . . 61
5.2.4 Marked Statistics . . . 62
5.3 Method . . . 62
5.3.1 Likelihood Free Inference . . . 62
5.3.2 Moment Network . . . 63
5.3.3 Neural Network Model . . . 64
5.3.4 Evaluators . . . 67

6 Results and Discussion **69**
6.1 f_{NL}^{equil} LH . . . 69
6.2 nwLH . . . 73

7 Conclusion **77**

Appendices **78**

A Window Function **79**

B Additional Figures **82**
B.1 f_{NL}^{equil} LH summary statistics combination plots . . . 82
B.2 nwLH summary statistics combination plots . . . 91

Chapter 1

Introduction

The standard ΛCDM model is both a boon and a bane for our current understanding of the universe. It quite well explains the formation of large-scale structures, the state of the early universe, and the cosmic abundance of different types of matter and energy. On the other hand, despite the standard ΛCDM model's continued success in observations, there is growing interest in expanding cosmology beyond it because of some profound theoretical challenges, like the enigmatic physical origin of dark matter and dark energy and the apparent discrepancies between the observed and predicted clustering properties of cold dark matter (CDM) on small scales, which leave a large unanswered gap.

The standard model involves an early period where the density perturbations are very close to the Gaussian field with all pertinent information contained in the power spectrum, which is the Fourier transform of the two-point correlation function. The evolution of the universe to low redshifts resulted in these fluctuations becoming non-Gaussian on small scales owing to the nonlinear gravitational evolution. This leads to information leakage from the two-point function to higher-order statistics.

Neutrinos with finite mass suppress the small-scale power spectrum below the neutrino free-streaming scale. The suppression is proportional to the sum of the neutrino masses. This allows us to measure the total neutrino mass (M_ν), determining which is one of the major challenges in cosmology as traditional power spectrum analysis is no longer applicable in the non-linear regime.

Another challenge is extracting primordial non-Gaussianity (PNG) information, especially in the case of equilateral shape, which we will focus on in this thesis from the non-Gaussianity produced by non-linear gravitational dynamics at low redshifts due to its comparatively weak signal. PNG is the deviation of the primordial perturbation distribution from Gaussianity. Improving the current constraint on the parameter $f_{\mathrm{NL}^{equil}}$ that characterizes the size of Equilateral PNG can help us develop a unique insight into the physics of the early universe.

The analysis of the large-scale structure of the Universe (LSS) at small, non-linear scales holds significant promise for such cosmological problems. Current and forthcoming cosmological surveys such as DESI, Euclid, SPHEREX, and the Roman space telescope will allow us to observe a large fraction of the sky and provide very detailed maps of the large-scale structure of the universe. The vast amounts of data collected from these surveys, i.e., all the small-scale modes, cannot be analyzed using fully analytical models hence leveraging into higher-statistical-significance measurements becomes imperative to extract meaningful insights.

In this work, we perform a thorough analysis of the dark matter halo field using various summary statistics and their combinations in Quijote N-body simulations [Villaescusa-Navarro et al., 2020a]. We explore simulations with PNG initial conditions (Quijote-PNG) [Coulton et al., 2023] and simulations with massive neutrinos at non-linear cosmological scales, up to $k_{\max} = 0.5\,h\,\mathrm{Mpc}^{-1}$, aiming to identify the most effective methods for constraining f_{NL}^{equil} and total neutrino mass M_ν respectively.

We use neural networks (NN) for likelihood-free inference using the moment network method developed by [Jeffrey and Wandelt, 2020] which is advantageous as it allows exploration across a broad range of parameter values without assuming a specific fiducial cosmology or relying on Gaussian likelihood assumptions. It directly maps the summary statistics to the final parameters, bypassing the need to evaluate covariance or assume a particular form for the likelihood function.

We compare different combinations of summary statistics such as the power spectrum, modal

bispectrum, halo mass function, marked power spectrum and marked modal bispectrum over a wide range of parameters in our neural network to determine their optimal combination in terms of $f_{NL}^{\rm equil}$ and M_ν parameters sensitivity and help break degeneracies of them and the cosmological parameters. This method represents some of the initial efforts in the direction of future applications to PNG and M_ν parameter inference on real data from galaxy surveys which will require the need to use more realistic galaxy mocks.

The thesis is organized as follows. In Chapter 2, we start with a basic overview of cosmology and then move on to discuss the early universe and the inflationary scenario. This chapter also includes discussions on the dynamics during that epoch, as well as brief introductions to primordial non-Gaussianity and massive neutrinos. Moving to Chapter 3, we discuss different cosmological perturbation theories, covering both linear and non-linear scales, alongside the insights from N-body simulations crucial for understanding density fluctuations, particularly at small, strongly non-linear scales. Chapter 4 explores the theory of summary statistics used in our analysis, examining their dependence on PNG conditions and massive neutrinos. In Chapter 5, we detail the dataset and methodology we employed, featuring the use of NN for analysis. Finally, Chapter 6 presents the findings derived from our study.

Chapter 2

Introduction to Cosmology

2.1 The Center of the Universe

One of the foundations of modern cosmology is the Cosmological Principle. If we view it over a large scale, the universe is independent of the position i.e. *homogeneous* and independent of direction i.e. *isotropic* to a first approximation. The spacetime geometry that is consistent with the universe being homogeneous and isotropic is given by the Friedmann-Robertson-Walker Metric(FRW) metric.

$$ds^2 = -c^2dt^2 + a(t)^2\left(\frac{dr^2}{1-kr^2} + r^2d\theta^2 + r^2\sin^2\theta d\phi^2\right), \tag{2.1}$$

where $k = 0$, $k = +1$ and $k = -1$ for flat, positively curved and negatively curved space-like 3-hypersurfaces, respectively. Whereas, $a(t)$ is the scale factor that incorporates the expansion of the universe.[1] For ease of notation, we will restrict most of our discussion to the case $k = 0$.
We can write the FRW metric in the flat space i.e. taking the form of the Minkowski metric with the scale factor.

$$ds^2 = -\ dt^2 + a^2(t)\ d\sigma^2 \tag{2.2}$$

where

$$d\sigma^2 = \gamma_{ij}\ dx^i\ dx^j \tag{2.3}$$

with

$$\gamma_{ij} = \delta_{ij} + k\frac{x_ix_j}{1-k(x_kx^k)} \tag{2.4}$$

We can notice that by using the polar coordinates we can get to equation Eq. (2.1). The polar coordinate used in the equation (2.1) refers to the observer comoving with the expansion of the universe. We get the physical distance by multiplying the comoving distance with the scale factor.
We can redefine the radial coordinate as

$$d\chi = \frac{dr}{\sqrt{1-kr^2}} \tag{2.5}$$

such that,

$$ds^2 = -dt^2 + a(t)^2\left[d\chi^2 + f(\chi)^2(d\theta^2 + \sin^2\theta d\phi^2)\right] \tag{2.6}$$

[1]The line element in (2.1) has a scaling symmetry i.e $a \to \lambda a$, $r \to \frac{r}{\lambda}$, $k \to \lambda^2 k$. We can use this freedom to set the scale factor to unity today i.e $a_0 = 1$

where,

$$f(\chi) = \begin{cases} sinh\chi, & k = -1 \\ \chi, & k = 0 \\ sin\chi, & k = +1 \end{cases} \tag{2.7}$$

This form of metric is particularly convenient for studying the propagation of light. Light travels along the null geodesic $ds^2 = 0$ so to make the propagation of light in FRW similar to the Minkowski space we introduce conformal time,

$$d\tau = \frac{dt}{a(t)}, \tag{2.8}$$

Substituting it in (2.6) we get

$$ds^2 = -a(\tau)^2 \left[d\tau^2 - \left(d\chi^2 + f(\chi)^2 (d\theta^2 + sin^2\theta d\phi^2) \right) \right] \tag{2.9}$$

2.2 Friedmann Equation

The dynamics of the way the universe evolves are determined by the Einstein equation.

$$G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = 8\pi G T_{\mu\nu} \tag{2.10}$$

This relates to the Einstein tensor $G_{\mu\nu}$ which is a measure of the "space-time curvature" of the FRW universe to the stress-energy tensor where $T_{\mu\nu}$ describes the "matter content" of the universe.

2.2.1 Space-time Curvature

Let us compute the Einstein tensor on the l.h.s. of the Einstein equation $G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R$
The Ricci tensor is given by :

$$R_{\mu\nu} = \Gamma^{\alpha}_{\mu\nu,\alpha} - \Gamma^{\alpha}_{\mu\alpha,\nu} + \Gamma^{\beta}_{\mu\nu}\Gamma^{\alpha}_{\alpha\beta} - \Gamma^{\beta}_{\mu\alpha}\Gamma^{\alpha}_{\nu\beta} \tag{2.11}$$

Let us see the Christoffel symbols for the FRW metric:

$$\Gamma^0_{\mu\nu} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & \frac{\dot{a}a}{1-kr^2} & 0 & 0 \\ 0 & 0 & r^2 a\dot{a} & 0 \\ 0 & 0 & 0 & r^2 a\dot{a}\sin^2\theta \end{bmatrix} \tag{2.12a}$$

$$\Gamma^1_{\mu\nu} = \begin{bmatrix} 0 & \dot{a}/a & 0 & 0 \\ \dot{a}/a & \frac{kr}{(1-kr^2)} & 0 & 0 \\ 0 & 0 & (kr^2-1)r & 0 \\ 0 & 0 & 0 & (kr^2-1)r\sin^2\theta \end{bmatrix} \tag{2.12b}$$

$$\Gamma^2_{\mu\nu} = \begin{bmatrix} 0 & 0 & \dot{a}/a & 0 \\ 0 & 0 & 1/r & 0 \\ \dot{a}/a & 1/r & 0 & 0 \\ 0 & 0 & 0 & -\sin\theta\cos\theta \end{bmatrix} \tag{2.12c}$$

$$\Gamma^3_{\mu\nu} = \begin{bmatrix} 0 & 0 & 0 & \dot{a}/a \\ 0 & 0 & 0 & 1/r \\ 0 & 0 & 0 & \cos\theta/\sin\theta \\ \dot{a}/a & 1/r & \cos\theta/\sin\theta & 0 \end{bmatrix} \tag{2.12d}$$

Here we used simplifications that the FRW metric is diagonal, and it does not depend on φ.

The components of the Ricci tensor are:

$$R_{00} = -3\partial_t(\frac{\dot{a}}{a}) - 3(\frac{\dot{a}}{a})^2 \tag{2.13a}$$
$$= -3\left(\frac{\ddot{a}}{a} - (\frac{\dot{a}}{a})^2 + (\frac{\dot{a}}{a})^2\right) \tag{2.13b}$$
$$= -3\frac{\ddot{a}}{a} \tag{2.13c}$$

$$\begin{aligned} R_{11} = {} & \partial_t(\frac{\dot{a}a}{1-kr^2}) + \partial_r(\frac{kr}{1-kr^2}) - \partial_r(\frac{kr}{1-kr^2}) - 2\partial_r(\frac{1}{r}) \\ & + \frac{\dot{a}a}{1-kr^2}3\frac{\dot{a}}{a} + \frac{kr}{1-kr^2}(\frac{kr}{1-kr^2} + \frac{2}{r}) \\ & - 2\frac{\dot{a}}{a}\frac{\dot{a}a}{1-kr^2} - (\frac{kr}{1-kr^2})^2 - 2(\frac{1}{r})^2 \end{aligned} \tag{2.14a}$$
$$= \frac{\ddot{a}a + \dot{a}^2}{1-kr^2} + 3\frac{\dot{a}^2}{1-kr^2} + 2\frac{k}{1-kr^2} - 2\frac{\dot{a}^2}{1-kr^2} \tag{2.14b}$$
$$= \frac{\ddot{a}a + 2\dot{a}^2 + 2k}{1-kr^2} \tag{2.14c}$$

$$\begin{aligned} R_{22} = {} & r^2\partial_t(a\dot{a}) + \partial_r((kr^2-1)r) - \partial_\theta(\frac{\cos\theta}{\sin\theta}) \\ & + 3\Gamma^t_{\theta\theta}\Gamma^\theta_{t\theta} + \Gamma^r_{\theta\theta}(\Gamma^r_{rr} + 2\Gamma^\theta_{r\theta}) - 2(\Gamma^t_{\theta\theta}\Gamma^\theta_{t\theta} + \Gamma^r_{\theta\theta}\Gamma^\theta_{\theta r}) - \frac{\cos^2\theta}{\sin^2\theta} \end{aligned} \tag{2.15a}$$
$$= r^2(\ddot{a}a + \dot{a}^2) + 3kr^2 - 1 + \frac{1}{\sin^2\theta} + r^2\dot{a}^2 - kr^2 - \frac{\cos^2\theta}{\sin^2\theta} \tag{2.15b}$$
$$= r^2(\ddot{a}a + 2\dot{a}^2 + 2k) \tag{2.15c}$$

$$R_{33} = \partial_\alpha\Gamma^\alpha_{\varphi\varphi} - \partial_\varphi\Gamma^\alpha_{\alpha\varphi} + \Gamma^\alpha_{\varphi\varphi}\Gamma^\beta_{\alpha\beta} - \Gamma^\beta_{\varphi\alpha}\Gamma^\alpha_{\varphi\beta} \tag{2.16a}$$
$$= r^2\sin^2\theta(\ddot{a}a + 2\dot{a}^2 + 2k) \tag{2.16b}$$

The Ricci scalar then comes out to be

$$\begin{aligned} R = g^{\mu\nu}R_{\mu\nu} = {} & 3\frac{\ddot{a}}{a} + \frac{1-kr^2}{a^2}\frac{\ddot{a}a + 2\dot{a}^2 + 2k}{1-kr^2} \\ & + \frac{1}{a^2r^2}r^2(\ddot{a}a + 2\dot{a}^2 + 2k) + \frac{1}{a^2r^2\sin^2\theta}r^2\sin^2\theta(\ddot{a}a + 2\dot{a}^2 + 2k) \end{aligned} \tag{2.17a}$$
$$= 3\frac{\ddot{a}}{a} + 3\frac{\ddot{a}a + 2\dot{a}^2 + 2k}{a^2} \tag{2.17b}$$
$$= 6\left[\frac{\ddot{a}}{a} + (\frac{\dot{a}}{a}) + \frac{k}{a^2}\right]. \tag{2.17c}$$

We find that the non-zero components of the Einstein tensor $G^\mu_\nu = g^{\mu\lambda}G_{\lambda\nu}$

$$G^0_0 = 3\left[\left(\frac{\dot{a}}{a}^2\right) + \frac{k}{a^2}\right] \tag{2.18a}$$

$$G^i_j = \left[2\frac{\ddot{a}}{a} + \left(\frac{\dot{a}}{a}^2\right) + \frac{k}{a^2}\right]\delta^i_j \tag{2.18b}$$

*

2.2.2 Stress-energy Tensor

For the universe to be isotropic and homogeneous it is forced to be a perfect fluid. Perfect fluids have a stress-energy tensor like :

$$T^{\mu\nu} = (\rho + P)u^{\mu}u^{\nu} + Pg^{\mu\nu} \tag{2.19}$$

where u^{μ} is the 4-velocity of the fluid element. It is diagonal in the *comoving frame*, in which $u^{\mu} = (1, \vec{0})$.
If we take the covariant divergence of the Einstein tensor $G_{\mu\nu}$ we get zero; so the stress-energy tensor must also have $\nabla_{\mu}T^{\mu\nu} = 0$. One key emphasis here is that the covariant derivative of the stress-energy tensor is not a conservation equation unlike in special relativity where the equation $\partial_{\mu}T^{\mu\nu}$ describes a conservation equation, a local one.
In GR we get a conserved quantity if the metric doesn't depend on some coordinate ($\equiv$ something constant). We denote Killing Vectors, a vector oriented in the direction of symmetry. But in cosmology, we do not have symmetry with respect to time translation, so there is no time-like Killing vector ξ_{μ} such that $\xi_{\nu}\nabla_{\mu}T^{\mu\nu}$ represents the conservation of energy.
This equation, $\nabla_{\mu}T^{\mu\nu}$ follows from the fact that our fluid follows its equations of motion. Let us explore the meaning of these equations. If, in the equation $\nabla_{\mu}T^{\mu}_{0} = 0$, we find

$$\nabla_{\mu}T^{\mu}_{0} = \partial_{\mu}T^{\mu}_{0} + \Gamma^{\mu}_{\mu\lambda}T^{\lambda}_{0} - \Gamma^{\lambda}_{\mu 0}T^{\mu}_{\lambda} = 0 \tag{2.20a}$$

the term T^{i}_{0} vanishes by isotropy. So we get,

$$\partial_{t}\rho + \Gamma^{\mu}_{\mu 0}\rho - \Gamma^{\lambda}_{\mu 0}T^{\mu}_{\lambda} = 0 \tag{2.20b}$$

From (2.12b),(2.12c),(2.12d) we see that $\Gamma^{\lambda}_{\mu 0}$ vanishes unless λ and μ are spatial indices equal to each other, in which case it is $\frac{\dot{a}}{a}$.The above equation therefore reads :

$$\dot{\rho} + 3\frac{\dot{a}}{a}(\rho + P) = 0 \tag{2.20c}$$

To find the other two Friedmann equations we compare the time-time equation and the space-space equation of the Einstein equations to the stress-energy tensor i.e comparing (2.19) to (2.18a) and (2.18b) respectively To relate the sources of matter to the evolution of the scale factor in the FRW metric (2.9),

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}(\rho + 3P) \tag{2.21a}$$

$$\left(\frac{\dot{a}}{a}\right)^{2} = \frac{8\pi G}{3}\rho - \frac{k}{a^{2}} \tag{2.21b}$$

The space-space equation is not a dynamical equation, since it contains no second-time derivatives: it is a *constraint* on the evolution of the system. However, the three Friedmann equations are not independent: the time-time one can be found from (2.21a) and (2.20c).
We define $\left(\frac{\dot{a}}{a}\right) \equiv H$, the Hubble parameter which is the measure of the expansion rate of the universe, measured in $\mathrm{Km.s^{-1}.Mpc^{-1}}$.
The above equations (2.20c),(2.21a),(2.21b) describe how the universe expands and lay the groundwork for all further discussion of cosmology. We can solve the Friedmann equations by an equation of state that relates ρ to P.

$$w = \frac{P}{\rho}. \tag{2.22}$$

Plugging it in the continuity equation (2.20c) we get:

$$\frac{\log \rho}{\log a} = -3(1+w) \implies \rho \propto a^{-3(1+w)} \tag{2.23}$$

$$\rho(a) \propto \begin{cases} a^{-4}, & \text{radiation: } w = \frac{1}{3} \\ a^{-3}, & \text{matter: } w = 0 \\ \text{constant}, & \text{dark energy: } w = -1 \end{cases} \tag{2.24}$$

It makes intuitive sense that the energy density of relativistic particles decreases as the universe expands, not only because the average number of particles per unit volume decreases ($V \propto a^3$), as it does for non-relativistic matter, but also because of the relativistic Doppler shift, which increases the negative power of the scaling factor by one. In contrast, dark energy has a constant energy density. Assuming the universe is composed of three different kinds of fluids, radiation, matter, and dark energy, then each has successively dominated the energy budget of the universe and shaped its history at various epochs.
Now, combining (2.23) with the first Friedmann equation (2.12b) we get the time dependence of the scale factor.

$$a(t) \propto \begin{cases} t^{\frac{2}{3(1+w)}}, & \text{if,} w \neq -1 \\ e^{Ht}, & \text{if,} w = -1 \end{cases} \tag{2.25}$$

We can observe that for $w = 0$, which corresponds to a Universe dominated by non-relativistic matter, we get $a(t) \propto t^{2/3}$. Now, for $w = 1/3$, which corresponds to relativistic particles (radiation), we get $a(t) \propto t^{1/2}$. The last and very relevant case is $w = -1$, corresponding to a fluid exerting negative pressure, which leads to the second solution in Eq. (2.25), i.e. an exponential expansion describes the Dark Energy.

2.3 The Cosmological Constant

When Einstein developed his theory of general relativity, the prevailing belief was that the Universe was static. However, this contradicted the Friedmann equations, which showed that a Universe evolving according to them could not be static unless the acceleration $\vec{a}$ was zero. namely

$$\rho = -3P \tag{2.26}$$

Since a fluid with such property did not seem to be physically reasonable, Einstein modified Eq. (2.10) by introducing the cosmological constant term Λ to counterbalance the gravitational attraction

$$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = 8\pi G T_{\mu\nu} - \Lambda g_{\mu\nu}, \tag{2.27}$$

in such a way that it does not change the covariant character of the equations. It can be shown that for an appropriate choice of Λ, one indeed obtains a static cosmological model. In order to recover a form similar to the Eq. (2.10), we rewrite the stress-energy tensor in a more compact way

$$\bar{T}_{\mu\nu} = T_{\mu\nu} - \frac{\Lambda}{8\pi G} g_{\mu\nu} \tag{2.28}$$

$$= (\bar{P} + \bar{\rho}) u_\mu u_\nu + \bar{P} g_{\mu\nu}, \tag{2.29}$$

so that

$$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = 8\pi G \bar{T}_{\mu\nu} \tag{2.30}$$

In Eq. (2.29), the effective pressure $\bar{P}$ and the effective density $\bar{\rho}$ are related to the corresponding quantities for a perfect fluid by

$$\bar{P} = P - \frac{\Lambda}{8\pi G}, \quad \bar{\rho} = \rho + \frac{\Lambda}{8\pi G} \tag{2.31}$$

Although after the discovery of the expansion of the Universe in the late 1920s, there was no need for a term that made the Universe static, the cosmological constant remained a subject of great interest and it is today the simplest possible explanation for the observed accelerated expansion of the Universe. The standard model of cosmology that incorporates the effects of the cosmological constant is the ΛCDM model.

One cosmological model that involves the cosmological constant is the de Sitter Universe. In this model, the Universe is dominated by a positive cosmological constant, which suppresses all other matter contributions, making the Universe empty ($\rho = 0$),($P = 0$) and flat ($k = 0$). Under these conditions, the effective pressure and density are related to the cosmological constant through equation (2.31) we get,

$$\bar{P} = -\bar{\rho} = -\frac{\Lambda}{8\pi G}, \tag{2.32}$$

which, if replaced in Eq. (2.21b) gives

$$\frac{\dot{a}^2}{a^2} = H^2 = \frac{\Lambda}{3}, \tag{2.33}$$

corresponding to a Hubble parameter constant in time.The solution to this equation is an exponential expansion, where the scale factor increases as a function of time

$$a \propto \exp\left(\sqrt{\frac{\Lambda}{3}}t\right), \tag{2.34}$$

This exponential expansion means that test particles move away from each other due to the repulsive gravitational effect of the positive cosmological constant.

2.4 Cosmic Distances

The Universe is observed to be expanding, and the scale factor $a(t)$ is increasing with time. This expansion also affects the photons that were emitted a long time ago far away and reach us today, which become redshifted due to the stretching of space. Therefore, it is convenient to define a quantity that describes the distance - and time - of a patch of the Universe we are observing based on this physical effect. For that, we define the redshift z as

$$z + 1 = \frac{a(t0)}{a(t)} \tag{2.35}$$

where t is the time when the signal was emitted and t_0 is today. The scale factor is defined up to a multiplicative constant, and it's usually assumed $a(t_0) = 1$, so $z + 1 = a(t)^{-1}$. As we know, the causal connection of points in space-time is determined by whether they are inside or outside each other's lightcone. Lightcones are the solutions to the equation

$$ds^2 = 0, \tag{2.36}$$

where ds^2 is the space-time interval between two events. In other words, the lightcone of an event consists of all the events that could potentially receive a signal from the first event, assuming the signal travels at or below the speed of light.[2] Any event inside the past lightcone of another event can be causally influenced by it, while any event outside the future lightcone cannot be influenced by

[2]In this section we use c = 1

it. With ds^2 a suitable metric, in our case the FRW metric Eq. (2.6), assuming a homogeneous and isotropic Universe and taking $d\Omega = 0$,the comoving particle horizon is then,

$$r_p(t) \equiv \chi(t) = \int_0^t \frac{dt'}{a(t')} = \int_0^a d\ln a \left(\frac{1}{aH}\right) = \tau \tag{2.37}$$

where in the last equality we exploited the relation we found earlier for the conformal time Eq. (2.8)[3].

The comoving horizon is therefore the logarithmic integral of the comoving Hubble radius $\frac{1}{aH}$, which we will define below. The related physical particle horizon is

$$d_p(t) = a(t)r_p(t) = a(t)\int_0^t \frac{dt'}{a(t')} \tag{2.38}$$

due to the scale factor in the denominator and the fact that $a(t) \to 0$ as t goes to 0, the particle horizon could become infinite, meaning that an observer would be in causal connection with the whole Universe. Using Eq. (2.25), we can find an approximate expression for $d_p(t)$ for $w > -1/3$, or else we can notice that the integral diverges. Using the Hubble parameter in the FRW model which is

$$H = \frac{2}{3(1+w)t} \tag{2.39}$$

With $w = 0$, a spatially flat matter-dominated universe, $H = 2/3t$ and $a \propto t^{\frac{2}{3}}$. With $w = 1/3$, a spatially flat radiation-dominated universe, we have $H = 1/2t$ and $a \propto t^{\frac{2}{3}}$. we can write Eq. (2.38) as

$$d_p(t) \simeq \frac{3(1+w)}{1+3w}t = \frac{2}{1+3w}\frac{1}{H} \tag{2.40}$$

We take into note that $w > -1/3$ implies $\ddot{a} < 0$, a decelerating expansion which means we have a finite particle horizon only in the case of a primordial universe characterized by a decelerating universe. Now, let us look into the Hubble time and the comoving Hubble radii.

$$t_H = \frac{a}{\dot{a}} = \frac{1}{H} \tag{2.41}$$

$$r_H = \frac{R_H}{a} = \frac{1}{\dot{a}} = \frac{1}{aH}^4 \tag{2.42}$$

The Hubble radius, which is given by $(aH)^{-1}$, represents the distance that particles can travel during one expansion time. It is a measure of whether particles can communicate with each other within a given time frame, based on their comoving separation λ. In contrast, the particle horizon represents the maximum distance from which particles can reach us within the age of the Universe. Thus, while both the Hubble radius and particle horizon are related to the causal connection of particles, they differ conceptually in that the former measures whether particles can communicate within a given time frame, while the latter measures the maximum distance from which particles can reach us.

1. if λ >particle horizon, then the particles could never have communicated.
2. if $\lambda > (aH)^{-1}$ then the particles can't communicate now.

[3]The main benefit of working with conformal time: light rays correspond to straight lines at 45° angles in the $\chi - \tau$ coordinates. If instead, we had used physical time t, then we would find the light cones for curved space-times would be curved.

2.5 Cosmic Accounting: The Universe's Energy Budget

As we saw previously in solving the Friedmann equations using the equation of state w, the universe is composed of various ingredients, which prompts us to consider their individual contributions to the cosmic energy budget. In this section, we explore the cosmic accounting of the universe's energy budget and investigate the different components that make up the cosmic inventory. We follow [Baumann, 2022] for the discussion. Let us look back into Eq. (2.21b) and re-write it using the Hubble parameter H, we get

$$H^2 = \frac{8\pi G}{3}\rho - \frac{k}{a^2} \tag{2.43}$$

We will use the subscript '0' to denote the quantities evaluated today, at $t = t_0$. Considering a flat universe($k = 0$) the density required for it corresponds to the critical density today denoted by ρ_c

$$\rho_{crit,0} = \frac{3H_0^2}{8\pi G} \tag{2.44}$$

We use critical density to define dimensionless density parameters

$$\Omega_{I,0} \equiv \frac{\rho_{I,0}}{\rho_{crit,0}} \tag{2.45}$$

We can write the Friedmann equation Eq. (2.43) as

$$H^2(a) = H_0^2\left[\Omega_{r,0}\left(\frac{a_0}{a}\right)^4 + \Omega_{m,0}\left(\frac{a_0}{a}\right)^3 + \Omega_{k,0}\left(\frac{a_0}{a}\right)^2 + \Omega_{\Lambda,0}\right] \tag{2.46}$$

where $I = (r, m, k, \Lambda)$ stands for radiation, matter, curvature, and vacuum energy contribution respectively and we defined a *curvature* density parameter $\Omega_{k,0} \equiv \frac{-k}{(a_0 H_0)^2}$. From now on we will drop the subscript "0" and just write it as Ω_r which means the radiation density today. So Eq. (2.46) becomes

$$\frac{H^2}{H_0^2} = [\Omega_r a^4 + \Omega_m a^3 + \Omega_k a^2 + \Omega_\Lambda] \tag{2.47}$$

At the current time, the scale factor is $a(t_0) = 1$ and $H = H_0$.Substituting the values in (2.47) we get

$$\Omega_{total} = \Omega_r + \Omega_m + \Omega_k + \Omega_\Lambda = 1 \tag{2.48}$$

so, we can say that the density parameter represents the energy content of the universe. Observations show that the universe is filled with radiation(r), matter (m), and dark energy(Λ).

$$|\Omega_k| \leq 0.01, \Omega_r = 9.4 \times 10^{-5}, \Omega_m = 0.32, \Omega_\Lambda = 0.68 \tag{2.49}$$

As we can see, the curvature only accounts for 1% of the total energy budget. Curvature's effects were inconsequential earlier(the curvature contribution is only 10^{-2}).The matter density parameter is further split into Baryons(b) and DM(c).
The predictions of primordial nucleosynthesis, which postulates that the density of protons and neutrons in the early Universe affects the efficiency with which fusion occurs, provide the best current constraints on the baryon density of the universe. Deuterium and other elements found in primordial gas clouds have been studied, and the results suggest that the density parameter of baryonic matter is $\Omega_b = 0.05$. The majority of the matter density parameter is contained in $non-baryonic$ matter called DM $\Omega_c = 0.27$.This 27% of the universe's energy budget is one of the most enigmatic substances in the universe, which has been an active topic of research.
Of the radiation contribution majority of it comes in the form of the photons from the cosmic microwave background (CMB) however, experiments such as the *Planck* satellite have found clear evidence for relativistic particles in addition to the known Photons and is compatible with the expected cosmic neutrino contribution[Akrami et al., 2019]. This will be explored in more detail in Chapter 2.9.

2.6 Thermal History of the Universe

In this section, we outline the thermal history of the standard hot big bang universe as the temperature of the plasma decreased from approximately $10^{11}\,\mathrm{K}$ to $10^{8}\,\mathrm{K}$, following [Weinberg, 2008]. During this period, two significant events occurred: neutrinos decoupled from the plasma, and shortly afterwards, electrons and positrons annihilated, leading to the heating of photons relative to neutrinos. Our goal is to understand how these events influenced the evolution of the energy densities of photons and neutrinos.

For massless particles described by the Fermi-Dirac or Bose-Einstein distributions, the energy density is given by

$$\rho(T) = \begin{cases} g\frac{\pi^2}{30}T^4 & \text{Boson} \\ \frac{7}{8}g\frac{\pi^2}{30}T^4 & \text{Fermion} \end{cases} \tag{2.50}$$

where g counts the number of distinct spin states. The entropy density for massless particles is given by

$$s(T) = \frac{4\rho(T)}{3T} \tag{2.51}$$

It is convenient to define a quantity $g_\star$ which counts the spin states for all particles and antiparticles, with an additional factor $\frac{7}{8}$ for fermions. With this definition, the total energy density and entropy density of the Universe during radiation domination are given by

$$\rho(T) = g_\star \frac{\pi^2}{30}T^4 \tag{2.52}$$

$$s(T) = \frac{4}{3}g_\star \frac{\pi^2}{30}T^3 \tag{2.53}$$

In an expanding universe, the first law of thermodynamics implies that for particles in equilibrium, the comoving entropy density is conserved:

$$a^3 s(T) = \text{const} \tag{2.54}$$

A straightforward consequence of this conservation is that for radiation in free expansion, the temperature evolves inversely with the scale factor:

$$T \propto \frac{1}{a} \tag{2.55}$$

Let's now apply this understanding to the early universe.

At a temperature of $10^{11}\,\mathrm{K}(T \sim 10\,\mathrm{MeV})$, the universe was filled with photons, electrons, positrons, and neutrinos (and antineutrinos) of three species, all in thermal equilibrium with negligible chemical potential.[5]. As the temperature of the plasma dropped below approximately $10^{10}\,\mathrm{K}$ (around 1 second after the end of inflation), the rate of collisions between neutrinos and electrons/positrons could no longer keep up with the universe's expansion rate. Consequently, neutrinos fell out of equilibrium and began free expansion. Meanwhile, electrons and positrons remained in equilibrium with photons, and their number densities decreased with the falling temperature, effectively vanishing by the time the temperature reached $T \sim 10\,\mathrm{keV}$.

For simplicity, we will assume that neutrinos decoupled instantaneously before electron-positron annihilation and will later discuss how more detailed calculations might modify the results. Non-zero neutrino masses can be neglected here as long as $m_\nu \lesssim 1\,\mathrm{keV}$, a condition supported by current observational bounds [Aghanim et al., 2020].

From this point forward, we will differentiate the temperature of neutrinos, T_ν, from that of photons, T_γ. Before neutrino decoupling, frequent interactions maintained thermal equilibrium

[5]There was also a much smaller density of baryons and dark matter, but they are not significant for the discussion above

between neutrinos and photons, ensuring a common temperature. After the universe became transparent to neutrinos, they retained their relativistic Fermi-Dirac distribution, with a temperature decreasing inversely with the scale factor. In contrast, photons were heated by the annihilation of electrons and positrons. The conservation of comoving entropy allows us to compute the relative temperatures at later times. After neutrino decoupling, but before electron-positron annihilation, the thermal plasma contained two spin states of photons, plus two spin states each of electrons and positrons, which means that during this period,

$$g_\star^{\text{before}} = 2 + \frac{7}{8}(2+2) = \frac{11}{2} \tag{2.56}$$

After electron-positron annihilation, only the two spin states of photons remained, and so

$$g_\star^{\text{after}} = 2 \tag{2.57}$$

Since $T_\nu \propto a^{-1}$ during this period, the comoving entropy conservation can be expressed as

$$\frac{g_\star^{\text{before}} T_{\gamma,\text{before}}^3}{T_{\nu,\text{before}}^3} = \frac{g_\star^{\text{after}} T_{\gamma,\text{after}}^3}{T_{\nu,\text{after}}^3} \tag{2.58}$$

Using the fact that $T_{\gamma,\text{before}} = T_{\nu,\text{before}}$, we find

$$\frac{T_{\gamma,\text{after}}}{T_{\nu,\text{after}}} = \left(\frac{11}{4}\right)^{1/3} \tag{2.59}$$

We see that in the instantaneous neutrino decoupling limit, the annihilation of electrons and positrons raised the temperature of photons relative to that of neutrinos by a factor of $(11/4)^{1/3} \simeq 1.40$.

After electron-positron annihilation, assuming three species of light neutrinos and antineutrinos, each with one spin state, the radiation density of the Universe is

$$\rho_r = \frac{\pi^2}{30}\left[2T_\gamma^4 + 6\frac{7}{8}T_\nu^4\right] = \frac{\pi^2}{15}\left[1 + 3\frac{7}{8}\left(\frac{4}{11}\right)^{4/3}\right]T_\gamma^4 \tag{2.60}$$

It is conventional to define a quantity N_{eff} which gives the radiation energy density in terms of the effective number of neutrino species as

$$\rho_r = \frac{\pi^2 k_B^4}{15\hbar^3 c^3}\left[1 + \frac{7}{8}\left(\frac{4}{11}\right)^{4/3} N_{\text{eff}}\right]T_\gamma^4 \tag{2.61}$$

where we have restored k_B, c and $\hbar$ for completeness. In the instantaneous neutrino decoupling approximation described above, we found $N_{\text{eff}} = 3$. In the real Universe, however, the decoupling of neutrinos is not instantaneous, and the residual coupling of neutrinos at the time of electron-positron annihilation increases N_{eff} by a small amount in the Standard Model.

Unlike photon decoupling at temperature $k_B T \sim 0.2\text{eV}$, active neutrino decoupling at $T \sim 10\text{MeV} - 0.1\text{MeV}$ takes place over many tens of Hubble times, with the result that we expect distortions in the relic neutrino energy spectra relative to the thermal relativistic Fermi-Dirac distribution. Standard Model Boltzmann neutrino transport calculations show that these distortions change N_{eff} at the percent level, with the current best estimate predicting $N_{\text{eff}} = 3.046$ [Mangano et al., 2005]. In section 2.9 we will see how neutrinos leave a detectable imprint on cosmological observations.

2.7 The Early Universe

The shortcoming of FRW cosmology is that it is unable to explain why the universe we observe is homogeneous and isotropic on larger scales without a finely tuned set of initial conditions, and how the initial seed perturbations for structure formation were generated.

In this chapter, we delve into two of the primary problems of the Hot Big Bang model: the horizon problem and the flatness problem. We explore how inflation, an early period of accelerated expansion,

can drive the primordial Universe towards homogeneity and isotropy, even if it begins in a more generic initial state, and also how quantum fluctuations that arise during the inflationary period can give rise to the formation of primordial black holes.

2.7.1 The Horizon Problem

Previously we introduced Hubble radii ($r_H = (aH)^{-1}$) Eq.(2.42) and comoving particle horizon τ (2.37) as an integral of the comoving Hubble radii. We could observe that for a universe dominated by the fluid equation of state, $w > -1/3$ the Hubble radius and the comoving particle horizon grow monotonically with time which implies that the comoving scales entering the horizon today have been far outside the horizon at CMB decoupling. However, we find the CMB to be extremely homogeneous. In other words, at the time of the last scattering, the universe was expected to be homogeneous only on small scales since wider scales would not have been causally connected. However, the observation of the nearly homogeneous CMB suggests that the universe was extremely uniform even on larger scales that should have been independent of each other which was surprising.
To overcome the horizon problem, it's important to consider the causal contact between particles. If a region, denoted by λ, has a typical size (constant in comoving scales) that is smaller than $(a_I H_I)^{-1}$, then the particles within that region were in causal contact. However, if λ becomes larger than $(a_I H_I)^{-1}$ after a sufficient period of inflation, then these particles can no longer communicate. Therefore, before crossing the Hubble radius and becoming causally disconnected, these particles had the opportunity to communicate with each other and reach similar conditions. This implies that everything within the Hubble sphere at the beginning of inflation, which is $(a_I H_I)^{-1}$, was in causal contact.
We can hence observe that the horizon problem can be resolved if the comoving Hubble radius at the beginning of inflation, $(a_I H_I)^{-1}$, exceeds the radius of the observable universe, $(a_0 H_0)^{-1}$. In this scenario, the entire observable universe would have been contained within the comoving Hubble radius at the onset of inflation.
The duration of the inflation to solve the horizon problem is given by the means of e-folds which is defined as,

$$\mathcal{N} = ln\left(\frac{a_E}{a_I}\right) \tag{2.62}$$

from the requirement $(a_I H_I)^{-1} > (a_0 H_0)^{-1}$ we get $\mathcal{N} \approx 60$

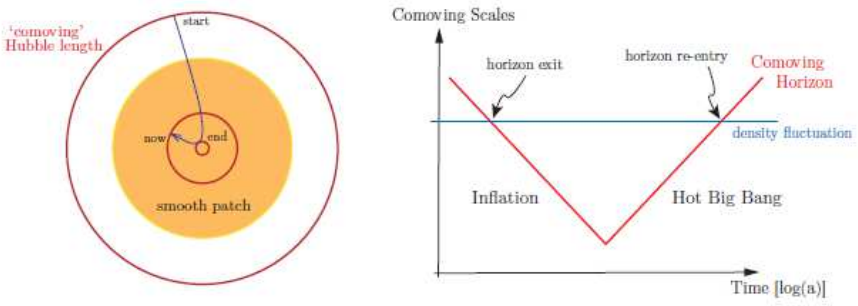

Figure 2.1: *Left*: Development of the inflationary universe's $(aH)^{-1}$ comoving Hubble radius. Inflation causes the co-moving Hubble sphere to contract and then expand. Hence, inflation serves as a tool to "zoom in" on a smooth sub-horizon patch. *Right*: Solution of the horizon problem. All scales that are relevant to cosmological observations today were larger than the Hubble radius until $a \sim 10^{-5}$. These scales were, however, smaller than the Hubble radius at sufficiently early eras and were consequently causally related. Similarly, the scales of cosmological interest came back within the Hubble radius in relatively recent times.[Baumann, 2012]

2.7.2 The Flatness Problem

We saw in section 2.5 that the current measurements of the density parameters are compatible with a flat universe $\Omega_{total} \simeq 1$.But the question arises how does Ω evolve with time? We can write the Friedmann equation Eq. (2.43) as

$$\Omega(t) - 1 = \frac{k}{H(t)^2 a(t)^2} \propto \frac{k}{a^{2-3(1+w)}} = k a^{3w+1} \tag{2.63}$$

We observe that for both radiation ($w = 1/3$) and matter ($w = 0$), we see that if $\Omega - 1 \neq 0$ at a certain early time then it will grow as a^2 or a^4 respectively and then the geometry of the present Universe is expected to be curved. So the near-flatness, we observe today requires the value of $\Omega(t)$ to be one. By manipulating the Friedmann equation Eq. (2.43) we find

$$\Omega^{-1}(z) - 1 = (\Omega_0^{-1} - 1)\left(\frac{T_0}{T(z)}\right)^2 . \tag{2.64}$$

Let us extend this to the Planck epoch because the density at a time $t_P \approx 10^{-43} s$ must have been very close to the critical density. The temperature at that time was $T_P = 10^{32} K$. T_0 is the temperature today.

If we compute $T_{\rm Pl}/T_0$ we get approximately 10^{32}.

This means that

$$\Omega^{-1}(z_{\rm Pl}) - 1 \approx (\Omega_0^{-1} - 1) 10^{-64} . \tag{2.65}$$

.As a result, the Universe should have needed to be extraordinarily finely tuned in order to imitate the flatness we observe today.

Using the Hubble radius, we can write to examine how the inflationary phase responds to this issue.

$$\Omega(t) - 1 = \frac{k}{H(t)^2 a(t)^2} = k r_H^2 \tag{2.66}$$

In standard cosmology, the comoving Hubble radius grows with time, therefore one expects the quantity $|\Omega(t) - 1|$ to grow with time and the geometry of the present Universe to be curved. But, if the universe undergoes a phase where the Hubble radius shrinks, this would bring $\Omega(t)$ closer to 1. It can be shown that the universe must expand by 60 e-foldings to achieve the flatness we observe today, as in the case of the horizon problem.

2.7.3 Dynamics of Inflation

In the previous section, we saw that for inflation to take place we need a phase of accelerated expansion ($\ddot{a} > 0$) where the Hubble radius(r_H) shrinks.
We can see from the Friedmann Eq. (2.21a) that for an accelerated expansion

$$\ddot{a} = -\frac{4\pi G}{3}(\rho + 3P) > 0 \Rightarrow P < \frac{-\rho}{3} \tag{2.67}$$

which we previously encountered in the de-Sitter model where we found $a(t) \propto exp(\sqrt{\frac{\Lambda}{3}} t)$ and the expansion is driven by cosmological constant Λ, so we can have a calculated guess that the inflation can't be driven by matter or radiation. In the current interpretation, Λ is linked to the quantum fluctuations of the vacuum. The vacuum expectation value of the stress-energy tensor gives the energy generated by these fluctuations.Using Eq. (2.19) and Eq. (2.32) we can write

$$\langle T_{\mu\nu} \rangle = -\langle P_\Lambda \rangle g_{\mu\nu} = \frac{\Lambda}{8\pi G} g_{\mu\nu} \tag{2.68}$$

As a result, in this instance, the stress-energy tensor's vacuum expectation value functions as the cosmological constant that promotes the expansion. This provides an indication of what to anticipate from a theory describing the inflationary mechanism.

Scalar(Inflaton) Field

To satisfy Eq. (2.67) let us introduce a minimally coupled(i.e not coupled with gravity or any other field) scalar field φ with a suitable potential $V(\varphi)$ which is a self-interaction of the field. We can write its Lagrangian as

$$\mathscr{L}_\varphi = -\frac{1}{2}g^{\mu\nu}\partial_\mu\varphi\partial_\nu\varphi - V(\varphi) \tag{2.69}$$

The action is in the form of

$$S = S_{EH} + S_\varphi + S_{\text{matter}} \tag{2.70}$$

where S_{EH} is the Einstein-Hilbert action for the metric, S_φ is the action for the field φ, while "matter" encompasses all the other fields but we can ignore it thanks to the *no-hair cosmic theorem.*

$$S = \frac{1}{16\pi G}\int \mathrm{d}^4x\,\sqrt{-g}R + \int \mathrm{d}^4x\,\sqrt{g}\mathscr{L}\varphi[\varphi, g\mu\nu],. \tag{2.71}$$

We are using the invariant volume element $\mathrm{d}^4x\,\sqrt{-g}$ representing the physical 4-volume regardless of the coordinates. Now, we can find the equation of motion for the inflaton field by the Klein-Gordon equation by varying the action with respect to φ.

$$\Box\varphi = \frac{\partial V}{\partial\varphi}\,. \tag{2.72}$$

where $\Box$ is the covariant D'Alembert operator.

$$\Box\varphi = \frac{1}{\sqrt{-g}}(g^{\mu\nu}\sqrt{g}\varphi_{,\mu})_{,\nu}\,. \tag{2.73}$$

and in a flat FRW metric Eq. (2.1) $\sqrt{-g} = a^3$ the evolution of φ becomes

$$\Box\varphi = \frac{1}{a^3}\partial_0(g^{00}a^3\partial_0\varphi) + \frac{1}{a^3}\partial_i(g^{ii}a^3\partial_i\varphi) = \partial_\varphi V \tag{2.74}$$

$$-\ddot{\varphi} - \dot{\varphi}3\frac{\dot{a}}{a} + \frac{\nabla^2}{a^2}\varphi = \partial_\varphi V \tag{2.75}$$

$$\ddot{\varphi} + 3H\dot{\varphi} - \frac{\nabla^2\varphi}{a^2} = -\partial_\varphi V\,. \tag{2.76}$$

where $3H\ddot{\varphi}$ appears as a friction term that is represented as a scalar field rolling down its potential suffering friction due to the expansion of the universe. If we consider the homogeneous background field, it will be constant in space, so the above equation evolves as

$$\ddot{\varphi} + 3H\dot{\varphi} = -\partial_\varphi V\,. \tag{2.77}$$

The stress-energy tensor associated with the scalar field can be defined by

$$T^{(\varphi)}_{\mu\nu} = -\frac{2}{\sqrt{-g}}\frac{\delta S_\varphi}{\delta g^{\mu\nu}} = -2\frac{\partial\mathscr{L}_\varphi}{\partial g^{\mu\nu}} + \frac{2}{\sqrt{-g}}\mathscr{L}_\varphi\frac{\partial\sqrt{-g}}{\partial g^{\mu\nu}} = -2\frac{\partial\mathscr{L}_\varphi}{\partial g^{\mu\nu}} + \mathscr{L}_\varphi g_{\mu\nu} \tag{2.78}$$

$$= \partial_\mu\varphi\partial_\nu\varphi + g_{\mu\nu}\left[-\frac{1}{2}g^{\alpha\beta}\partial_\alpha\varphi\partial_\beta\varphi - V(\varphi)\right]. \tag{2.79}$$

If we compare it with Eq. (2.19) we find that $\varphi(t)$ behaves like a perfect fluid with

$$P = -\frac{1}{2}g^{\alpha\beta}\partial_\alpha\varphi\partial_\beta\varphi - V(\varphi) \tag{2.80}$$

$$\rho = -\frac{1}{2}g^{\alpha\beta}\partial_\alpha\varphi\partial_\beta\varphi + V(\varphi) \tag{2.81}$$

$$u_\mu = \frac{\partial_\mu\varphi}{|\partial\varphi|} \tag{2.82}$$

$$|\partial\varphi| = \sqrt{-g^{\alpha\beta}\partial_\alpha\varphi\partial_\beta\varphi}\,. \tag{2.83}$$

We start by considering $\varphi(t,x)$ and split it as

$$\varphi(\mathbf{x},t) = \varphi(t) + \delta\varphi(\mathbf{x},t) \tag{2.84}$$

where $\varphi(t)$ is the classical field that is the expectation value of the inflaton field($\langle\varphi(t,x)\rangle = \varphi(t)$) and $\delta\varphi(\mathbf{x},t)$ represents the quantum fluctuations around $\varphi(t)$. Now,

$$\left|\frac{\delta\varphi(\mathbf{x},t)}{\varphi(t)}\right| \ll 1 \tag{2.85}$$

as quantum fluctuation is negligible in comparison to the classical value. These fluctuations are what generated the density fluctuation which creates anisotropies in the CMB photons. Let us drop the "t" and indicate the value of the classic inflaton field by φ. On explicitly computing the energy-momentum tensor of the classical background φ we find,

$$T^0_0 = -\left(\frac{1}{2}\dot\varphi(t)^2 + V(\varphi)\right) = -\rho_\varphi \tag{2.86}$$

$$T^i_j = \left(\frac{1}{2}\dot\varphi^2(t) - V(\varphi)\right)\delta^i_j = P_\varphi\delta^i_j\,. \tag{2.87}$$

This is the perfect fluid energy-momentum tensor. Therefore, if

$$V(\varphi) \gg \dot\varphi^2, \tag{2.88}$$

we get $P_\varphi \simeq -\rho_\varphi \implies w_\varphi \simeq -1$ i.e the quasi-de Sitter expansion.
So, we see that if the potential energy is greater than the kinetic energy this scalar field gives inflation. For better intuition let us simplify the system and assume the vacuum expectation value of the inflaton to be constant i.e $\langle\varphi(t,x)\rangle = \bar\varphi$ then the stress-energy tensor becomes

$$\langle T_{\mu\nu}\rangle = -g_{\mu\nu}V(\varphi), \tag{2.89}$$

Comparing it to Eq. (2.68) we see that the potential of the inflaton field $V(\varphi)$ represents the vacuum energy associated with φ which drives the acceleration.

Slow Roll Conditions

Let us try to understand under what conditions the scalar field may initiate inflation. The equation of motion for the homogeneous scalar field is given in Eq. (2.77). To satisfy the condition given in Eq. (2.88) the scalar field should slowly roll down its potential. The easiest way to satisfy the slow roll condition is to require that there exist regions of field-configuration space where the potential is sufficiently flat. Considering the potential to be flat the acceleration of the field should be negligible i.e

$$\ddot\varphi \ll 3H\dot\varphi \tag{2.90}$$

We can also say that at a sufficiently late time the scalar field is driven by friction term. Therefore,

$$3H\dot\varphi \approx -\partial_\varphi V \tag{2.91}$$

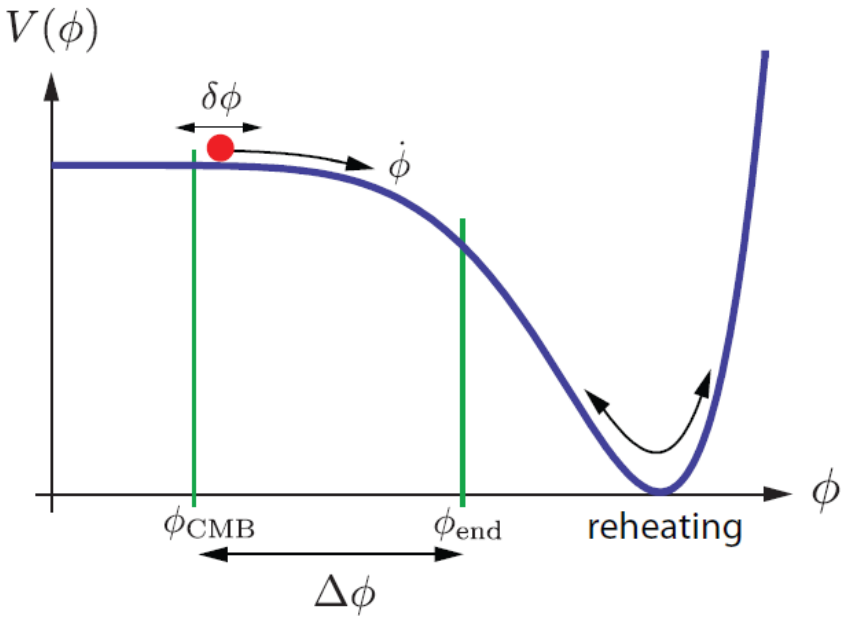

Figure 2.2: A possible shape of the potential for Slow-Roll Inflation [Baumann, 2012]

We expect that V and all of its derivatives change very slowly with φ. This means that in this equation we have $\partial_\varphi V \approx const$, as well as $H \approx const$: this is the same equation that is obeyed by a particle under a constant force and friction: it will then reach the asymptotic "terminal velocity" and move with a constant $\dot{\varphi}_0$.

So, the slow-roll condition is,

$$\ddot{\varphi} \ll (3H\dot{\varphi}), (-\partial_\varphi V) \tag{2.92}$$

Let us now combine it with the Friedmann equation (2.43),

$$H^2 = \frac{8\pi G}{3}(\rho_\varphi + \rho_m + \rho_r) - \frac{k}{a^2}\,. \tag{2.93}$$

The matter and radiation densities scale like a^{-3} for ρ_m, a^{-4} for ρ_r; in this early phase the scalar field dominates the dynamics, so the equation will simplify to

$$H^2 \approx \frac{8\pi G}{3}V(\varphi)\,. \tag{2.94}$$

so, for a slow roll case, the Hubble parameter is nearly a constant and the scale factor is given by $a(t) \propto exp(Ht)$.
We saw the slow roll conditions which are necessary for successful inflation. Next, we will parameterize it.

These are the parameters we need to quantify in order to determine how closely the potential matches our expectations. It is given by ϵ and η. Let us start with the first parameter.
We define ϵ as,

$$\epsilon = -\frac{\dot{H}}{H^2}. \tag{2.95}$$

As we saw previously for inflation to take place the Hubble radius shrinks. So let us write this in terms of the ϵ parameter.

$$\frac{\mathrm{d}(aH)^{-1}}{\mathrm{d}t} = \frac{-\dot{a}H + a\dot{H}}{(aH)^2} = -\frac{1}{a}\left(1 - \left(\frac{-\dot{H}}{H^2}\right)\right) = -\frac{1}{a}(1-\epsilon) \tag{2.96}$$

so we see that if $\dot{r_H} < 0$ implies that $\epsilon \ll 1$.
Using Eq. (2.91) and Eq. (2.94) we can write ϵ as,

$$\epsilon = -\frac{\dot{H}}{H^2} = +4\pi G\frac{\dot{\varphi}^2}{H^2} \approx \frac{3}{2}\frac{\dot{\varphi}^2}{V} = \frac{1}{16\pi G}\left(\frac{\partial_\varphi V}{V}\right)^2, \tag{2.97}$$

so the condition $\epsilon \ll 1$ can also be written as

$$\frac{(\partial_\varphi V)^2}{16\pi G V^2} \ll 1 \tag{2.98}$$

$$\frac{(\partial_\varphi V)^2}{V} \ll 16\pi G V = \frac{2}{3}H^2 \tag{2.99}$$

$$\frac{1}{V}\left(\frac{\partial V}{\partial \varphi}\right)^2 \ll H^2\,. \tag{2.100}$$

on using Eq. (2.95) we can see that it is exactly the slow roll condition in Eq. (2.88), So, we see that ϵ gives a bound on the first derivative of the potential and corresponds to the conditions of the potential being flat and the kinetic energy being small compared to the potential.
The second derivative of the potential is controlled by the parameter η which is defined as,

$$\eta = -\frac{\ddot{\varphi}}{H\dot{\varphi}}\,, \tag{2.101}$$

and we can also define

$$\eta = \frac{1}{3}\frac{\partial_\varphi^2 V}{H^2} = \frac{1}{8\pi G}\frac{\partial_\varphi^2 V}{V}\,. \tag{2.102}$$

We can see that $\eta \ll 1$ is equivalent to

$$\frac{\partial^2 V}{\partial \varphi^2} \ll H^2 \tag{2.103}$$

We can show that these three parameters are related $\delta = \eta - \epsilon$.
We start from Eq. (2.91) and differentiate it with respect to the time we get

$$\ddot{\varphi} \approx -\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial_\varphi V}{3H}\right) \tag{2.104}$$

$$= -\frac{1}{3H}\partial_\varphi^2 V\dot{\varphi} - \frac{\partial_\varphi V}{3}\underbrace{\left(-\frac{\dot{H}}{H^2}\right)}_{\epsilon} \tag{2.105}$$

$$= -\dot{\varphi}H\frac{\partial_\varphi^2 V}{3H^2} - \frac{\partial_\varphi V}{3}\epsilon \tag{2.106}$$

$$= -\dot{\varphi}H\eta - \frac{\partial_\varphi V}{3}\epsilon \tag{2.107}$$

$$-\frac{\ddot{\varphi}}{H\ddot{\varphi}} = \eta - \epsilon \tag{2.108}$$

$$\delta = \eta - \epsilon,. \tag{2.109}$$

which is the desired result. We see that $\delta \ll 1$ corresponds to the condition $\ddot{\varphi} \ll -\partial_\varphi V$, which is required in order to neglect the acceleration term in the Klein-Gordon equation. So, the condition $\delta \ll 1$ ensures that we move towards an attractor solution in the friction-dominated regime i.e., it will then reach the asymptotic "terminal velocity" and move with a constant $\dot{\varphi}$. Also, for inflation to solve the horizon and the flatness problem we need a phase of accelerating expansion that lasts sufficiently long. For this to happen, we need $\epsilon \sim const$, since $\epsilon \sim \dot{\varphi}$ while $\delta \sim \ddot{\varphi}$, so requiring $\delta \ll 1$ also ensures this. .
In summary, the slow-roll approximation described in Eq. (2.88) and Eq. (2.92) implies the inflationary potential's flatness under conditions Eq. (2.97) and Eq. (2.102).

2.7.4 Inflation-Induced Cosmological Perturbations

Inflationary cosmology relies on understanding the evolution of quantum fluctuations of the inflaton field $\delta\varphi(\mathbf{x}, t)$, which give rise to primordial energy density perturbations that persist after inflation and form the basis of the large-scale structures observed in the Universe. These fluctuations arise on extremely small scales within the comoving Hubble radius during inflation, but the rapid expansion of space during this epoch stretches them out to cosmological scales. As the Hubble radius begins to increase faster than the scale factor after inflation ends, these fluctuations eventually re-enter the Hubble radius during the radiation or matter-dominated eras. The fluctuations that re-enter around 60 e-foldings before reheating have physical wavelengths that can be observed through various methods, such as the analysis of CMB anisotropies. The resulting inflationary spectrum provides a unique and distinct signature of inflation that can help us understand the origins of structure in the Universe.

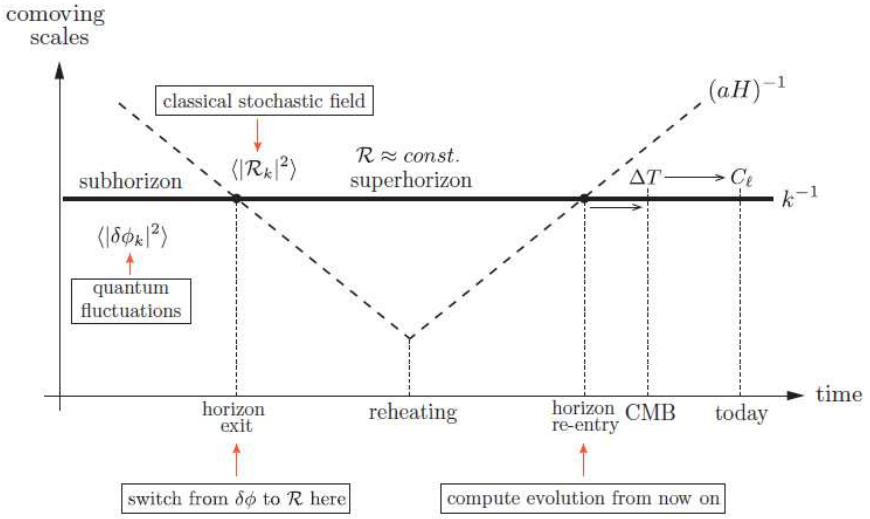

Figure 2.3: Creation and evolution of perturbations in the inflationary universe. On subhorizon scales, quantum mechanics produces fluctuations. During inflation, comoving scales, k^{-1}, stay constant in the comoving Hubble radius, but $(aH)^-1$ diminishes and the perturbations leave the horizon, where they remain frozen until horizon re-entry at later times. The fluctuations change into anisotropies in the CMB and disturbances in the LSS after horizon re-entry. Taken from [Baumann, 2012]

Dynamics of Quantum Fluctuations: Qualitative Analysis

Let us now describe qualitatively how the quantum fluctuations of a generic scalar field evolve during an inflationary stage. We saw the dynamics of the scalar field obey the Klein-Gordon Eq. (2.76) which we Taylor expand up to linear order around the background value for both φ Eq. (2.84) and for the potential $V(\varphi) \approx V_0 + \delta\varphi \left(\frac{\partial V}{\partial \varphi}\right)\Big|_{\varphi_0}$. Assuming the perturbation is indeed small we can write the Klein-Gordon equation for the fluctuation as

$$\ddot{\delta\varphi} + 3H\dot{\delta\varphi} - \frac{\nabla^2}{a^2}\delta\varphi = -\frac{\partial^2 V(\varphi)}{\partial\varphi^2}\delta\varphi\,. \tag{2.110}$$

Let us move to the Fourier space which is more convenient for the problem. We can write,

$$\delta\varphi(\mathbf{x}, t) = \frac{1}{(2\pi)^3}\int \mathrm{d}^3k\, e^{i\mathbf{k}\cdot\mathbf{x}}\delta\varphi_{\mathbf{k}}(t)\,. \tag{2.111}$$

Since the field is real, $\delta\varphi_{\mathbf{k}} = \delta\varphi^*_{-\mathbf{k}}$.
Now we will write the Klein-Gordon equation for the quantum fluctuation in Fourier space.

$$\ddot{\delta\varphi_{\mathbf{k}}} + 3H\dot{\delta\varphi_{\mathbf{k}}} + \frac{k^2}{a^2}\delta\varphi_{\mathbf{k}} = -\partial^2_{\varphi}(V)\delta\varphi_{\mathbf{k}}\,. \tag{2.112}$$

If we consider a massless scalar field $\partial^2_{\varphi}(V) \approx 0$ which corresponds to the requirement of the parameter $\eta \ll 1$ which we discussed in the previous section. So, we can write the above equation as

$$\ddot{\delta\varphi_{\mathbf{k}}} + 3H\dot{\delta\varphi_{\mathbf{k}}} + \frac{k^2}{a^2}\delta\varphi_{\mathbf{k}} \simeq 0 \tag{2.113}$$

Now let us distinguish between the **sub-horizon**($\lambda < r_H$) and the **super-horizon**($\lambda > r_H$) regimes on the basis of the comoving wavelength $\lambda_{com}a(t) = \lambda_{physical} \simeq 2\pi/k_{physical}$

- **sub-horizon**($\lambda < r_H$):

$$\lambda \ll \frac{1}{aH} \implies \frac{k}{aH} \gg 1 \tag{2.114}$$

We can write Eq. (2.113) as

$$\ddot{\delta\varphi_{\mathbf{k}}} + (3H^2 + \frac{k^2}{a^2})\delta\varphi_{\mathbf{k}} = \ddot{\delta\varphi_{\mathbf{k}}} + \frac{k^2}{a^2}\delta\varphi_{\mathbf{k}} \simeq 0 \tag{2.115}$$

where we used the Hubble time scale $t_{\mathrm{H}} = H^{-1}$ as the time reference and the sub-horizon condition Eq. (2.114). So we observe that in the sub-horizon regime, the Fourier transform of the quantum fluctuations of the scalar field $\delta\varphi_{\mathbf{k}}$ can be described by a harmonic oscillator equation with a frequency term given by the factor $k/a(t)$ where the scale factor depends upon time as $a(t) \propto e^{Ht}$. This behaviour is expected, as at small scales inside the horizon, the local space-time appears flat like Minkowski space-time, and the expansion of the universe is negligible.

- **super-horizon**($\lambda > r_H$):

$$\lambda \gg \frac{1}{aH} \implies \frac{k}{aH} \ll 1, \tag{2.116}$$

In this regime, we proceed as in the sub-horizon regime and observe that for the above condition given in Eq. (2.116)

$$3H^2\delta\varphi_{\mathbf{k}} \ll \frac{1}{a^2}k^2\delta\varphi_{\mathbf{k}} \tag{2.117}$$

which shows that the friction term is dominant with respect to the Laplacian, therefore we can write Eq.(2.113) as

$$\ddot{\varphi_{\mathbf{k}}} + 3H\dot{\delta}\varphi_{\mathbf{k}} \simeq 0 \tag{2.118}$$

The solution for the above second-order differential Eq. (2.118) is given by

$$\delta\varphi_{\mathbf{k}} = A + Be^{-3Ht}. \tag{2.119}$$

from this, we can interpret that as the exponential term decays and we get towards a constant fluctuation. In other words, as time increases the fluctuation oscillates until the wavelength gets larger than the Hubble horizon (maximum distance of causal connection), as soon as it is larger, the fluctuations can't interact nor grow and they cease to oscillate and get *"frozen-in"*

Dynamics of Quantum Fluctuations: exact solution

In this section, we will discuss the exact solution to the dynamics of quantum fluctuation of a more generic scalar field $\hat{\delta\varphi}(\tau, \mathbf{x})$ that includes quantum field theory effects and the mass term given by $\partial^2_\varphi V$. As done in the previous section, we explore the solution in the sub-horizon and super-horizon limits.

We start by redefining the field as,

$$\hat{\delta\varphi}(\tau, \mathbf{x}) = a(\tau)\delta\varphi(\tau, \mathbf{x}), \tag{2.120}$$

here we need to take into note that we are using conformal time τ instead of cosmic time. Now, we can write the generic scalar field $\hat{\delta\varphi}(\tau, \mathbf{x})$ as a linear combination of the creation-annihilation operators $\left(a_{\mathbf{k}}, a^\dagger_{\mathbf{k}}\right)$

$$\hat{\delta\varphi}(\tau, \mathbf{x}) = \frac{1}{(2\pi)^3}\int \mathrm{d}^3\mathrm{k}\left[u_{\mathbf{k}}(\tau)a_{\mathbf{k}}e^{i\mathbf{k}\cdot\mathbf{x}} + u^*_{\mathbf{k}}(\tau)a^\dagger_{\mathbf{k}}e^{-i\mathbf{k}\cdot\mathbf{x}}\right] \tag{2.121}$$

and time-dependent mode functions $u_{\mathbf{k}}(\tau)$ that obeys a normalization condition

$$u'^*_{\mathbf{k}}(\tau)u_{\mathbf{k}}(\tau) - u^*_{\mathbf{k}}(\tau)u'_{\mathbf{k}}(\tau) = -i \tag{2.122}$$

such that commutation rules are given by

$$[a_{\mathbf{k}}, a_{\mathbf{k}'}] = \left[a^\dagger_{\mathbf{k}}, a^\dagger_{\mathbf{k}'}\right] = 0, \tag{2.123}$$

$$\left[a_{\mathbf{k}}, a^\dagger_{\mathbf{k}'}\right] = \hbar\delta^{(3)}\left(\mathbf{k} - \mathbf{k}'\right). \tag{2.124}$$

In the Minkowski space-time the solutions i.e the mode functions are described by plane waves such as

$$u_{\mathbf{k}}(\tau) \sim \frac{e^{-i\omega_{\mathbf{k}}\tau}}{\sqrt{2\omega_{\mathbf{k}}}}, \quad \omega_{\mathbf{k}} = \sqrt{k^2 + m^2}. \tag{2.125}$$

However, in the case of an expanding FRW Universe, we have a curved space-time, and we expect a more complicated expression for $u_{\mathbf{k}}(\tau)$. Since, in quantum field theory on curved space-time, the choice of the vacuum state is not unique enough to determine the mode functions due to the presence of the gravitational field. This means that different choices of vacuum state can result in different sets of mode functions, which can affect the physical properties of the theory. So, in short in quantum field theory on curved space-time there is an ambiguity with the choice of the vacuum state, therefore $u_{\mathbf{k}}(\tau)$ is not a priori fixed.
Looking at the equivalence principle as a guiding principle, modes $u_k(\tau)$ at very short distances must reproduce the form for the ordinary flat space-time quantum field theory, which is the plane waves. So, we require

$$\frac{k}{aH} \to \infty \implies u_{\mathbf{k}}(\tau) \sim \frac{e^{-i\mathbf{k}\tau}}{\sqrt{2\mathbf{k}}}, \quad \sqrt{k^2 + m^2} \simeq k \tag{2.126}$$

which is the Bunch-Davies condition on the vacuum state, which states that in the limit of small scales and for initial times the mode functions are given by

$$u_{\mathbf{k}}(\tau) \to \frac{e^{-i\mathbf{k}\tau}}{\sqrt{2\mathbf{k}}}. \tag{2.127}$$

To motivate the inflationary vacuum state we try to recall our previous discussion that at a sufficiently early time all the modes of cosmological interest were deep inside the horizon i.e, which implies

$k/aH \gg 1$ so, we can write $w_k \simeq k$ and there the mode function is given by Eq. (2.127)
Before rewriting equation 1.2.45 in the Fourier space, we will explicitly change the time coordinate going from the reference time t to the conformal time τ such that

$$\frac{d}{dt} \quad \longrightarrow \quad \frac{d}{dt}\frac{d\tau}{d\tau} = \frac{1}{a}\frac{d}{d\tau} \tag{2.128}$$

let us perform the calculations term by term using the coordinate change in equation (2.128). On the left-hand side of equation (2.112) we have

$$\delta\ddot{\varphi}_{\mathbf{k}} = \frac{1}{a}\frac{d}{d\tau}\left[\frac{1}{a}\frac{d}{d\tau}\left(\frac{\delta\hat{\varphi}_{\mathbf{k}}}{a}\right)\right] \tag{2.129}$$

$$= \frac{1}{a}\frac{d}{d\tau}\left[\frac{1}{a}\left(\frac{\delta\hat{\varphi}'_{\mathbf{k}}}{a} - \frac{a'}{a^2}\delta\hat{\varphi}_{\mathbf{k}}\right)\right] \tag{2.130}$$

$$= \frac{1}{a}\left(\frac{\delta\hat{\varphi}''_{\mathbf{k}}}{a^2} - 2\frac{a'}{a^3}\delta\hat{\varphi}'_{\mathbf{k}} - \frac{a''}{a^3}\delta\hat{\varphi}_{\mathbf{k}} - 3\frac{a'^2}{a^4}\delta\hat{\varphi}_{\mathbf{k}} - \frac{a'}{a^3}\delta\hat{\varphi}'_{\mathbf{k}}\right) \tag{2.131}$$

$$= \frac{\delta\hat{\varphi}''_{\mathbf{k}}}{a^3} - 2\frac{a'}{a^4}\delta\hat{\varphi}'_{\mathbf{k}} - \frac{a''}{a^4}\delta\hat{\varphi}_{\mathbf{k}} + 3\frac{a'^2}{a^5}\delta\hat{\varphi}_{\mathbf{k}} - \frac{a'}{a^4}\delta\hat{\varphi}'_{\mathbf{k}}, \tag{2.132}$$

and

$$3H\delta\dot{\varphi}_{\mathbf{k}} = 3\frac{1}{a^2}\frac{da}{d\tau}\frac{1}{a}\frac{d}{d\tau}\left(\frac{\delta\hat{\varphi}_{\mathbf{k}}}{a}\right) \tag{2.133}$$

$$= 3\frac{a'}{a^4}\delta\hat{\varphi}'_{\mathbf{k}} - 3\frac{a'^2}{a^5}\delta\hat{\varphi}_{\mathbf{k}} \tag{2.134}$$

Finally, putting all the results together, we obtain

$$\delta\hat{\varphi}''_{\mathbf{k}} - \frac{a''}{a}\delta\hat{\varphi}_{\mathbf{k}} + k^2\delta\hat{\varphi}_{\mathbf{k}} = -\partial^2_{\varphi}Va^2\delta\hat{\varphi}_{\mathbf{k}} \tag{2.135}$$

which in terms of the mode functions can be written as

$$u''_{\mathbf{k}}(\tau) + \left(k^2 - \frac{a''}{a} + \partial^2_{\varphi}Va^2\right)u_{\mathbf{k}}(\tau) = 0. \tag{2.136}$$

From this, we can see why we chose to use the re-scaled version $\delta\hat{\varphi}_{\mathbf{k}}$. We can notice that the *ansatz* is basically a harmonic oscillator with a time-dependent frequency changing according to the accelerated expansion of the universe. Let us look at the behaviour of the mode function in the sub-horizon and super-horizon regimes for a de Sitter Universe and a quasi-de Sitter Universe.

Solution in de-Sitter

To solve the equation we will first consider the de-Sitter expansion where $H = constant$, $\epsilon \to 0$ and we assume the inflation is massless i.e $m^2_{\varphi} = \partial^2_{\varphi}V = 0$ Under these conditions, we can write the conformal time given by Eq. (2.8) as

$$d\tau = \frac{dt}{a(t)} = dte^{-Ht} \tag{2.137}$$

where we used the $a(t) \propto e^{Ht}$. After integration, we obtain the corresponding scale factor which reads as

$$a(\tau) = -\frac{1}{H\tau}(\tau < 0) \tag{2.138}$$

and

$$\frac{a''}{a} = 2a^2H^2 \tag{2.139}$$

which on plugging it into Eq. (2.136) gives

$$u''_{\mathbf{k}}(\tau) + \left(k^2 - 2a^2H^2\right) u_{\mathbf{k}}(\tau) = 0. \tag{2.140}$$

In the **sub-horizon**($k \gg aH$) limit, Eq. (2.140) becomes

$$u''_{\mathbf{k}}(\tau) + k^2 u_{\mathbf{k}}(\tau) = 0 \quad \Rightarrow \quad u_{\mathbf{k}}(\tau) = \frac{e^{-ik\tau}}{\sqrt{2k}} \tag{2.141}$$

where we chose the Bunch-Davies vacuum state. As predicted in the qualitative analysis in the previous section, the behaviour of quantum fluctuations with wavelength within the cosmological horizon oscillates as we saw before in a flat space-time quantum field theory. This is what was expected for wavelengths much smaller than the Hubble radius scale where it is a good approximation to approximate the space-time as flat.

In the **super-horizon regime** ($k \ll aH$), Eq. (2.140) approximates as

$$u''_{\mathbf{k}}(\tau) - \frac{a''}{a} u_{\mathbf{k}}(\tau) = 0 \tag{2.142}$$

whose solution is simply given by

$$u_{\mathbf{k}}(\tau) = \underbrace{B(k)a(\tau)}_{\text{growing mode}} + \underbrace{C(k)a(\tau)^{-2}}_{\text{decaying mode}}. \tag{2.143}$$

We see that the solution has a growing and a decaying mode where the decaying mode decays in time as $a(\tau)^{-2}$, so we can neglect the decaying mode because even if the amplitude increases it will decay quickly. Therefore the amplitude of the physical fluctuation reads as

$$|\delta\varphi_{\mathbf{k}}| \propto \frac{|u_{\mathbf{k}}|}{a(\tau)} = |B(k)| \tag{2.144}$$

In the above equation, we can see that the "growing mode" is actually asymptotically constant on super-horizon scales since $|B(k)|$ is independent of τ. To determine the scale of $B(k)$ we need to match the super-horizon and sub-horizon solutions, so it is basically a link between the quantum perturbation at horizon exit time during inflation and perturbation at horizon re-entry time during radiation or matter-dominated era. We thus evaluate $|B(k)|$ at the time of horizon crossing during inflation and we get

$$|B(k)|a = \left|\frac{e^{-ik\tau}}{\sqrt{2k}}\right| \tag{2.145}$$

$$|\delta\varphi_{\mathbf{k}}| = |B(k)| = \frac{1}{a\sqrt{2k}} = \frac{H}{\sqrt{2k^3}}. \tag{2.146}$$

The fluctuation gets "frozen in" at horizon crossing:

$$|\delta\varphi_{\mathbf{k}}| = \frac{H}{\sqrt{2k^3}}\,. \tag{2.147}$$

Without proof, we will state the following result: the exact solution for the perturbation growth in De Sitter space-time is

$$u_{\mathbf{k}}(\tau) = \frac{e^{ik\tau}}{\sqrt{2k}}\left(1 - \frac{i}{k\tau}\right). \tag{2.148}$$

As shown above, the physics underlying small scales below the horizon is described by special relativity, while quantum fluctuations around the vacuum expectation value of the scalar field are treated using quantum field theory in flat space-time. These fluctuations have a mean value of zero due to the nature of the vacuum state, which is characterized by the creation and annihilation of particles with a net number of particles equal to zero.

During inflation, the comoving Hubble radius shrinks, causing the rapid expansion to stretch the wavelength of the quantum fluctuations. This results in all fluctuations generated at sub-horizon scales exiting the horizon, where their amplitudes become frozen and are not affected by causal contact. After inflation ends, the comoving horizon begins to grow again, causing all fluctuations to re-enter the horizon with an imprint of their primordial fluctuations generated by inflation but with a much larger physical wavelength.

Solution in Quasi- de Sitter

In this section, we will see how to solve the equation (2.136) explicitly

$$u''_{\mathbf{k}}(\tau) + \left[k^2 - \frac{a''}{a} + a^2m^2\right]u_{\mathbf{k}}(\tau) = 0\,, \tag{2.149}$$

where τ is the conformal time, while $m^2 = \partial^2 V/\partial\varphi^2$. The stage in which $m^2 = 0$ is the quasi-de-Sitter stage. Now we will consider that H is not a constant, instead, we will consider the effect of a nonzero slow-roll parameter $\epsilon = -\dot{H}/H^2 \ll 1$. .We will use the relation

$$\frac{\ddot{a}}{a} = H^2(1-\epsilon)\,, \tag{2.150}$$

which can be proven by straightforward manipulation.

Using the definition of the conformal time, we can show that the scale factor for small values of ϵ becomes

$$\tau = -\frac{1}{aH(1-\epsilon)} \tag{2.151}$$

Using this, we find that to first order

$$(aH)^2 \approx (1+2\epsilon)/\tau^2 \tag{2.152}$$

therefore we can re-write Eq. (2.150) using Eq. (2.152) in terms of conformal time as,

$$\frac{a''}{a} = a^2H^2(2-\epsilon) = 2\frac{(1+2\epsilon)}{\tau^2}\left(1 - \frac{\epsilon}{2}\right) \approx \frac{2}{\tau^2}\left(1 + \frac{3}{2}\epsilon\right). \tag{2.153}$$

Substituting this for a''/a the equation becomes

$$u''_{\mathbf{k}}(\tau) + \left[k^2 - \frac{\nu^2 - 1/4}{\tau^2}\right] u_{\mathbf{k}}(\tau) = 0\,, \tag{2.154}$$

where $\nu^2 = 9/4 + 3\epsilon$. This is a Bessel equation: these equations are generally in the form

$$z^2 y''(z) + z y'(z) + (z^2 - \nu^2) y(z) = 0\,. \tag{2.155}$$

The solutions are called Hankel functions, assuming ν as constant the general solution is given as

$$u_{\mathbf{k}}(\tau) = \sqrt{-\tau}\Big[c_1(k) H_\nu^{(1)}(-k\tau) + c_2(k) H_\nu^{(2)}(-k\tau)\Big] \tag{2.156}$$

where $H_\nu^{(2)} = H_\nu^{(1)*}$. We want to impose the asymptotic behaviour of the solution: let us start with

- **Sub-horizon ($k/aH \gg 1$):**
  . This means that

  $$u_{\mathbf{k}}(\tau) = \frac{1}{\sqrt{2k}} e^{-ik\tau}\,,$$

  and the asymptotics of the Hankel functions are

  $$H_\nu^{(1)}(x) \sim \sqrt{\frac{2}{\pi x}} \exp\Big(i\Big(x - \frac{\pi}{2}\nu - \frac{\pi}{4}\Big)\Big) \sim \frac{e^{ix}}{\sqrt{x}} \tag{2.157}$$

  for $x \gg 1$. This works well for us: we can set $c_2(k) = 0$ and only use $c_1(k)$.
  We must choose

  $$c_1(k) = \frac{\sqrt{\pi}}{2} \exp\left(i\left(\nu + \frac{1}{2}\right)\frac{\pi}{2}\right) \tag{2.158}$$

  in order to have the correct normalization. Then, our solution will read

  $$u_{\mathbf{k}}(\tau) = \frac{\sqrt{\pi}}{2} \exp\left(i(\nu + \frac{1}{2})\frac{\pi}{2}\right) \sqrt{-\tau} H_\nu^{(1)}(-k\tau)\,, \tag{2.159}$$

- **Super-horizon ($k/aH \ll 1$):**
  We have the asymptotic $x \ll 1$ expansion of the Hankel function:

  $$H_\nu^{(1)}(x) \sim \sqrt{\frac{2}{\pi}} e^{-i\pi/2} 2^{\nu - 3/2} \frac{\Gamma(\nu)}{\Gamma(3/2)} x^{-\nu} \sim x^{-\nu}\,, \tag{2.160}$$

  therefore

  $$u_{\mathbf{k}}(\tau) \approx \exp\left(i(\nu - \frac{1}{2})\frac{\pi}{2}\right) 2^{\nu - 3/2} \frac{\Gamma(\nu)}{\Gamma(3/2)} (-k\tau)^{1/2 - \nu} \frac{1}{\sqrt{2k}}\,, \tag{2.161}$$

so for getting the modulus $|\delta\varphi_{\mathbf{k}}| = |u_{\mathbf{k}}|/a$ we can use the fact that $\tau \sim -1/aH$, therefore $|\delta\varphi_k| \approx -H\tau|u_k|$. Inserting the expression we have for u_k we get,[6]

$$|\delta\varphi_{\mathbf{k}}| \approx 2^{\nu-3/2}\frac{\Gamma(\nu)}{\Gamma(3/2)}\frac{H}{\sqrt{2k^3}}\left(\frac{k}{aH}\right)^{3/2-\nu} \tag{2.162}$$

$$\approx \frac{H}{\sqrt{2k^3}}\left(\frac{k}{aH}\right)^{3/2-\nu} \approx \frac{H}{\sqrt{2k^3}}\left(\frac{k}{aH}\right)^{-\epsilon}, \tag{2.163}$$

which holds at super-horizon scales. In the De Sitter case, the exact result reads

$$u_{\mathbf{k}}(\tau) \propto \sqrt{-\tau}H^{(1)}_{3/2}(-k\tau) = \frac{e^{-ik\tau}}{\sqrt{2k}}\left(1 - \frac{i}{k\tau}\right). \tag{2.164}$$

Let's take a step forward in generalizing the scalar field with non-zero mass, denoted by $m^2 = \frac{\partial^2 V}{\partial\varphi^2}$. We require this mass to be much smaller than the Hubble parameter, H. This condition ensures that the scalar field can drive inflation through slow-roll. Specifically, we need the slow-roll parameter $\eta_V = \frac{m^2}{3H^2} \ll 1$.

The reason for this requirement is that a massive scalar field (compared to H^2) would not be able to maintain the slow-roll conditions needed for inflation. In other words, its mass would dominate over the Hubble friction term and prevent the scalar field from smoothly rolling down its potential.

Then, the m^2a^2 term in the equation represents the mass term of the scalar field in an expanding universe which can be written as

$$m^2a^2 = 3\eta a^2H^2 = \frac{3\eta}{\tau^2}, \tag{2.165}$$

but we also know that

$$\frac{a''}{a} = \frac{2}{\tau^2}\left[1 + \frac{3}{2}\epsilon\right], \tag{2.166}$$

Hence, the equation can still be recast into the same Bessel form but this time only the coefficient of the τ^{-2} term changes.

The equation will read

$$u''_{\mathbf{k}}(\tau) + \left[k^2 - \frac{\nu^2 - 1/4}{\tau^2}\right]u_{\mathbf{k}}(\tau) = 0, \tag{2.167}$$

with $\nu^2 = 9/4 + 3\epsilon - 3\eta$. Then, the solution has exactly the same form as before:
on super-horizon scales, with $k \ll aH$, we have

$$|\delta\varphi_k| = \frac{H}{\sqrt{2k^3}}\left(\frac{k}{aH}\right)^{3/2-\nu} \qquad \nu \approx \frac{3}{2} + \epsilon - \eta. \tag{2.168}$$

It's worth noting that the previous computation doesn't only apply to an inflaton field, but also to any scalar field evolving during this phase of the universe's expansion. However, it's important to consider the field's mass - it can be shown that if m^2 is much greater than H^2, the super-horizon fluctuations of the field will have difficulty being excited and the field will remain in its vacuum state. So, if it's another scalar field, it needs to be light to be able to generate super-horizon fluctuations.

[6] Expanded to first order in ϵ, so that $3/2 - \nu \approx -\epsilon$.

2.7.5 The Power Spectrum

The power spectrum is a useful quantity to characterize the statistical properties of a Gaussian perturbation field which helps us to connect theoretical predictions with observational.[7]

The properties of a Gaussian random field $\delta(t, \mathbf{x})$ that represents a fluctuation in a point of space-time, such as the density fluctuation, can be described by an infinite set of correlation functions. These functions are defined as:

$$\xi\left(\mathbf{x}_1, \mathbf{x}_2\right) = \left\langle \delta\left(t, \mathbf{x}_1\right) \delta\left(t, \mathbf{x}_2\right) \right\rangle, \tag{2.169}$$

$$\xi\left(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\right) = \left\langle \delta\left(t, \mathbf{x}_1\right) \delta\left(t, \mathbf{x}_2\right) \delta\left(t, \mathbf{x}_3\right) \right\rangle, \tag{2.170}$$

$$\vdots \tag{2.171}$$

$$\xi\left(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N\right) = \left\langle \delta\left(t, \mathbf{x}_1\right) \delta\left(t, \mathbf{x}_2\right) \ldots \delta\left(t, \mathbf{x}_N\right) \right\rangle, \tag{2.172}$$

where the $\langle \cdot \rangle$ symbol refers to the ensemble average.

In the case of Gaussian random fields, they are determined by the two-point correlation function, i.e. the first definition in equation (2.169). For reflection-symmetric Gaussian processes, the correlation function for odd N is identically zero, whereas those characterized by even N can be written as a combination of $\xi\left(\mathbf{x}_1, \mathbf{x}_2\right)$.

The homogeneity condition implies that the statistical properties of the field are the same at every point in space. This means that the two-point correlation function, which characterizes the statistical dependence between two points in space, should only depend on the distance between the two points and not on their absolute positions. The isotropy condition applied to the two-point correlation function means that it should be invariant under spatial rotations. This implies that it should only depend on the magnitude of the separation vector between two points, and not on its direction. Therefore, the correlation function should only depend on the relative distance, and not on the relative orientation or direction of the separation vector between the two points $\xi(r) = \xi\left(\left|\mathbf{x}_1 - \mathbf{x}_2\right|\right)$.

Knowing that the Fourier transform of the fluctuation reads as

$$\delta(t, \mathbf{x}) = \frac{1}{(2\pi)^3} \int \mathrm{d}^3 k e^{i\mathbf{k}\cdot\mathbf{x}} \delta_{\mathbf{k}}(t) \tag{2.173}$$

we define the power spectrum $P(k)$ as

$$\left\langle \delta_{\mathbf{k}_1}(t) \delta_{\mathbf{k}_2}(t) \right\rangle = (2\pi)^3 \delta^{(3)}\left(\mathbf{k}_1 + \mathbf{k}_2\right) P(k), \tag{2.174}$$

and we show that the power spectrum is the Fourier transform of the two-point correlation function $\xi(r)$. In fact

$$\xi(r) = \langle \delta(t, x + r) \delta(t, x) \rangle \tag{2.175}$$

$$= \left\langle \frac{1}{(2\pi)^3} \int \mathrm{d}^3 k e^{i\mathbf{k}\cdot(\mathbf{x}+\mathbf{r})} \delta_{\mathbf{k}}(t) \frac{1}{(2\pi)^3} \int \mathrm{d}^3 k' e^{i\mathbf{k}'\cdot\mathbf{x}} \delta_{\mathbf{k}'}(t) \right\rangle \tag{2.176}$$

$$= \frac{1}{(2\pi)^6} \int \mathrm{d}^3 k \int \mathrm{d}^3 k' e^{i\mathbf{k}\cdot(\mathbf{x}+\mathbf{r})} e^{\mathbf{k}'\cdot\mathbf{x}} \left\langle \delta_{\mathbf{k}}(t) \delta_{\mathbf{k}'}(t) \right\rangle \tag{2.177}$$

$$= \frac{1}{(2\pi)^3} \int \mathrm{d}^3 k \int \mathrm{d}^3 k' e^{i\mathbf{k}\cdot(\mathbf{x}+\mathbf{r})} e^{i\mathbf{k}'\cdot\mathbf{x}} \delta^{(3)}\left(\mathbf{k} + \mathbf{k}'\right) P(k) \tag{2.178}$$

$$= \frac{1}{(2\pi)^3} \int \mathrm{d}^3 k e^{i\mathbf{k}\cdot\mathbf{r}} P(k). \tag{2.179}$$

The plus is there since we wrote $\delta(k)\delta(k')$ instead of $\delta(k)\delta^*(k')$: if we included the conjugate then the argument of the δ would be $k - k'$, due to the fact that since $\delta(x)$ is real we have $\delta^*(k) = d(-k)$.

[7] Perturbations close to Gaussian.

Another relevant quantity is the variance of the fluctuations defined as

$$\sigma^2 \equiv \langle \delta_{\rm k}^2 \rangle = \frac{1}{(2\pi)^2} \int {\rm d}k k^2 P(k) \tag{2.180}$$

We further introduce the dimensionless power spectrum

$$\mathscr{P}(k) \equiv \frac{k^3}{2\pi^2} P(k), \tag{2.181}$$

by which Eq. (2.180) becomes

$$\sigma^2 = \int_0^\infty \frac{{\rm d}k}{k} \mathscr{P}(k) \tag{2.182}$$

The scale dependence of $\mathscr{P}(k)$ is given by the spectral index

$$n_s(k) - 1 = \frac{d \ln \mathscr{P}(k)}{d \ln k}. \tag{2.183}$$

If the spectral index is scale-invariant, i.e. $n_s = $ const, then the power spectrum $\mathscr{P}(k)$ can be generally written as

$$\mathscr{P}(k) = \mathscr{A}_s \left(\frac{k}{k_0} \right)^{n_s - 1}, \tag{2.184}$$

where k_0 is a pivot scale. A useful way to categorize different power spectra is through the spectral index, denoted by n_s. The Harrison-Zel'dovich power spectrum, corresponding to $n_s = 1$, is scale-invariant and does not depend on the cosmological scale k and $\mathscr{A}_s = \mathscr{P}(k_0)$. A blue-tilted power spectrum, with $n_s > 1$, implies that perturbations have more power on small scales than on large scales. Conversely, a red-tilted power spectrum, with $n_s < 1$, indicates that there is less power on small scales compared to large scales. The power spectrum of the fluctuations of the inflaton field at lowest order in the slow-roll parameters. Using the fact that $\delta\varphi_{\mathbf{k}}(t)^* = \delta\varphi_{-\mathbf{k}}(t)$ and that it can be written as a combination of creation-annihilation operators [8] we obtain

$$\langle \delta\varphi_{\mathbf{k}}(t) \delta\varphi^*_{\mathbf{k}'}(t) \rangle = (2\pi)^3 \delta^{(3)} \left(\mathbf{k} - \mathbf{k}' \right) |\delta\varphi_{\mathbf{k}}|^2 . \tag{2.185}$$

Comparing this result with equation (2.174) we have

$$P_{\delta\varphi}(k) = |\delta\varphi_{\mathbf{k}}(t)|^2 , \quad \mathscr{P}_{\delta\varphi}(k) = \frac{k^3}{2\pi^2} |\delta\varphi_{\mathbf{k}}|^2 . \tag{2.186}$$

In particular, using the result in (2.168) we find that at super-horizon scales the power spectrum reads as

$$\mathscr{P}_{\delta\varphi}(k) = \left(\frac{H}{2\pi} \right)^2 \left(\frac{k}{aH} \right)^{3-2\nu} . \tag{2.187}$$

[8] Writing the L.H.S of (2.185) in terms of creation and annihilation operators: we have

$$\langle 0 | aa | 0 \rangle = \langle 0 | a^\dagger a | 0 \rangle = \langle 0 | a^\dagger a^\dagger | 0 \rangle = 0 ,$$

while

$$\langle 0 | aa^\dagger | 0 \rangle = \underbrace{\langle 0 | \left[a, a^\dagger \right] | 0 \rangle}_{\delta^{(3)}(\mathbf{k}-\mathbf{k}')} - \underbrace{\langle 0 | a^\dagger | 0 \rangle}_{=0}$$

so,

$$\langle \delta\varphi_{\mathbf{k}} \delta\varphi^*_{\mathbf{k}'} \rangle = (2\pi)^3 |\delta\varphi_{\mathbf{k}}|^2 \delta^3(\mathbf{k} - \mathbf{k}')$$

2.7.6 From Quantum to Cosmological Fluctuations

During inflation, the inflaton field fluctuations $\delta\varphi$ arise at sub-horizon scales and are stretched by the rapid expansion of the Universe until they exit the horizon and freeze at super-horizon scales, as shown in section (2.7.4). Later on, these fluctuations re-enter the region of causal connection and interact through physical processes. In this section, we investigate how the quantum perturbations during inflation relate to the primordial density perturbations that act as the seeds for the current cosmic structure.

Due to the presence of these fluctuations, different regions of the Universe complete inflation at slightly different times. This is demonstrated by considering the equation of motion of the homogeneous inflaton field (2.77) and splitting the scalar field as

$$\varphi(\mathbf{x}, t) = \varphi_0(t) + \delta\varphi(\mathbf{x}, t)\,, \tag{2.188}$$

Now taking its derivative with respect to cosmic time we get

$$\ddot{\varphi}\dot{\varphi} + 3H\dot{\varphi}^2 = -\partial_\phi^2 V(\varphi)\dot{\varphi} \implies \partial_t^2(\dot{\varphi}) + 3H\partial_t(\dot{\varphi}) = -\partial_\phi^2 V(\varphi)\dot{\varphi} \tag{2.189}$$

we notice that the resulting equation has the same form as Eq. (2.112) except for the Laplacian term. In the limit where this term can be neglected, $\delta\varphi$ and φ obey the same equation. In particular, taking the Fourier transform of the Laplacian term

$$\frac{1}{a^2}\nabla^2(\delta\varphi) \xrightarrow{\mathscr{F}} -\frac{k^2}{a^2}\delta\varphi_{\mathbf{k}} \tag{2.190}$$

we have that it vanishes in the limit of large scales, i.e. for k that tends to 0. This implies that the Laplacian term can be ignored on large scales, and we can assume a coarse-graining process that is valid for the discussion of inflationary fluctuations.

The Wronskian operator, a mathematical tool that enables us to determine whether a group of solutions to a differential equation are linearly independent or linearly dependent, is used to analyze this system. Given two homogeneous fields $\varphi(t)$ and $\psi(t)$, the Wronskian is defined as

$$W(\varphi, \psi) = \dot{\varphi}\psi - \dot{\psi}\varphi \begin{cases} \neq 0 & \text{two linearly independent solutions,} \\ = 0 & \text{two linearly dependent solutions.} \end{cases} \tag{2.191}$$

In our case, the Wronskian has the following form

$$W(\dot{\varphi}_0, \delta\varphi) = \ddot{\varphi}_0\delta\varphi - \dot{\varphi}_0\dot{\delta\varphi}\,. \tag{2.192}$$

If we further derive with respect to time, we obtain that

$$\dot{W} = \dddot{\varphi}_0\delta\varphi + \ddot{\varphi}_0\dot{\delta\varphi} - \ddot{\varphi}_0\dot{\delta\varphi} - \dot{\varphi}_0\ddot{\delta\varphi} \tag{2.193}$$

$$= \delta\varphi\big(-3H\ddot{\varphi}_0 - V''\dot{\varphi}_0\big) - \dot{\varphi}_0\Big(-3H\dot{\delta\varphi} - V''\delta\varphi\Big) \tag{2.194}$$

$$= -3H\Big(\ddot{\varphi}_0\delta\varphi - \dot{\varphi}_0\dot{\delta\varphi}\Big) = -3HW(\dot{\varphi}_0, \delta\varphi)\,. \tag{2.195}$$

which has a solution decaying with time given by

$$W(\dot{\varphi}_0, \delta\varphi) = e^{-3Ht} \tag{2.196}$$

Thus, using the conditions (2.191), at large scales $\delta\varphi$ and $\dot{\varphi}_0$ are related. In particular

$$\delta\varphi(t,\mathbf{x}) \propto \dot{\varphi_0} \Rightarrow \delta\varphi(t,\mathbf{x}) \simeq (-\delta t(\mathbf{x}))\dot{\varphi_0} \tag{2.197}$$

where the proportional constant $-\delta t(\mathbf{x})$ must have the dimensionality of time as $\delta\varphi(t,\mathbf{x})$is space dependent($\mathbf{x}$ that represents a large region of the Universe) whereas $\dot{\varphi_0}$ is not.So, by reverse-Taylor expansion, the full field is given by

$$\varphi(t,\mathbf{x}) = \varphi_0(t) + \delta\varphi(t,\mathbf{x}) \tag{2.198}$$
$$\simeq \varphi_0[t - \delta t(\mathbf{x})]. \tag{2.199}$$

The inflaton field is subject to a temporal shift given by

$$\delta t(\mathbf{x}) \simeq -\frac{\delta\varphi(t,\mathbf{x})}{\dot{\varphi}(t)} \tag{2.200}$$

The expression (2.199) tells us that at each point in space $\mathbf{x}$, the field φ takes on the same values as the field φ_0 at a slightly earlier time $t - \delta t(\mathbf{x})$. In other words, the value of the field at each point evolves in time, but the evolution is slightly delayed at each point. This is because the field at each point feels the influence of its surroundings, and fluctuations in the surrounding field cause the field at that point to evolve differently than if it were completely isolated.

This interpretation is intuitive if we consider the analogy of waves on a water surface. Imagine that we drop a pebble in the middle of a still pond, creating a circular wave pattern. As the waves propagate outward, the shape of the wave pattern at any point in the pond depends on the position and shape of the waves at all other points. Similarly, in a fluctuating field like φ, the value of the field at any point depends on the values of the field at all other points.

In this sense, the delayed evolution of the field at each point can be thought of as a result of the time it takes for information to propagate through the field. Information about fluctuations in the field has to propagate from one point to another, and this takes time. Therefore, the evolution of the field at each point is slightly delayed relative to the evolution of the field as a whole. As a consequence, also the expansion of the Universe is different in different parts of the Universe.
We introduce the fluctuations in the number of e-folds $\mathscr{N}$ Eq. (2.7.1) given by the primordial curvature perturbation ζ

$$\zeta = \delta\mathscr{N} = H\delta t \simeq -H\frac{\delta\varphi}{\dot{\varphi}} \tag{2.201}$$

where in the last passage we used Eq. (2.200). We now show that during inflation the primordial curvature perturbation induces fluctuations in the density field. Let us consider the expression of ρ in equation (2.86), then

$$-H\frac{\delta\rho}{\dot{\rho}} \simeq -H\frac{\partial_\varphi V(\varphi)\delta\varphi}{-3H(\rho+p)}, \tag{2.202}$$
$$\simeq H\frac{3H\dot{\varphi}\delta\varphi}{-3H\dot{\varphi}^2} \simeq -H\frac{\delta\varphi}{\dot{\varphi}}, \tag{2.203}$$

where in the first passage we used the continuity equation (2.20c) during inflation. Therefore, $\zeta \simeq -H\delta\rho/\dot{\rho}$. An important property of ζ is that it remains constant outside the horizon for adiabatic matter perturbations[Bartolo et al., 2004]. Therefore, given a scale k, we can evaluate ζ at the time of horizon crossing $t(k)_{\text{exit}}$, knowing that it maintains the same value until it re-enters the horizon at $t(k)_{\text{enter}}$ during radiation or matter domination.

The power spectrum of the curvature perturbation can be derived using Eq. (2.201)

$$\mathscr{P}_\zeta \simeq \frac{H^2}{\dot{\varphi}^2}\mathscr{P}_{\delta\varphi} = \left(\frac{H^2}{2\pi\dot{\varphi}}\right)^2\left(\frac{k}{aH}\right)^{3-2\nu} \tag{2.204}$$

and at the time of horizon exit ($k \approx aH$), we have

$$\mathscr{P}_\zeta = \left(\frac{H^2}{2\pi\dot{\varphi}}\right)^2\Bigg|_{t(k)_{exit}}. \tag{2.205}$$

The curvature perturbation ζ introduced in Eq.(2.201) is a particular form of the gauge-invariant curvature perturbation on uniform-density hyper-surfaces defined as

$$\zeta \equiv -\psi - H\frac{\delta\rho}{\dot{\rho}} \tag{2.206}$$

where ψ is the scalar perturbation of the diagonal spatial component of the FRW metric known as the Bardeen Potential. In fact, in Eq. (2.201) we considered the gauge choice $\psi = 0$, called **uniform curvature gauge**.
To calculate the curvature perturbation on large scales we can consider the curvature perturbation on comoving surfaces which in the case of the scalar field reads as,

$$\mathscr{R}_c = \psi + H\frac{\delta\varphi}{\dot{\varphi}}, \tag{2.207}$$

which is the metric perturbation in the comoving space-time slicing(observer which sees the cosmic expansion as isotropic). It can be shown that both $\mathscr{R}_c$ and ζ are gauge invariant and on large scales they are equal,

$$\mathscr{R}_c = -\zeta = \psi + H\frac{\delta\varphi}{\dot{\varphi}}, \tag{2.208}$$

The scale index of the power spectrum can be parameterized using a scale dependence of the spectral index n_s (2.184):

$$\mathscr{P}_\zeta = A_s(\frac{k}{k_0})^{n_s-1+\frac{1}{2}\alpha_s ln(k/k_0)+\frac{1}{3!}\beta_s ln^2(k/k_0)+....}, \tag{2.209}$$

where A_s is the scalar amplitude and the coefficients α_s,β_s are, respectively, called "running of spectral index" and "running of running" parameters. They are independent of k and are obtained by Taylor expansion of the spectral index. In the slow roll approximation, we can write $\mathscr{P}_\zeta(k)$ as a function of the slow roll potential and its derivatives, hence we can write the running parameters as a function of the slow roll parameters[Baumann, 2012].

$$1 - n_s = 2\eta - 6\epsilon \tag{2.210}$$

$$\alpha_s = -2\xi^2 + 16\eta\epsilon - 24\epsilon^2 \tag{2.211}$$

$$\beta_s = 2\sigma^3 + +2\xi^2(\eta - 12\epsilon) - 32\epsilon(\eta^2 - 6\eta\epsilon + 6\epsilon^2). \tag{2.212}$$

In the standard slow-roll inflation models, α_s and β_s are highly suppressed by slow-roll parameters. However, in the large-running mass inflation model, these parameters would be relatively large.

2.8 Primordial Non-Gaussianity

Previously, we saw that for a single field inflationary model, the action considered is that of a single field coupled to gravity Eq. (2.71) and the leading order term in this expansion is quadratic in the small fluctuations around the homogeneous background. Since a free field can be viewed as a collection of harmonic oscillators, and these oscillators begin in their ground state, the resulting fluctuations are Gaussian to the leading order. However, there can be deviations from the Gaussian initial conditions [Maldacena, 2003] when the conditions for the standard inflationary model are not met,

like, in the case of the multi-field inflation [Polarski and Starobinsky, 1994] or the curvaton scenario [Lyth and Wands, 2002]. For a detailed review of the generation of primordial non-Gaussianity (PNG) from different inflationary models refer to [Bartolo et al., 2004] and the references therein .

In the context of Gaussian random fields, all information is encapsulated in the power spectrum. The bispectrum, which is the Fourier transform of the three-point correlation function, is significant as it serves as the lowest-order statistic capable of distinguishing non-Gaussian features from Gaussian perturbations.

Each inflationary model makes specific predictions regarding the bispectrum of the primordial perturbations in the gravitational potential $\Phi_{\mathbf{k}}$ hence accurately calculating the primordial bispectrum of cosmological perturbations is a crucial area of research for enhancing our understanding of the early universe. The bispectrum $B_\Phi\left(k_1, k_2, k_3\right)$ of these perturbations is defined as:

$$\left\langle \Phi\left(\mathbf{k}_1\right) \Phi\left(\mathbf{k}_2\right) \Phi\left(\mathbf{k}_3\right)\right\rangle \equiv (2\pi)^3 \delta_D\left(\mathbf{k}_{123}\right) B_\Phi\left(k_1, k_2, k_3\right) \tag{2.213}$$

where $\mathbf{k}_{ij} \equiv \mathbf{k}_1 + \mathbf{k}_2$ so that the Dirac delta function here is $\delta_D\left(\mathbf{k}_{123}\right) \equiv \delta_D\left(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3\right)$. Together with the assumption of statistical homogeneity and isotropy for the primordial perturbations, this implies that the bispectrum is a function of the triplet defined by the magnitude of the wavenumbers k_1, k_2 and k_3 forming a closed triangular configuration. Additionally. we can write

$$B_\Phi\left(k_1, k_2, k_3\right) \equiv f_{\mathrm{NL}} F\left(k_1, k_2, k_3\right) \tag{2.214}$$

where $F\left(k_1, k_2, k_3\right)$ encodes the functional dependence of the primordial bispectrum on the specific triangle configurations and f_{NL} is the dimensionless PNG amplitude that measures the strength of the PNG signal.[9]

Predictions for both the amplitude f_{NL} and the functional dependence $F\left(k_1, k_2, k_3\right)$ vary significantly between different inflationary models. Below is a brief description of the three main templates of primordial bispectrum shapes. For more details refer to [Meerburg et al., 2019] and references therein.

The parameter space of single-field models is well-described by the so-called equilateral and orthogonal templates.[Creminelli et al., 2006, Chen et al., 2007]

$$\begin{aligned} B_\Phi^{\mathrm{equil}}\left(k_1, k_2, k_3\right) = 6A^2 f_{\mathrm{NL}}^{\mathrm{equil}} \times \Bigg\{ & -\frac{1}{k_1^{4-n_s} k_2^{4-n_s}} - \frac{1}{k_2^{4-n_s} k_3^{4-n_s}} - \frac{1}{k_3^{4-n_s} k_1^{4-n_s}} - \frac{2}{\left(k_1 k_2 k_3\right)^{2(4-n_s)/3}} \\ & + \left[\frac{1}{k_1^{(4-n_s)/3} k_2^{2(4-n_s)/3} k_3^{4-n_s}} + 5 \text{ perms.}\right]\Bigg\} \end{aligned} \tag{2.215}$$

$$\begin{aligned} B_\Phi^{\mathrm{ortho}}\left(k_1, k_2, k_3\right) = 6A^2 f_{\mathrm{NL}}^{\mathrm{ortho}} \times \Bigg\{ & -\frac{3}{k_1^{4-n_s} k_2^{4-n_s}} - \frac{3}{k_2^{4-n_s} k_3^{4-n_s}} - \frac{3}{k_3^{4-n_s} k_1^{4-n_s}} - \frac{8}{\left(k_1 k_2 k_3\right)^{2(4-n_s)/3}} \\ & + \left[\frac{3}{k_1^{(4-n_s)/3} k_2^{2(4-n_s)/3} k_3^{4-n_s}} + 5 \text{ perms.}\right]\Bigg\} \end{aligned} \tag{2.216}$$

The bispectrum for multi-field models is typically of the local type(for a review refer [Byrnes and Choi, 2010]) given by

$$\begin{aligned} B_\Phi^{\mathrm{local}}\left(k_1, k_2, k_3\right) &= 2 f_{\mathrm{NL}}^{\mathrm{local}}\left[P_\Phi\left(k_1\right) P_\Phi\left(k_2\right) + P_\Phi\left(k_1\right) P_\Phi\left(k_3\right) + P_\Phi\left(k_2\right) P_\Phi\left(k_3\right)\right] \\ &= 2A^2 f_{\mathrm{NL}}^{\mathrm{local}}\left[\frac{1}{k_1^{4-n_s} k_2^{4-n_s}} + \text{cycl.}\right]. \end{aligned} \tag{2.217}$$

[9] "NL" stands for non-linear

To use non-Gaussianities as a probe of the early Universe and the details of inflation, it is essential to measure both the shape of the primordial bispectrum and the magnitude of its signal, specifically by constraining the f_{NL} parameter. The detection of non-Gaussianities can provide a way to distinguish between the different classes of inflationary models and eliminate those that don't predict such an amount of deviation from the exact Gaussian distribution.

2.9 Neutrino Mass

In section 2.6, we discussed how neutrinos decoupled from the plasma in the early universe. These relic neutrinos leave detectable imprints on cosmological observations. The energy distribution and dynamics of these relic neutrinos have significant consequences for the evolution of the Universe.

Initially, the relic neutrinos were relativistic from the time of their decoupling through to recombination. Due to the expansion and cooling of the Universe, their momenta redshifted as $p_\nu \propto 1/a$, eventually causing the energy of most relic neutrinos to be dominated by their rest mass rather than their momentum. Consequently, the energy density in non-relativistic neutrinos contributes to the matter budget of the Universe today. However, since neutrinos were relativistic for much of the Universe's history, their gravitational clustering differs qualitatively from that of cold dark matter (CDM) particles. This difference allows us to distinguish between the contributions of neutrinos and cold dark matter to the overall matter density [LESGOURGUES and PASTOR, 2006]. In this section, we review how neutrino mass affects the evolution of the neutrino energy density and the gravitational clustering of matter following [LESGOURGUES and PASTOR, 2006, Abazajian et al., 2016].

Current constraints from the cosmic microwave background (CMB) on N_{eff} [Aghanim et al., 2020] align closely with the Standard Model expectation of three species of neutrinos and antineutrinos, each described by a relativistic thermal Fermi-Dirac distribution. The distribution function for each species of neutrinos and antineutrinos is given by

$$f_\nu(p) = \frac{1}{e^{ap/T_{\nu 0}} + 1} \tag{2.218}$$

where $T_{\nu 0} \approx 1.7 \times 10^{-4}\,\text{eV}(1.95\,\text{K})$ is the temperature today [Workman et al., 2022]. In Standard Model physics, the spectral shape of the neutrino phase space distribution is preserved with the expansion, so relic neutrinos have retained the relativistic Fermi-Dirac momentum distribution inherited from decoupling even as the individual neutrinos became non-relativistic.

Now, we write the neutrino energy density as

$$\rho_\nu = \sum_i \int \frac{d^3\mathbf{p}}{(2\pi)^3} \frac{\sqrt{p^2 + m_{\nu i}^2}}{e^{ap/T_{\nu 0}} + 1} \tag{2.219}$$

where $m_{\nu i}$ are the three neutrino mass eigenstates. For $T_\nu/a \gg m_{\nu i}$ the neutrino energies are dominated by their momenta and the total energy density behaves like radiation

$$\rho_\nu|_{\text{early}} \approx \frac{7\pi^2}{40} (T_{\nu 0})^4 \frac{1}{a^4} \propto a^{-4} \tag{2.220}$$

While for $T_{\nu 0}/a \ll m_{\nu i}$ the energy density behaves like matter

$$\rho_\nu|_{\text{late}} \approx \sum_i m_{\nu i} \bar{n}_\nu \propto a^{-3} \tag{2.221}$$

where $\bar{n}_\nu$ is the number of neutrinos and antineutrinos in each mass eigenstate

$$\bar{n}_\nu = \int \frac{d^3\mathbf{p}}{(2\pi)^3} \frac{2}{e^{ap/T_{\nu 0}} + 1} \approx \frac{113}{a^3}\,\text{cm}^{-3} \tag{2.222}$$

For a neutrino of mass $m_{\nu i}$, the transition between these two regimes $\left(k_B T_\nu(a) \sim m_{\nu i}^2\right)$ occurs at redshift $z_{\text{nr}} \sim 300\,(m_{\nu i}/0.05\,\text{eV})$. Using Eq. 2.221, the fractional energy density in neutrinos today can be written as

$$\Omega_\nu h^2 \approx \frac{\sum_i m_{\nu i}}{93\,\mathrm{eV}} \tag{2.223}$$

The individual masses of the neutrino states are unknown but neutrino oscillation data specifies the square of two mass splittings $\Delta m^2_{12} = 7.54 \times 10^{-5}\,\mathrm{eV}$, $\left|\Delta m^2_{13}\right| \approx 2.4 \times 10^{-3}\,\mathrm{eV}$ [Olive et al., 2014]. These mass splittings, in combination with the neutrino number density, give a lower limit on the contribution of neutrinos to the cosmic energy budget [Abazajian et al., 2016]

$$\Omega_\nu h^2 \gtrsim 0.0006 \tag{2.224}$$

At $z \ll z_{\mathrm{nr}}$, the matter density of the Universe, which enters into the Hubble equation, is the sum of the CDM, baryon, and massive neutrino energy densities $\Omega_m = \Omega_c + \Omega_b + \Omega_\nu$. Whereas, at $z \gg z_{\mathrm{nr}}$, the matter density is solely made up of the baryon and CDM parts while neutrinos contribute to the radiation density.

Neutrinos do not contribute to gravitational collapse until they transition to a nonrelativistic state. Before this transition, neutrinos free-stream out of gravitational wells, leaving behind cold dark matter (CDM) and baryons. As a result, primordial fluctuations in the neutrino density are damped on scales smaller than the horizon at z_{nr}. In comoving coordinates, this scale is represented by a wave number.

$$k_{\mathrm{nr}} \equiv a_{\mathrm{nr}} H\left(a_{\mathrm{nr}}\right) \approx 0.003 \left(\frac{\Omega_m}{0.3}\frac{m_\nu}{0.05\,\mathrm{eV}}\right)^{1/2} h/\mathrm{Mpc} \tag{2.225}$$

Once the neutrinos are non-relativistic, their finite velocity dispersion still prevents them from clustering on scales smaller than the typical distance a neutrino travels in a Hubble time, $v_\nu/H(a)$ where $v_\nu \approx 3.15 T_{\nu 0}/\left(a m_\nu\right)$ the mean neutrino velocity. In analogy with the Jeans criterion for gravitational collapse, the neutrino free-streaming scale is defined in comoving coordinates as

$$k_{\mathrm{fs}}(a) \equiv \sqrt{\frac{3}{2}}\frac{aH(a)}{v_\nu(a)} \approx 0.04 a^2 \sqrt{\Omega_m a^{-3} + \Omega_\Lambda}\left(\frac{m_\nu}{0.05\,\mathrm{eV}}\right)\, h/\mathrm{Mpc} \tag{2.226}$$

On scales larger than k_{nr}, adiabatic perturbations in the densities of neutrinos, baryons, and cold dark matter (CDM) are coherent and can be described by a single perturbation to the total matter density $\delta_m = \delta\rho_m/\rho_m$. On smaller scales, where neutrino perturbations have decayed, only perturbations in CDM and baryons remain, such that $\delta_m = \delta_{cb}\left(\Omega_c + \Omega_b\right)/\Omega_m$. The remaining perturbations in CDM and baryons grow more slowly because the neutrino energy density affects the expansion rate but not the source potentials. These two effects lead to a suppression in the amplitude and growth rate of matter perturbations with wavenumbers $k > k_{\mathrm{fs}}$ compared to a universe with massless neutrinos and to density perturbations with $k < k_{\mathrm{nr}}$. The net change in the amplitude of perturbations with $k > k_{\mathrm{nr}}$ mainly depends on the fractional energy density in massive neutrinos (assuming $\Omega_c + \Omega_b$ is fixed) but also retains a slight sensitivity to the individual neutrino masses through dependence on a_{nr}.

An estimate of the effect of massive neutrinos on the growth of structure can be made by studying the evolution of matter perturbations in the two regimes $k \ll k_{\mathrm{fs}}$ and $k \gg k_{\mathrm{fs}}$. In the synchronous gauge, linear perturbations to the matter density with wavenumbers $k \ll k_{\mathrm{fs}}$ evolve as

$$\ddot{\delta}_m + 2H(a)\dot{\delta}_m - \frac{3}{2}\Omega_m H_0^2 a^{-3}\delta_m = 0 \quad \text{for} \quad k \ll k_{\mathrm{nr}} \tag{2.227}$$

which has solutions $\delta_m \propto a, a^{-\frac{3}{2}}$ during the matter dominated era.

On scales where the neutrino perturbations have decayed, perturbations to matter density are just in the CDM and baryon components

$$\delta_m(k \gg k_{\mathrm{fs}}) \approx (\delta\rho_c + \delta\rho_b)/\rho_m = (1 - f_\nu)\delta_{cb} \tag{2.228}$$

where $f_\nu = \Omega_\nu/\Omega_m$ and $\delta_{cb} = (\delta\rho_c + \delta\rho_b)/(\rho_c + \rho_b)$, but the neutrino energy density still contributes to the Hubble friction. In this limit, linear perturbations to the CDM and baryon density evolve as

$$\ddot{\delta}_m + 2H(a)\dot{\delta}_m - \frac{3}{2}\Omega_{cb}H_0^2 a^{-3}\delta_m = 0 \quad \text{for} \quad k \gg k_{\rm fs} \tag{2.229}$$

where $\Omega_{cb} = \Omega_c + \Omega_b$ and $\Omega_{cb} < \Omega_m$ for a cosmology with massive neutrinos. Equation (2.229) has the approximate solutions during the matter-dominated era of $\delta_{cb} \propto a^{1-\frac{3}{5}f_\nu}, a^{-\frac{3}{2}+\frac{3}{5}f_\nu}$ for $f_\nu \ll 1$ [LESGOURGUES and PASTOR, 2006].

The matter-dominated solutions provide a straightforward estimation of how massive neutrinos affect the amplitude of matter perturbations. In a cosmology with fixed $\Omega_c h^2$, the evolution of perturbations behaves similarly whether neutrinos have mass ($f_\nu \neq 0$) or are massless ($f_\nu = 0$) until the scale factor reaches $a_{\rm nr}$. After $a_{\rm nr}$, perturbations with wavenumbers $k \gg k_{\rm fs}$ exhibit slower growth (as indicated by Eq. (2.229), where the growing mode scales as $\propto a^{1-\frac{3}{5}f_\nu}$) compared to those with $k \ll k_{\rm fs}$ (as per Eq. (2.227), which scales as $\propto a$). Thus, at a given scale factor a during the matter-dominated era, the overall difference in the growth of perturbations with $k \gg k_{\rm fs}$ can be approximated as follows:

$$\frac{\delta_{cb}(k \gg k_{\rm fs}, a \mid f_\nu)}{\delta_{cb}(k \gg k_{\rm fs}, a \mid f_\nu = 0)} \sim \left(\frac{a}{a_{\rm nr}}\right)^{-\frac{3}{5}f_\nu} \tag{2.230}$$

The difference in the amplitude of the matter's power spectra is then

$$\frac{P_{mm}(k \gg k_{\rm fs}, a \mid f_\nu)}{P_{mm}(k \gg k_{\rm fs}, a \mid f_\nu = 0)} \sim (1 - 2f_\nu)\,\frac{P_{cb}(k \gg k_{\rm fs}, a \mid f_\nu)}{P_{cb}(k \gg k_{\rm fs}, a \mid f_\nu = 0)} \sim \left(1 - 2f_\nu - \frac{6}{5}f_\nu \ln\left(\frac{a}{a_{\rm nr}}\right)\right) \tag{2.231}$$

On the other hand, the evolution of the large-scale modes is identical,

$$\frac{P_{mm}(k \ll k_{\rm fs}, a \mid f_\nu)}{P_{mm}(k \ll k_{\rm fs}, a \mid f_\nu = 0)} = 1 \tag{2.232}$$

where P_{mm} represents the power spectrum of total matter fluctuations including CDM, neutrinos, and baryons, while P_{cb} denotes the power spectrum involving only CDM and baryons. The expression above tends to overestimate the impact of neutrino mass by assuming an instantaneous transition from relativistic to non-relativistic states. Furthermore, it neglects the influence of the cosmological constant at later epochs. Incorporating the actual evolution of δ_{cb} through $a_{\rm nr}$ and accounting for the cosmological constant leads to a more accurate assessment.

$$\frac{P_{mm}(k \gg k_{\rm fs} \mid f_\nu)}{P_{mm}(k \gg k_{\rm fs} \mid f_\nu = 0)} \approx 1 - 6f_\nu \tag{2.233}$$

at $a = 1$. Note that this expression assumes fixed $\Omega_c h^2, \Omega_b h^2$ so that matter-radiation equality is not changed by neutrino mass and that and $\Omega_\Lambda = 1 - \Omega_m$ is fixed by adjusting h so that the onset of cosmological constant domination is also unchanged. Alternatively, assuming fixed Ω_m and decreasing Ω_{cb} to account for Ω_ν makes matter-radiation equality, which occurs while the neutrinos are relativistic, slightly later so that the suppression is increased to

$$\frac{P_{mm}(k \gg k_{\rm fs} \mid f_\nu)}{P_{mm}(k \gg k_{\rm fs} \mid f_\nu = 0)} \approx (1 - 8f_\nu) \tag{2.234}$$

Chapter 3

Perturbation Theory and the Formation of Large-Scale Structures

One of the most sought-after questions in modern cosmology is understanding the large-scale structures of the universe we see today. Previously in section 2.7.4 we saw how quantum fluctuations during inflation that grow to super-horizon scales induce the primordial density perturbations. These perturbations are induced in the dark matter field once they re-enter the horizon. As they grow over the age of the universe, baryons accumulate in the gravitational potential wells of these dark matter density fluctuations eventually forming the structures we observe due to gravitational instabilities.

Cosmological Perturbation Theory (PT) aims at comprehending the dynamical evolution of the density and velocity fields of matter perturbations. This perturbative technique works especially well when characterizing the dynamics of gravitational instabilities at large scales, where density fluctuations are still tiny. On these large scales, structure creation is often thought to be primarily driven by gravity inside the PT framework. At small scales, however, this assumption falters as non-gravitational factors start to affect the distribution of matter for example the effects of the pressure forces of the baryons.

Furthermore, the predictive ability of PT is ultimately limited by the strong non-linear nature of gravitational evolution. At small scales ($k_{NL}(z = 0) \approx 0.25h/Mpc$ [Dodelson and Schmidt, 2020]), strong non-linearities and non-perturbative effects dominate the dynamics, making perturbative approaches ineffective. These scales are characterized by complex interactions and feedback mechanisms that require more advanced modelling techniques, frequently involving numerical simulations and semi-analytical methods.

However, there is a quasi non-linear regime in which the non-linear evolution of matter can still be adequately described by higher-order corrections to linear PT that could provide important information about the change from linear to fully non-linear scales. This approach allows the utility of PT beyond its strictly linear domain, bridging the gap between large-scale predictions and small-scale complexities.

Moreover, PT offers a framework for comprehending the statistical properties of matter distribution, like the power spectrum and bispectrum, which hold information about the universe and are crucial for restricting cosmological models and understanding data from extensive surveys.

PT operates under two main assumptions:

- On large scales baryonic pressure is neglected i.e. baryon is treated as cold as it cools rapidly after recombination and is approximated as collisionless. cold dark matter (CDM) and baryons are treated as pressureless single-field.

- Since CDM particles are non-relativistic, at scales smaller than the Hubble radius (r_H) the equations of motion reduce to those of Newtonian gravity.

In this chapter, we will discuss the Eulerian and Lagrangian framework of the standard PT and follow Ref. [Bernardeau et al., 2002] and references within it for our discussion.

3.1 Euler Perturbation Theory

3.1.1 The Vlasov Equation

When we consider gravitational instabilities in an expanding universe we need to account for the departures from the homogeneous universe. The positions are described by their comoving coordinates $\mathbf{x}$ ($\mathbf{r} = a(\tau)\mathbf{x}$ where r is the physical coordinate) and the equations of motion are described in terms conformal time τ. The conformal expansion rate is defined as $\mathscr{H} \equiv \mathrm{d}\ln a/\mathrm{d}\tau = Ha$ where H is the Hubble constant. We then define the density contrast $\delta(\mathbf{x})$ by,

$$\rho(\mathbf{x},\tau) \equiv \bar{\rho}(\tau)[1+\delta(\mathbf{x},\tau)], \tag{3.1}$$

the peculiar velocity $\mathbf{u}$ with

$$\mathbf{v}(\mathbf{x},\tau) \equiv \mathscr{H}\mathbf{x} + \mathbf{u}(\mathbf{x},\tau), \tag{3.2}$$

and the cosmological gravitational potential Φ with

$$\phi(\mathbf{x},\tau) \equiv -\frac{1}{2}\frac{\partial\mathscr{H}}{\partial\tau}x^2 + \Phi(\mathbf{x},\tau), \tag{3.3}$$

so that the latter is sourced only by density fluctuations, as expected. The Poisson equation reads,

$$\begin{aligned}\nabla^2\Phi(\mathbf{x},\tau) &= \frac{3}{2}\Omega_m(\tau)\mathscr{H}^2(\tau)\delta(\mathbf{x},\tau)\\ &= 4\pi G\bar{\rho}(\tau)\delta(\mathbf{x},\tau),\end{aligned} \tag{3.4}$$

In the following, the spatial variable will only be comoving coordinates, therefore all space derivatives should be interpreted as being performed with respect to $\mathbf{x}$. Then, the equation of motion can be written as,

$$\frac{\mathrm{d}\mathbf{p}}{\mathrm{d}\tau} = -am\nabla\Phi(\mathbf{x}), \tag{3.5}$$

where,

$$\mathbf{p} = am\mathbf{u}, \tag{3.6}$$

Now we define the Vlasov equation which is a a non-linear partial differential equation that describes the evolution of the particle number density in phase space. The particle number density in phase space is defined by the distribution function $f(\mathbf{x},\mathbf{p},\tau)$; phase-space conservation implies the Vlasov equation,

$$\frac{\mathrm{d}f}{\mathrm{d}\tau} = \frac{\partial f}{\partial\tau} + \frac{\mathbf{p}}{ma}\cdot\nabla f - am\nabla\Phi\cdot\frac{\partial f}{\partial\mathbf{p}} = 0, \tag{3.7}$$

The non-linearity is induced by the fact that the potential Φ depends through the Poisson equation on the integral of the distribution function over momentum (i.e. momentum average) (which gives the density field, see Eq.(3.8) below).

3.1.2 Eulerian Dynamics

In practice, we are usually interested in solving the evolution of the spatial distribution rather than the full phase-space dynamics. This is obtained by taking momentum moments of the distribution function.

$$\int \mathrm{d}^3\mathbf{p} f(\mathbf{x},\mathbf{p},\tau) \equiv \rho(\mathbf{x},\tau), \tag{3.8}$$

$$\int \mathrm{d}^3\mathbf{p} \frac{\mathbf{p}}{am} f(\mathbf{x},\mathbf{p},\tau) \equiv \rho(\mathbf{x},\tau)\mathbf{u}(\mathbf{x},\tau), \tag{3.9}$$

$$\int \mathrm{d}^3\mathbf{p} \frac{p_i p_j}{a^2 m^2} f(\mathbf{x},\mathbf{p},\tau) \equiv \rho(\mathbf{x},\tau)\mathbf{u}_i(\mathbf{x},\tau)\mathbf{u}_j(\mathbf{x},\tau) + \sigma_{ij}(\mathbf{x},\tau), \tag{3.10}$$

The zeroth order moment Eq.(3.8) relates the phase space density to the local mass density field, The higher order moments Eq.(3.9) and Eq.(3.10) define the peculiar velocity flow $\mathbf{u}(\mathbf{x},\tau)$ and the stress tensor $\sigma_{ij}(\mathbf{x},\tau)$. The equation for these fields follows from taking moments of the Vlasov equation. The zeroth moment gives the continuity equation describing the mass conservation,

$$\frac{\partial \delta(\mathbf{x},\tau)}{\partial \tau} + \nabla \cdot \{[1+\delta(\mathbf{x},\tau)]\mathbf{u}(\mathbf{x},\tau)\} = 0, \tag{3.11}$$

and subtracting $\mathbf{u}(\mathbf{x},\tau)$ times the continuity equation from the first moment gives the Euler equation which describes the conservation of momentum

$$\frac{\partial \mathbf{u}(\mathbf{x},\tau)}{\partial \tau} + \mathscr{H}(\tau)\mathbf{u}(\mathbf{x},\tau) + \mathbf{u}(\mathbf{x},\tau)\cdot\nabla\mathbf{u}(\mathbf{x},\tau) = - \nabla\Phi(\mathbf{x},\tau) - \frac{1}{\rho}\nabla_j\left(\rho\sigma_{ij}\right). \tag{3.12}$$

We can observe in the above equations that the continuity equation couples the zeroth (ρ) to the first moment ($\mathbf{u}$) of the distribution function similarly, the Euler equation couples the first moment ($\mathbf{u}$) to the second moment (σ_{ij}), and so on.
Now that the phase-space information has been integrated out, an ansatz is considered for the cosmological fluid's equation of state i.e. the stress tensor (σ_{ij}).

The first assumption that we considered at the beginning of the chapter is that the cosmological structure formation is driven by cold matter which has negligible pressure and velocity dispersion. The stress tensor in Eq.(3.10) characterizes the deviation of particle motions from a single coherent flow (single stream), for which the first term will be the dominant contribution. Therefore, it is a good approximation to set $\sigma_{ij} \approx 0$, at the initial stages of gravitational instability when structures did not have time to collapse and virialize. As time goes on, this approximation will break down at progressively larger scales, but at present times at the scales relevant to large-scale structure, this approximation is extremely handy. The breakdown describes the generation of velocity dispersion due to multiple streams, known as *shell crossing*

At large scales, when fluctuations are small, linear perturbation theory provides an adequate description of cosmological fields. In this regime, different Fourier modes evolve independently which conserves the primordial statistics. Therefore we combine the equations eq.(3.4), eq.(3.11) and eq.(3.12) and transform it into Fourier space to get the equation of motion for gravitational instabilities which are characterised by the overdensity $\tilde{\delta}(\mathbf{k},\tau)$ and velocity divergence $\tilde{\theta}(\mathbf{k},\tau)$. These are the continuity and Euler equations in Fourier space.

$$\frac{\partial\tilde{\delta}(\mathbf{k},\tau)}{\partial\tau}+\tilde{\theta}(\mathbf{k},\tau)=-\int \mathrm{d}^3\mathbf{k}_1\ \mathrm{d}^3\mathbf{k}_2\delta_D\left(\mathbf{k}-\mathbf{k}_{12}\right)\alpha\left(\mathbf{k}_1,\mathbf{k}_2\right)\tilde{\theta}\left(\mathbf{k}_1,\tau\right)\tilde{\delta}\left(\mathbf{k}_2,\tau\right), \tag{3.13}$$

$$\frac{\partial\tilde{\theta}(\mathbf{k},\tau)}{\partial\tau}+\mathscr{H}(\tau)\tilde{\theta}(\mathbf{k},\tau)+\frac{3}{2}\Omega_m\mathscr{H}^2(\tau)\tilde{\delta}(\mathbf{k},\tau)=-\int \mathrm{d}^3\mathbf{k}_1\ \mathrm{d}^3\mathbf{k}_2\delta_D\left(\mathbf{k}-\mathbf{k}_{12}\right)\times\beta\left(\mathbf{k}_1,\mathbf{k}_2\right)\tilde{\theta}\left(\mathbf{k}_1,\tau\right)\tilde{\theta}\left(\mathbf{k}_2,\tau\right), \tag{3.14}$$

(δ_D denotes the three-dimensional Dirac delta distribution) where $\mathbf{k}_{ij}=\mathbf{k}_i+\mathbf{k}_j$ the functions.The left-hand side of the above equations is the linear part, while the right-hand side contains the non-linear evolution that generates mode couplings, imprinted in functions

$$\alpha\left(\mathbf{k}_1,\mathbf{k}_2\right)\equiv\frac{\mathbf{k}_{12}\cdot\mathbf{k}_1}{k_1^2},\quad \beta\left(\mathbf{k}_1,\mathbf{k}_2\right)\equiv\frac{k_{12}^2\left(\mathbf{k}_1\cdot\mathbf{k}_2\right)}{2k_1^2k_2^2}. \tag{3.15}$$

Now, coming back to our assumption for the vanishing stress tensor we will try to investigate further the Poisson, continuity and Euler equations i.e. eq.(3.4), eq.(3.11) and eq.(3.12) respectively for the three unknown ρ, $\mathbf{u}$ and Φ. in the next section.

3.1.3 Linear Solution

At large scales[1], the fluctuation fields can be assumed to be small compared to the homogeneous contribution described by the first terms in eq.(3.11) and eq.(3.12). Therefore, it follows that we can linearize them to obtain the equations of motion in the linear regime.

$$\frac{\partial\delta(\mathbf{x},\tau)}{\partial\tau}+\theta(\mathbf{x},\tau)=0, \tag{3.16}$$

$$\frac{\partial\mathbf{u}(\mathbf{x},\tau)}{\partial\tau}+\mathscr{H}(\tau)\mathbf{u}(\mathbf{x},\tau)=-\nabla\Phi(\mathbf{x},\tau), \tag{3.17}$$

where $\theta(\mathbf{x},\tau)\equiv\nabla\cdot\mathbf{u}(\mathbf{x},\tau)$ is the divergence of the velocity field as mentioned before.
Now a velocity field, as any vector field, can be completely described by its divergence $\theta(\mathbf{x},\tau)$ and its vorticity $\mathbf{w}(\mathbf{x},\tau)\equiv\nabla\times\mathbf{u}(\mathbf{x},\tau)$, whose equations of motion follow from Eq.(3.17)

$$\frac{\partial\theta(\mathbf{x},\tau)}{\partial\tau}+\mathscr{H}(\tau)\theta(\mathbf{x},\tau)+\frac{3}{2}\Omega_m(\tau)\mathscr{H}^2(\tau)\delta(\mathbf{x},\tau)=0 \tag{3.18}$$

$$\frac{\partial\mathbf{w}(\mathbf{x},\tau)}{\partial\tau}+\mathscr{H}(\tau)\mathbf{w}(\mathbf{x},\tau)=0 \tag{3.19}$$

On solving Eq. (3.19) we see that vorticity evolution is given by, w $(\tau)\propto a^{-1}$, i.e. in the linear regime any initial vorticity decays away due to the expansion of the Universe. To get the density contrast evolution we will take the time derivative of Eq.(3.18) and replace it in Eq.(3.16) and use the Poisson equation Eq.(3.4) to get,

$$\frac{\mathrm{d}^2D_1(\tau)}{\mathrm{d}\tau^2}+\mathscr{H}(\tau)\frac{\mathrm{d}D_1(\tau)}{\mathrm{d}\tau}=\frac{3}{2}\Omega_m(\tau)\mathscr{H}^2(\tau)D_1(\tau) \tag{3.20}$$

where $\delta(\mathbf{x},\tau)=D_1(\tau)\delta^1(\mathbf{x},0)$, and $D_1(\tau)$ is the linear growth factor which describes the growth of the overdensity field from some reference time $\tau=0$ to some later time τ. The second term in the above equation is the friction term that arises due to the Hubble flow whereas the third term is the force term. Since it is a second-order differential equation, it has two independent solutions, let's denote the fastest growing mode $D_1^{(+)}(\tau)$ and the slowest one $D_1^{(-)}(\tau)$. The evolution of the density is then

$$\delta(\mathbf{x},\tau)=D_1^{(+)}(\tau)A(\mathbf{x})+D_1^{(-)}(\tau)B(\mathbf{x}), \tag{3.21}$$

[1]not to confuse with large scale modes

where $A(\mathbf{x})$ and $B(\mathbf{x})$ are two arbitrary functions of position describing the initial density field configuration, whereas the velocity divergence [using Eq.(3.16)] is given by

$$\theta(\mathbf{x},\tau) = -\mathscr{H}(\tau)\left[f\left(\Omega_m,\Omega_\Lambda\right)A(\mathbf{x}) + g\left(\Omega_m,\Omega_\Lambda\right)B(\mathbf{x})\right], \tag{3.22}$$

$$f\left(\Omega_m,\Omega_\Lambda\right) \equiv \frac{\mathrm{d}\ln D_1^{(+)}}{\mathrm{d}\ln a} = \frac{1}{\mathscr{H}}\frac{\mathrm{d}\ln D_1^{(+)}}{\mathrm{d}\tau} \quad g\left(\Omega_m,\Omega_\Lambda\right) = \frac{1}{\mathscr{H}}\frac{\mathrm{d}\ln D_1^{(-)}}{\mathrm{d}\tau}, \tag{3.23}$$

Let us discuss three different cases -

- The Einstein-de Sitter Universe i.e when $\Omega_m = 1, \Omega_\Lambda = 0$,

$$D_1^{(+)} = a, \quad D_1^{(-)} = a^{-3/2}, \quad f(1,0) = 1 \tag{3.24}$$

  We see that the density fluctuations grow as the scale factor which is consistent with our intuition that, as time evolves overdense regions attract more matter and become overdense

- When $\Omega_m < 1, \Omega_\Lambda = 0$ we have $(x \equiv 1/\Omega_m - 1)$.

$$D_1^{(+)} = 1 + \frac{3}{x} + 3\sqrt{\frac{1+x}{x^3}}\ln[\sqrt{1+x} - \sqrt{x}] \quad D_1^{(-)} = \sqrt{\frac{1+x}{x^3}}, \tag{3.25}$$

  and the logarithmic derivative can be approximated by [Peebles, 1976]

$$f\left(\Omega_m,0\right) \approx \Omega_m^{3/5}, \tag{3.26}$$

  As $\Omega_m \to 0 (x \gg 1), D_1^{(+)} \to 1$ and $D_1^{(-)} \to x^{-1}$ and perturbations cease to grow.

- Where there is only matter and vacuum energy

  In this case the linear growth factor admits the integral representation as a function of Ω_m and Ω_Λ

$$D_1^{(+)} = H(a)\frac{5\Omega_m}{2}\int_0^a \frac{\mathrm{d}a}{a^3 H(a)}, \tag{3.27}$$

So, we see that the growth factor describes the scale-independent growth at late times. Now, we would like to relate the potentials during this time to the primordial curvature perturbations. Using the Poisson equation, Eq. (3.4) in Fourier space, we relate the linear density contrast to the gravitational potential fluctuations :

$$k^2\Phi_{(\mathbf{k})} = 4\pi G\bar{\rho}\delta^1_{\mathbf{k}} = \frac{3}{2}\mathscr{H}^2\Omega_m\delta^1_{\mathbf{k}}. \tag{3.28}$$

Now the curvature perturbations can be related to the gravitational potential of the perturbed FLRW metric, for components with a barotropic equation of state, as:

$$\Phi(\mathbf{k}) = -\frac{3(1+w)}{(5+3w)}\zeta(\mathbf{k}), \tag{3.29}$$

Using Eq. (3.28) and (3.29) we get the following equation which relates over-densities to primordial potential.

$$\delta^1(\mathbf{k}) = -\frac{2(1+w)}{(5+3w)\Omega_m}\frac{k^2}{a^2H^2}T(k,\tau)D_1^{(+)}\zeta(\mathbf{k}). \tag{3.30}$$

where $D_1(\tau)$ is the growth factor discussed above and $T(k)$ is the transfer function that defines the evolution of perturbations across the horizon crossing and radiation/matter transition epochs. For the matter-dominated era, it reduces to [2]

[2]Notice we have changed to redshift(z) from scale factor

$$\delta^1_{(\mathbf{k})} = -\frac{2}{3}\frac{k^2}{\Omega_m H_0^2} T(k) D_1^{(+)} \Phi_{(\mathbf{k})} = M(\mathbf{k}, z)\Phi_{(\mathbf{k},z)} \tag{3.31}$$

3.1.4 Eulerian Non-Linear Perturbation Theory

We will now consider the density and velocity fields evolution beyond the linear approximation[3]. To do so, we shall first make a self-consistent approximation, that is, we will characterize the velocity field by its divergence, and justify that the vorticity degrees of freedom can be neglected.

Taking the curl of Eq. (3.12) and temporarily restoring the stress tensor contribution (σ_{ij}) to the conservation of momentum we can write the vorticity equation of motion as

$$\frac{\partial \mathbf{w}(\mathbf{x},\tau)}{\partial \tau} + \mathscr{H}(\tau)\mathbf{w}(\mathbf{x},\tau) - \nabla \times [\mathbf{u}(\mathbf{x},\tau) \times \mathbf{w}(\mathbf{x},\tau)] = \nabla \times \left(\frac{1}{\rho}\nabla \cdot \vec{\sigma}\right), \tag{3.32}$$

Let us now look into two cases. In the first case, if the primordial vorticity vanishes, we see that $\sigma_{ij} \approx 0$, as in the case of a pressureless perfect fluid. So, in the absence of primordial vorticity σ_{ij} remains zero at all times. In the next case, if the initial vorticity is non-zero, we saw in the previous section that in the linear regime vorticity decays due to the expansion of the Universe; however, it can be amplified non-linearly through the third term in Eq. (3.31). In what follows, we shall assume that the initial vorticity vanishes, thus Eq. (3.31) together with the equation of state $\sigma_{ij} \approx 0$ guarantees that vorticity remains zero throughout the evolution.

PT assumes that it is possible to expand the density and velocity fields about the linear solutions by effectively treating the variance of the linear fluctuations as a small parameter and assuming no vorticity in the velocity field [[Dodelson and Schmidt, 2020][Bernardeau et al., 2002]]. Linear solutions correspond to simple (time-dependent) scaling of the initial density field; which we can write as

$$\delta(\mathbf{x},t) = \sum_{n=1}^{\infty} \delta^{(n)}(\mathbf{x},t), \quad \theta(\mathbf{x},t) = \sum_{n=1}^{\infty} \theta^{(n)}(\mathbf{x},t), \tag{3.33}$$

where $\delta^{(1)} = D_1(\tau)\delta(\mathbf{x},0)$ and $\theta^{(1)} = -\mathscr{H}\delta^{(1)}$ are linear in the initial density field as we saw before, $\delta^{(2)}$ and $\theta^{(2)}$ are quadratic in the initial density field, etc.

3.1.5 General Solutions in Einstein-de Sitter Cosmology

Let's first consider an Einstein-de Sitter Universe, for which $\Omega_m = 1$ and $\Omega_\Lambda = 0$ and scaling out an overall factor of $\mathscr{H}$ from the velocity field brings Eq. (3.13) and Eq. (3.14) into homogeneous form in τ or, equivalently, in $a(\tau)$. As a consequence, these equations can formally be solved with the following perturbative expansion [Goroff et al., 1986, Jain and Bertschinger, 1994],

$$\tilde{\delta}(\mathbf{k},\tau) = \sum_{n=1}^{\infty} a^n(\tau)\delta_n(\mathbf{k}), \quad \tilde{\theta}(\mathbf{k},\tau) = -\mathscr{H}(\tau)\sum_{n=1}^{\infty} a^n(\tau)\theta_n(\mathbf{k}), \tag{3.34}$$

where only the fastest-growing mode is taken into account. Remarkably, it implies that the PT expansions defined in Eq. (3.32) are expansions with respect to the linear density field with time-independent coefficients. At small a the series are dominated by their first term, and since $\theta_1(\mathbf{k}) = \delta_1(\mathbf{k})$ from the continuity equation, $\delta_1(\mathbf{k})$ completely characterizes the linear fluctuations.

The equations of motion, Eqs. (37-38) determine $\delta_n(\mathbf{k})$ and $\theta_n(\mathbf{k})$ in terms of the linear fluctuations to be:

$$\delta_n(\mathbf{k}) = \int \mathrm{d}^3\mathbf{q}_1 \ldots \int \mathrm{d}^3\mathbf{q}_n \delta_D\left(\mathbf{k} - \mathbf{q}_{1\ldots n}\right) F_n\left(\mathbf{q}_1, \ldots, \mathbf{q}_n\right) \delta_1\left(\mathbf{q}_1\right) \ldots \delta_1\left(\mathbf{q}_n\right), \tag{3.35}$$

$$\theta_n(\mathbf{k}) = \int \mathrm{d}^3\mathbf{q}_1 \ldots \int \mathrm{d}^3\mathbf{q}_n \delta_D\left(\mathbf{k} - \mathbf{q}_{1\ldots n}\right) G_n\left(\mathbf{q}_1, \ldots, \mathbf{q}_n\right) \delta_1\left(\mathbf{q}_1\right) \ldots \delta_1\left(\mathbf{q}_n\right), \tag{3.36}$$

[3]Given the smallness of the spacetime perturbations we can continue to work to linear order in the potential

here F_n and G_n are homogeneous functions of the wave vectors $\{\mathbf{q}_1, \ldots, \mathbf{q}_n\}$ with degree zero. They incorporate the non-linear mode coupling induced by gravity and are constructed from the fundamental mode coupling functions $\alpha(\mathbf{k}_1, \mathbf{k}_2)$ and $\beta(\mathbf{k}_1, \mathbf{k}_2)$ according to the recursion relations ($n \geq 2$, see [Goroff et al., 1986, Jain and Bertschinger, 1994] for a detailed derivation):

$$F_n(\mathbf{q}_1, \ldots, \mathbf{q}_n) = \sum_{m=1}^{n-1} \frac{G_m(\mathbf{q}_1, \ldots, \mathbf{q}_m)}{(2n+3)(n-1)} \left[(2n+1)\alpha(\mathbf{k}_1, \mathbf{k}_2) F_{n-m}(\mathbf{q}_{m+1}, \ldots, \mathbf{q}_n) \right.$$
$$\left. + 2\beta(\mathbf{k}_1, \mathbf{k}_2) G_{n-m}(\mathbf{q}_{m+1}, \ldots, \mathbf{q}_n)\right], \tag{3.37}$$

$$G_n(\mathbf{q}_1, \ldots, \mathbf{q}_n) = \sum_{m=1}^{n-1} \frac{G_m(\mathbf{q}_1, \ldots, \mathbf{q}_m)}{(2n+3)(n-1)} \left[3\alpha(\mathbf{k}_1, \mathbf{k}_2) F_{n-m}(\mathbf{q}_{m+1}, \ldots, \mathbf{q}_n) \right.$$
$$\left. + 2n\beta(\mathbf{k}_1, \mathbf{k}_2) G_{n-m}(\mathbf{q}_{m+1}, \ldots, \mathbf{q}_n)\right], \tag{3.38}$$

where $\mathbf{k}_1 \equiv \mathbf{q}_1 + \ldots + \mathbf{q}_m, \mathbf{k}_2 \equiv \mathbf{q}_{m+1} + \ldots + \mathbf{q}_n, \mathbf{k} \equiv \mathbf{k}_1 + \mathbf{k}_2$. Up to third order, the symmetrized kernels (i.e. sum of the n-th order kernels over all possible permutations of the modes) for the density field are:

$$F_1(\mathbf{k}) = 1, \tag{3.39}$$

$$F_2^{(s)}(\mathbf{k}_1, \mathbf{k}_2) = \frac{5}{7} + \frac{1}{2}\frac{\mathbf{k}_1 \cdot \mathbf{k}_2}{k_1 k_2}\left(\frac{k_1}{k_2} + \frac{k_2}{k_1}\right) + \frac{2}{7}\left(\frac{\mathbf{k}_1 \cdot \mathbf{k}_2}{k_1 k_2}\right)^2, \tag{3.40}$$

$$F_3^{(s)}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) = \frac{7}{54}\Big[F_2^{(s)}(\mathbf{k}_1, \mathbf{k}_2)\,\alpha(\mathbf{k}_3, \mathbf{k}_{12}) + F_2^{(s)}(\mathbf{k}_2, \mathbf{k}_3)\,\alpha(\mathbf{k}_1, \mathbf{k}_{23}) + F_2^{(s)}(\mathbf{k}_3, \mathbf{k}_1)\,\alpha(\mathbf{k}_2, \mathbf{k}_{31})$$
$$+ G_2^{(s)}(\mathbf{k}_1, \mathbf{k}_2)\,\alpha(\mathbf{k}_{12}, \mathbf{k}_3) + G_2^{(s)}(\mathbf{k}_2, \mathbf{k}_3)\,\alpha(\mathbf{k}_{23}, \mathbf{k}_1) + G_2^{(s)}(\mathbf{k}_3, \mathbf{k}_1)\,\alpha(\mathbf{k}_{31}, \mathbf{k}_2)$$
$$+ \frac{2}{27}\Big[G_2^{(s)}(\mathbf{k}_1, \mathbf{k}_2)\,\beta(\mathbf{k}_{12}, \mathbf{k}_3) + G_2^{(s)}(\mathbf{k}_2, \mathbf{k}_3)\,\beta(\mathbf{k}_{23}, \mathbf{k}_1) + G_2^{(s)}(\mathbf{k}_3, \mathbf{k}_1)\,\beta(\mathbf{k}_{31}, \mathbf{k}_2)\Big], \tag{3.41}$$

$$F_4^{(s)}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3, \mathbf{k}_4) = \frac{1}{792}\Big[2G_2^{(s)}(\mathbf{k}_1, \mathbf{k}_4)$$
$$\Big(18F_2^{(s)}(\mathbf{k}_2, \mathbf{k}_3)\,\alpha(\mathbf{k}_1 + \mathbf{k}_4, \mathbf{k}_2 + \mathbf{k}_3) + 8G_2^{(s)}(\mathbf{k}_2, \mathbf{k}_3)\,\beta(\mathbf{k}_2 + \mathbf{k}_3, \mathbf{k}_1 + \mathbf{k}_4)\Big)$$
$$+ 2G_2^{(s)}(\mathbf{k}_1, \mathbf{k}_3)\Big(18F_2^{(s)}(\mathbf{k}_2, \mathbf{k}_4)\,\alpha(\mathbf{k}_1 + \mathbf{k}_3, \mathbf{k}_2 + \mathbf{k}_4) + 8G_2^{(s)}(\mathbf{k}_2, \mathbf{k}_4)\,\beta(\mathbf{k}_1 + \mathbf{k}_3, \mathbf{k}_2 + \mathbf{k}_4)\Big)$$
$$+ 2G_2^{(s)}(\mathbf{k}_1, \mathbf{k}_2)\Big(18F_2^{(s)}(\mathbf{k}_3, \mathbf{k}_4)\,\alpha(\mathbf{k}_1 + \mathbf{k}_2, \mathbf{k}_3 + \mathbf{k}_4) + 8G_2^{(s)}(\mathbf{k}_3, \mathbf{k}_4)\,\beta(\mathbf{k}_1 + \mathbf{k}_2, \mathbf{k}_3 + \mathbf{k}_4)\Big)$$
$$+ 6G_3^{(s)}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3)\,(9\alpha(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3, \mathbf{k}_4) + 4\beta(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3, \mathbf{k}_4))$$
$$+ 6G_3^{(s)}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_4)\,(9\alpha(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_4, \mathbf{k}_3) + 4\beta(\mathbf{k}_3, \mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_4))$$
$$+ 6G_3^{(s)}(\mathbf{k}_1, \mathbf{k}_3, \mathbf{k}_4)\,(9\alpha(\mathbf{k}_1 + \mathbf{k}_3 + \mathbf{k}_4, \mathbf{k}_2) + 4\beta(\mathbf{k}_2, \mathbf{k}_1 + \mathbf{k}_3 + \mathbf{k}_4))$$
$$+ 6G_3^{(s)}(\mathbf{k}_2, \mathbf{k}_3, \mathbf{k}_4)\,(9\alpha(\mathbf{k}_2 + \mathbf{k}_3 + \mathbf{k}_4, \mathbf{k}_1) + 4\beta(\mathbf{k}_1, \mathbf{k}_2 + \mathbf{k}_3 + \mathbf{k}_4))$$
$$+ 36F_2^{(s)}(\mathbf{k}_1, \mathbf{k}_4)\,G_2^{(s)}(\mathbf{k}_2, \mathbf{k}_3)\,\alpha(\mathbf{k}_2 + \mathbf{k}_3, \mathbf{k}_1 + \mathbf{k}_4)$$
$$+ 36F_2^{(s)}(\mathbf{k}_1, \mathbf{k}_3)\,G_2^{(s)}(\mathbf{k}_2, \mathbf{k}_4)\,\alpha(\mathbf{k}_2 + \mathbf{k}_4, \mathbf{k}_1 + \mathbf{k}_3)$$
$$+ 36F_2^{(s)}(\mathbf{k}_1, \mathbf{k}_2)\,G_2^{(s)}(\mathbf{k}_3, \mathbf{k}_4)\,\alpha(\mathbf{k}_3 + \mathbf{k}_4, \mathbf{k}_1 + \mathbf{k}_2)$$
$$+ 54F_3^{(s)}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3)\,\alpha(\mathbf{k}_4, \mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3) + 54F_3^{(s)}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_4)\,\alpha(\mathbf{k}_3, \mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_4)$$
$$+ 54F_3^{(s)}(\mathbf{k}_1, \mathbf{k}_3, \mathbf{k}_4)\,\alpha(\mathbf{k}_2, \mathbf{k}_1 + \mathbf{k}_3 + \mathbf{k}_4) + 54F_3^{(s)}(\mathbf{k}_2, \mathbf{k}_3, \mathbf{k}_4)\,\alpha(\mathbf{k}_1, \mathbf{k}_2 + \mathbf{k}_3 + \mathbf{k}_4)\Big], \tag{3.42}$$

while for the velocity divergence field they are:

$$G_1(\mathbf{k}) = 1, \tag{3.43}$$

$$G_2^{(s)}(\mathbf{k}_1, \mathbf{k}_2) = \frac{3}{7} + \frac{1}{2}\frac{\mathbf{k}_1 \cdot \mathbf{k}_2}{k_1 k_2}\left(\frac{k_1}{k_2} + \frac{k_2}{k_1}\right) + \frac{4}{7}\left(\frac{\mathbf{k}_1 \cdot \mathbf{k}_2}{k_1 k_2}\right)^2, \tag{3.44}$$

$$\begin{aligned} G_3^{(s)}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) = \frac{1}{18}\Big[& F_2^{(s)}(\mathbf{k}_1, \mathbf{k}_2)\,\alpha(\mathbf{k}_3, \mathbf{k}_{12}) + F_2^{(s)}(\mathbf{k}_2, \mathbf{k}_3)\,\alpha(\mathbf{k}_1, \mathbf{k}_{23}) + F_2^{(s)}(\mathbf{k}_3, \mathbf{k}_1)\,\alpha(\mathbf{k}_2, \mathbf{k}_{31}) \\ & + G_2^{(s)}(\mathbf{k}_1, \mathbf{k}_2)\,\alpha(\mathbf{k}_{12}, \mathbf{k}_3) + G_2^{(s)}(\mathbf{k}_2, \mathbf{k}_3)\,\alpha(\mathbf{k}_{23}, \mathbf{k}_1) + G_2^{(s)}(\mathbf{k}_3, \mathbf{k}_1)\,\alpha(\mathbf{k}_{31}, \mathbf{k}_2)\Big] \\ & + \frac{2}{9}\Big[G_2^{(s)}(\mathbf{k}_1, \mathbf{k}_2)\,\beta(\mathbf{k}_{12}, \mathbf{k}_3) + G_2^{(s)}(\mathbf{k}_2, \mathbf{k}_3)\,\beta(\mathbf{k}_{23}, \mathbf{k}_1) + G_2^{(s)}(\mathbf{k}_3, \mathbf{k}_1)\,\beta(\mathbf{k}_{31}, \mathbf{k}_2)\Big], \end{aligned} \tag{3.45}$$

$$\begin{aligned} G_4^{(s)}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3, \mathbf{k}_4) = \frac{1}{792}\Big[& 2G_2^{(s)}(\mathbf{k}_1, \mathbf{k}_4) \\ & \Big(6F_2^{(s)}(\mathbf{k}_2, \mathbf{k}_3)\,\alpha(\mathbf{k}_1 + \mathbf{k}_4, \mathbf{k}_2 + \mathbf{k}_3) + 32G_2^{(s)}(\mathbf{k}_2, \mathbf{k}_3)\,\beta(\mathbf{k}_2 + \mathbf{k}_3, \mathbf{k}_1 + \mathbf{k}_4)\Big) \\ & + 2G_2^{(s)}(\mathbf{k}_1, \mathbf{k}_3)\Big(6F_2^{(s)}(\mathbf{k}_2, \mathbf{k}_4)\,\alpha(\mathbf{k}_1 + \mathbf{k}_3, \mathbf{k}_2 + \mathbf{k}_4) + 32G_2^{(s)}(\mathbf{k}_2, \mathbf{k}_4)\,\beta(\mathbf{k}_1 + \mathbf{k}_3, \mathbf{k}_2 + \mathbf{k}_4)\Big) \\ & + 2G_2^{(s)}(\mathbf{k}_1, \mathbf{k}_2)\Big(6F_2^{(s)}(\mathbf{k}_3, \mathbf{k}_4)\,\alpha(\mathbf{k}_1 + \mathbf{k}_2, \mathbf{k}_3 + \mathbf{k}_4) + 32G_2^{(s)}(\mathbf{k}_3, \mathbf{k}_4)\,\beta(\mathbf{k}_1 + \mathbf{k}_2, \mathbf{k}_3 + \mathbf{k}_4)\Big) \\ & + 6G_3^{(s)}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3)\,(3\alpha(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3, \mathbf{k}_4) + 16\beta(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3, \mathbf{k}_4)) \\ & + 6G_3^{(s)}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_4)\,(3\alpha(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_4, \mathbf{k}_3) + 16\beta(\mathbf{k}_3, \mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_4)) \\ & + 6G_3^{(s)}(\mathbf{k}_1, \mathbf{k}_3, \mathbf{k}_4)\,(3\alpha(\mathbf{k}_1 + \mathbf{k}_3 + \mathbf{k}_4, \mathbf{k}_2) + 16\beta(\mathbf{k}_2, \mathbf{k}_1 + \mathbf{k}_3 + \mathbf{k}_4)) \\ & + 6G_3^{(s)}(\mathbf{k}_2, \mathbf{k}_3, \mathbf{k}_4)\,(3\alpha(\mathbf{k}_2 + \mathbf{k}_3 + \mathbf{k}_4, \mathbf{k}_1) + 16\beta(\mathbf{k}_1, \mathbf{k}_2 + \mathbf{k}_3 + \mathbf{k}_4)) \\ & + 12F_2^{(s)}(\mathbf{k}_1, \mathbf{k}_4)\,G_2^{(s)}(\mathbf{k}_2, \mathbf{k}_3)\,\alpha(\mathbf{k}_2 + \mathbf{k}_3, \mathbf{k}_1 + \mathbf{k}_4) \\ & + 12F_2^{(s)}(\mathbf{k}_1, \mathbf{k}_3)\,G_2^{(s)}(\mathbf{k}_2, \mathbf{k}_4)\,\alpha(\mathbf{k}_2 + \mathbf{k}_4, \mathbf{k}_1 + \mathbf{k}_3) \\ & + 12F_2^{(s)}(\mathbf{k}_1, \mathbf{k}_2)\,G_2^{(s)}(\mathbf{k}_3, \mathbf{k}_4)\,\alpha(\mathbf{k}_3 + \mathbf{k}_4, \mathbf{k}_1 + \mathbf{k}_2) \\ & + 18F_3^{(s)}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3)\,\alpha(\mathbf{k}_4, \mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3) + 18F_3^{(s)}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_4)\,\alpha(\mathbf{k}_3, \mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_4) \\ & + 18F_3^{(s)}(\mathbf{k}_1, \mathbf{k}_3, \mathbf{k}_4)\,\alpha(\mathbf{k}_2, \mathbf{k}_1 + \mathbf{k}_3 + \mathbf{k}_4) + 18F_3^{(s)}(\mathbf{k}_2, \mathbf{k}_3, \mathbf{k}_4)\,\alpha(\mathbf{k}_1, \mathbf{k}_2 + \mathbf{k}_3 + \mathbf{k}_4)\Big], \end{aligned} \tag{3.46}$$

From above we can see that the first order is just the linear solution and hence $F_1(\mathbf{k}) = 1$ and $G_1(\mathbf{k}) = 1$. In an Einstein-de-Sitter Universe (where, f = 1), the n-th order growth factor is $D_n = D_1^n$. The difference in the above equations between ΛCDM and Einstein-de-Sitter Universe is a factor of Ω_m/f^2 and its value is very close to unity [Bernardeau et al., 2002].
Since the growth rate in ΛCDM is found to be $f \approx \Omega_m^{0.59}$ therefore, we can safely take all the higher order growth factors in ΛCDM to be $D_n = D_1^n$ and use the kernel results of Einstein-de-Sitter Universe.

So, the equations Eq. (3.34) and Eq. (3.35) allow us to calculate how the universe evolves non-linearly very neatly. An intuitive representation of the perturbative expansion which is analogous to Feynman's diagram of quantum field theory can be seen in Fig 3.1.

The second order density field $\delta^{(2)}$ is connected by joining two instances of initial (linear) density field with an F_2 kernel. Similarly, the n_{th} order field is made by joining n initial density fields with the n_{th} order kernel F_n. The analogous rules hold for the expansion of the velocity divergence.

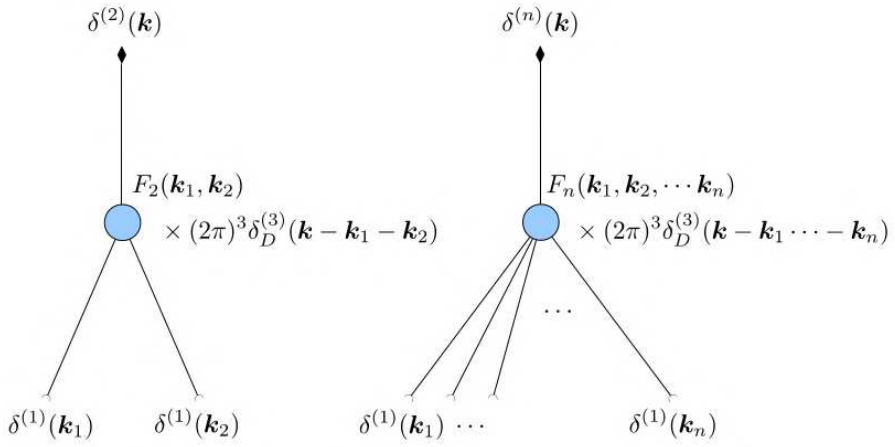

Figure 3.1: Diagrammatic representation of the second-order density field $\delta^{(2)}$ (left) and the n^{th} order density field(right). In each case, the final density field is connected to n initial density fields by the interaction kernel F_n (with n = 2 in the case of $\delta^{(2)}$). Analogous diagrams describe the velocity divergence $\theta(n)$ in terms of kernelsG_n. Here the time arguments are suppressed for clarity.[Dodelson and Schmidt, 2020]

3.2 Lagrangian Perturbation theory

3.2.1 Lagrangian Dynamics

We saw that the Eulerian PT describes the evolution of the density and velocity fields from a fixed comoving coordinate system. Another way to formulate non-linear perturbation theory is by using Lagrangian Perturbation Theory (LPT) which follows the trajectories of the fluid elements [Zel'dovich, 1970, Buchert, 1989, Bouchet, 1996]. The idea is to parameterise each particle in the cosmic fluid instead of taking the Lagrangian of all the particles. The displacement field $\Psi(\mathbf{q})$ is the dynamical parameter in this formalism. It connects the initial Lagrangian positions $\mathbf{q}$ of the fluid elements, with the final Eulerian positions $\mathbf{x}$. which is defined as:

$$\mathbf{x}(\tau) = \mathbf{q} + \mathbf{\Psi}(\mathbf{q}, \tau) \tag{3.47}$$

We can write the equation of motion for the particle's trajectory as

$$\frac{d^2\mathbf{x}(\tau)}{d\tau^2} + \mathscr{H}\frac{d\mathbf{x}}{d\tau} = \nabla_{\mathbf{x}}\Phi(\mathbf{x}), \tag{3.48}$$

where again Φ is the gravitational potential fluctuations and the subscript in the gradient subscript denotes the Eulerian position. The divergence of the above equation gives

$$J(\mathbf{q}, \tau)\boldsymbol{\nabla}_{\mathbf{x}}\left[\frac{d^2\mathbf{\Psi}}{d\tau^2} + \mathscr{H}\frac{d\mathbf{\Psi}}{d\tau}\right] = \frac{3}{2}\Omega_m\mathscr{H}^2(J-1), \tag{3.49}$$

Where we have used the Poisson equation Eq. (3.4) and conservation of particles in an infinitesimal volume between the two frames, i.e. $\bar{\rho}d^3q = \rho(\mathbf{x}, \tau)d^x = \bar{\rho}[1 + \delta(\mathbf{x})]d^3x$. The Jacobian transformation that connects the two frames is given by

$$Jd^3q = d^3x \Rightarrow J = \det\left|\frac{\partial x_i}{\partial q_i}\right| = \det\left|\delta_{ij} + \frac{\partial \Psi_i}{\partial q_j}\right|, \tag{3.50}$$

which gives

$$1 + \delta(\mathbf{x}, \tau) = \frac{1}{J(\mathbf{q}, \tau)}, \tag{3.51}$$

Using the chain rule on the divergence in Eulerian space [4], we can get the final equation of the displacement field as

$$[\delta_{ij} + \Psi_{i,j}(\mathbf{q},\tau)]^{-1}\left[\frac{d^2\Psi_{i,j}(\mathbf{q},\tau)}{d\tau^2} + \mathscr{H}\frac{d\Psi_{i,j}(\mathbf{q},\tau)}{d\tau}\right] = \frac{3}{2}\Omega_m\mathscr{H}^2\frac{J(\mathbf{q},\tau)-1}{J(\mathbf{q},\tau)} \tag{3.52}$$

where $\Psi_{i,j} = \frac{\partial\Psi_i}{\partial q_j}$. This approach, however, breaks down at shell crossing, a phenomenon that occurs when particles come very close to each other, eventually acquiring the same Eulerian coordinate $\mathbf{x}$ despite originating from different Lagrangian points $\mathbf{q}$. This convergence happens over time due to the time-dependent relationship between the Eulerian and Lagrangian coordinates Eq. (3.46). The Jacobian can be expanded as

$$\begin{aligned}
J &= \det\left(\delta_{ij} + \frac{\partial\Psi_i}{\partial q_j}\right) \\
&= 1 + \boldsymbol{\nabla}_{\mathbf{q}}\cdot\Psi(\mathbf{q},\tau) \\
&\quad + \frac{1}{2}\left[(\boldsymbol{\nabla}_{\mathbf{q}}\cdot\Psi(\mathbf{q},\tau))^2 - \sum_{i,j}\Psi_{i,j}\Psi_{j,i}\right] \\
&\quad + \frac{1}{6}\left[(\boldsymbol{\nabla}_{\mathbf{q}}\cdot\boldsymbol{\Psi}(\mathbf{q},\tau))^3 - 3\boldsymbol{\nabla}_{\mathbf{q}}\cdot\boldsymbol{\Psi}(\mathbf{q},\tau)\sum_{i,j}\Psi_{i,j}\Psi_{j,i}\right. \\
&\quad \left. +2\sum_{i,j,k}\Psi_{i,j}\Psi_{j,k}\Psi_{k,i}\right] + \dots
\end{aligned}$$

while the EOM of the fluid trajectory can be solved perturbatively as

$$\Psi(\mathbf{q},\tau) = \boldsymbol{\Psi}^{(1)}(\mathbf{q},\tau) + \Psi^{(2)}(\mathbf{q},\tau) + \dots \tag{3.53}$$

3.2.2 Linear Solution

The linear part of the Jacobian, i.e. the first two terms of Eq. (3.35), are used in this first approximation. The inverse of the Jacobian transformation matrix will be now

$$[\delta_{ij} + \Psi_{i,j}(\mathbf{q},\tau)]^{-1} \simeq \delta_{ij} - \Psi_{i,i}, \tag{3.54}$$ [5]

Using, Eq. (3.54) and the linear part of the Jacobian expansion, we can derive the linear solution as:

$$\begin{aligned}
&1 + \delta^{(1)}(\mathbf{x},\tau) = \frac{1}{J(\mathbf{q},\tau)} \simeq 1 - \nabla_{\mathbf{q}}\cdot\Psi(\mathbf{q},\tau) \Rightarrow \\
&\nabla_{\mathbf{q}}\cdot\Psi^{(1)}(\mathbf{q},\tau) = -D_1^{(+)}(\tau)\delta^{(1)}(\mathbf{q}),
\end{aligned} \tag{3.55}$$

where we split the time part from the spatial part in the growing linear solution. The linear density field $\delta^{(1)}(\mathbf{q})$ is the initial condition field, which evolves with the linear growth factor under the divergence of the displacement field. The linear growth factor in LPT is the same as in Eulerian PT. Therefore, in the first order, the position of the particle in Eulerian space can be written using Eq. (3.46) and Eq. (3.56) as :

[4]
$$\frac{\partial}{\partial x_i} = \frac{\partial q_i}{\partial x_i}\frac{\partial}{\partial q_i} = \left(\delta_{ij} + \frac{\partial\Psi_i}{\partial q_j}\right)^{-1}\frac{\partial}{\partial q_j}$$

[5] using $\det(I + A) = 1 + \mathrm{tr}(A) + \mathscr{O}\left(A^2\right) I$

$$\mathbf{x}(\tau) = \mathbf{q} - \nabla_{\mathbf{q}}^{-1}\delta^{(1)}(\mathbf{x},\tau), \tag{3.56}$$

while the velocity field is given by

$$\mathbf{u}(\tau)\frac{d\mathbf{x}}{d\tau} = -f\mathscr{H}\nabla_{\mathbf{q}}^{-1}\delta^{(1)}(\mathbf{x},\tau), \tag{3.57}$$

where f is the linear growth rate same as in the Eulerian PT. For the Poisson equation, we can relate the displacement field divergence with the gravitational potential in Lagrangian space, under the assumption of an irrotational gravity field, as

$$\nabla_{\mathbf{q}} \cdot \Psi^{(1)}(\mathbf{q},\tau) = -\nabla_{\mathbf{q}}^2\Phi^{(1)}(\mathbf{q},\tau) = -\delta^{(1)}(\mathbf{x},\tau), \tag{3.58}$$

which leads to

$$\Psi^{(1)}(\mathbf{q},\tau) = -\nabla_{\mathbf{q}}\Phi(\mathbf{q},\tau). \tag{3.59}$$

Zel'dovich [Zel'dovich, 1970] used the linear solution and tried to approximate the dynamical equation by extrapolating it into the non-linear regime which is known as the (Zel'dovich approximation (ZA). This was done by exchanging the divergence of the displacement field with the tidal tensor (traceless part). From Eq. (3.50) we get

$$\rho(\mathbf{x},\tau) = \frac{\bar{\rho}(\tau)}{\det\left|\delta_{ij} + \frac{\partial\Psi_i}{\partial q_j}\right|} = \frac{\bar{\rho}}{\left|(1-\lambda_1 D_1(\tau))\,(1-\lambda_2 D_1(\tau))\,(1-\lambda_3 D_1(\tau))\right|}, \tag{3.60}$$

where the variables λ_i are the eigenvalues of the tidal tensor field $\Psi_{i,j}$. The power of this result lies on the fact that, we can predict the future of a collapsing region (i.e. $(1-\lambda D_1(\tau)) = 0$) and determine the structure this point belongs to. If the eigenvalues are all positive, while one of them is larger than the rest (i.e. $\lambda_1 > \lambda_2, \lambda_3$), we get a pancake shape. This indicates that the element collapses in one direction. However, the ZA breaks down before the point reaches the final steps of collapse. A spherical collapse occurs in the case where all eigenvalues are positive, but now equal in size. If two of them are positive and one negative (i.e. $\lambda_1, \lambda_2 > 0, 0 < \lambda_3$), then the collapse happens in 2 dimensions and therefore the point belongs to a filament. In the case where two of them are negative and only one is positive (i.e. $\lambda_1, \lambda_2 < 0, \lambda_3 > 0$) the element belongs to a wall. Finally, negative eigenvalues correspond to a growing mode, which indicates that the point belongs to a void.

3.2.3 Second-order solution

For LPT there is no known recursive solution for the expression of higher order like in Eulerian PT. The solution must be performed order by order. The second-order Lagrangian PT (2LPT) can improve significantly the predictions, for the density and velocity fields over the linear solution.This is because the second-order solution to Eq. (3.51) accounts for the correction to the Zel'dovich approximation (ZA) displacement due to tidal effects [Buchert et al., 1994, Melott et al., 1995, Bouchet, 1996, Bernardeau et al., 2002]. Now, we will use Eq. (3.53) and consider terms up to second order. We will then substitute these terms into the equation of motion, Eq. (3.51), to get

$$\frac{d^2\Psi_{i,i}^{(2)}}{d\tau^2} + \mathscr{H}\frac{d\Psi_{i,i}^{(2)}}{d\tau} - \frac{3}{2}\mathscr{H}^2\Omega_m\Psi_{i,i}^{(2)} = -\frac{3}{4}\left[\left(\Psi_{k,k}^{(2)}\right)^2 - \Psi_{i,j}^{(2)}\Psi_{j,i}^{(2)}\right], \tag{3.61}$$

where the linear solution of the displacement field has been also used in the above derivation. Separating, as before, the second-order solution into a time and a spatial part, we get

$$\Psi_{k,k}^{(2)}(\mathbf{q},\tau) = \frac{D_2(\tau)}{2D_1^2(\tau)}\sum_{i\neq j}\left(\Psi_{i,i}^{(1)}(\mathbf{q},\tau)\Psi_{j,j}^{(1)}(\mathbf{q},\tau) - \Psi_{i,j}^{(1)}(\mathbf{q},\tau)\Psi_{j,i}^{(1)}(\mathbf{q},\tau)\right), \tag{3.62}$$

where the time-dependent part of $\Psi^{(2)}$ is denoted as $D_2(\tau)$ which is the second-order growth factor. It has been shown that in ΛCDM cosmology[Bouchet, 1996] ,

$$D_2(\tau) \simeq \frac{-3D_1^2(\tau)\Omega_m^{-1/143}}{7}. \tag{3.63}$$

By using the Poisson equation together with the displacement field relation at second order (i.e. $\Psi^{(2)}(\mathbf{q},\tau) = \boldsymbol{\nabla}_{\mathbf{q}}\Phi^{(2)}(\mathbf{q},\tau)\big)$ We can further simplify the second-order result as,

$$\nabla_{\mathbf{q}}\Phi^{(2)}(\mathbf{q},\tau) \simeq -\frac{3}{7}\Omega_m^{-1/143}\sum_{i>j}\left(\Phi^{(1)}_{,ii}(\mathbf{q},\tau)\Phi^{(1)}_{,jj}(\mathbf{q},\tau) - \left[\Phi^{(1)}_{,ij}(\mathbf{q},\tau)\right]^2\right), \tag{3.64}$$

On Expanding the linear results for the position and velocity of a fluid element we get,

$$\mathbf{x}(\mathbf{q},\tau) = \mathbf{q} - D_1\nabla_{\mathbf{q}}\Phi^{(1)}(\mathbf{q},\tau) + D_2\nabla_{\mathbf{q}}\Phi^{(2)}(\mathbf{q},\tau), \tag{3.65}$$

$$\mathbf{u}(\mathbf{q},\tau) = -f_1\mathscr{H}\nabla_{\mathbf{q}}^{-1}\Phi^{(1)}(\mathbf{q},\tau) + f_2\mathscr{H}\nabla_{\mathbf{q}}^{-1}\Phi^{(2)}(\mathbf{q},\tau), \tag{3.66}$$

For flat models with a non-zero cosmological constant Ω_Λ, we have for $0.01 \leq \Omega_m \leq 1$ [Bouchet, 1996]:

$$f_1 \approx \Omega_m^{5/9}, \quad f_2 \approx 2\Omega_m^{6/11}. \tag{3.67}$$

While extending the solution to the third order displacement field provides a slightly better description of under-dense regions and additional substructures in high-density regions [Bouchet, 1996], it offers a minimal improvement over 2LPT, as noted by [Buchert et al., 1994, Melott et al., 1995].

We can use LPT to generate initial conditions for N-body numerical simulations. The first step involves generating random Gaussian density fluctuations in Fourier space, by using the definition of the power spectrum i.e.

$$\delta_{\mathbf{k}} = \sqrt{P_m^L}A\exp(i\phi)\Big), \tag{3.68}$$

with a random amplitude (i.e. fluctuation around $\sqrt{P_m^L}$) and a phase. Then it can be connected to the linear part of the gravitational potential $\Phi^1_{\mathbf{k}}$ using Eq. (3.59). To go to the second-order effects we then solve Eq. (3.65) which is the corresponding equation for the second-order gravitational potential. These relations can be easily linked, in Fourier space, to the displacement field at each order. The final step is to inverse-Fourier transform these results, to get the linear and second-order perturbative solutions of the displacement field. Finally, these displacements can be used in Eq. (3.66) to move the particles from their starting grid points and assign to them an initial velocity Eq. (3.67).

3.3 N-body Simulation

The density fluctuation becomes strongly non-linear on small scales, at these scales the PT breaks down and one has to leverage numerical simulations to study their evolution. The current state-of-the-art numerical simulations can follow the dynamics up to about 10^{12} particles [Maksimova et al., 2021]. A thing to keep in mind is that when we say dark matter particles these do not stand for actual dark matter particles rather, they represent small elements of the dark matter distribution in phase space. The mass of the particle m is only a numerical parameter and it is determined by the total amount of matter in the simulation volume divided by the number of particles. So, a higher resolution simulation has more particles with correspondingly smaller mass. N-body simulations discretize the phase space into N elementary volumes "particles" with well-defined positions, velocities and masses and follow the evolution of these test particles due to the action of gravity.

The basic steps of an N-body simulation proceed as follows.

1. Implementation of initial conditions ([Michaux et al., 2020, Sirko, 2005] and references therein)
2. Calculating force by solving the Poisson equation.
3. Update position and velocities of particles using integrators that preserve the phase space volume. eg the Leapfrog method [Hockney and Eastwood, 1981]
4. Go back to step 2. until the simulation is completed.

The methods used for these simulations mainly differ in how the Poisson equation is solved which corresponds to gravitational forces(for a review check [Trenti and Hut, 2008]). There are two primary algorithmic approaches for this task.

The first one is the grid-based method also known as the Particle-Mesh method(PM) in which, each particle's mass is distributed onto a $3D$ grid to create a smooth density field. This can be achieved using various methods such as the Cloud-in-cell (CIC)[Birdsall and Fuss, 1997] method Piece-wise Spline (PCS) etc. The grid can either have a fixed resolution or be adaptively refined in regions with high particle density. The Poisson equation is then solved for the potential, Φ, using the Fast Fourier Transform (FFT)[Press et al., 2007], or other efficient numerical methods for adaptively refined grids. Finally, the potential gradient is interpolated to each particle's position to compute gravitational forces.

Next is the tree algorithm [Barnes and Hut, 1986] which expands gravitational forces into multipoles, focusing on the lowest-order contributions from distant particles. This method efficiently approximates forces by grouping distant particles and treating them collectively.

Both the grid-based and tree algorithms employ force smoothing [Dehnen, 2001], known as "softening," on small scales to prevent nonphysical particle-particle interactions. This adjustment is necessary because simulation particles represent aggregated mass distributions rather than discrete physical particles.

The computational expense of both tree algorithms and adaptive grid methods scales approximately as $\mathcal{O}(N \log N)$ with the number of particles $\mathcal{O}(N)$. This contrasts sharply with the direct summation of forces [Wang et al., 2015], which scales as $\mathcal{O}(N^2)$ and would be prohibitively expensive for simulations involving billions of particles.

Then there are hybrid methods such as the P3M and PM-Tree codes [Springel et al., 2005]. The P3M (Particle-Particle-Particle-Mesh) algorithm enhances the dynamic range of the Particle-Mesh (PM) algorithm by efficiently handling long-range gravitational interactions. However, in scenarios where particles cluster strongly, a substantial number of particles interact directly with each other, resulting in a computational complexity of $\mathcal{O}(N^2)$ and slows down the computations. To mitigate this issue, adaptive mesh techniques are employed. These adapt the spatial resolution of the mesh, refining it in regions of high particle density. Adaptive P3M [Norman and Bryan,] implementations thus maintain computational efficiency by scaling with $\mathcal{O}(N \log(N))$, similar to tree codes. This adaptive approach ensures that computational resources are focused where they are most needed, optimizing performance. Alternatively, another strategy involves hybridizing the PM algorithm with a tree code for short-range force calculations[Springel et al., 2005]. This hybrid PM-Tree scheme leverages the efficiency of PM for long-range interactions and the locality-handling capabilities of tree codes for short-range interactions.

So, once the N-body simulation is processed it generates "snapshots" of particle locations and velocity at various time points. We can compute the density field by assigning particles to a grid which allows us to measure various summary statistics like the power, spectrum Bi-spectrum and perform cosmological analysis. We can also search for gravitationally bound groups of particles, known as dark matter halos, that can reveal the location of galaxies.

Chapter 4

Statistical Analysis of Random Cosmic Field

4.1 Random fields

A significant challenge in modern cosmology is identifying appropriate tools to analyze the distribution of density fluctuations, their initial conditions, and their subsequent evolution. The current explanation for the large-scale structure of the Universe posits that the present distribution of matter on cosmological scales results from the growth of primordial, small seed fluctuations in an otherwise homogeneous Universe, amplified by gravitational instability.

Testing cosmological theories that characterize these primordial seeds is inherently statistical rather than deterministic. This is due to the lack of direct observational access to primordial fluctuations, which would provide definite initial conditions for deterministic evolution equations. Moreover, the timescale for cosmological evolution far exceeds the period over which we can make observations, preventing us from following the evolution of individual systems. Instead, we observe different objects at various stages of their evolution through our past light cone.

The observable Universe is modelled as a stochastic realization of a statistical ensemble of possibilities. The objective is to make statistical predictions based on the properties of the primordial perturbations that lead to the formation of large-scale structures. According to the cosmological principle, which asserts that the Universe is homogeneous and isotropic on large scales, cosmological models assume that the statistical properties of density fluctuations are uniform across widely separated regions of the Universe.

A crucial assumption in cosmology is that the observable part of the Universe is a fair sample of the whole. This implies that statistical measures, such as correlation functions, should be considered averages over the ensemble. However, we only have one realization of the Universe. The fair sample hypothesis posits that samples from well-separated regions of the Universe are independent realizations of the same physical process, and there are enough independent samples in the observable Universe to be representative of the statistical ensemble. The hypothesis of ergodicity further states that averaging over many realizations is equivalent to averaging over a sufficiently large volume. [Verde, 2008][Bernardeau et al., 2002]

Theories are limited to predicting the statistical characteristics of the density contrast $\delta(x)$. The cosmological principle requires that this density contrast forms an isotropic and homogeneous random field, of which the observable Universe is one realization. Among the models explaining the large-scale structure of the Universe, those based on the inflationary paradigm are most widely considered. In the simplest single-field models, these generate adiabatic, Gaussian initial fluctuations, with the origin of stochasticity lying in quantum fluctuations generated in the early Universe.

In cosmology, the scalar field $\delta(x)$ suffices to define the initial fluctuations field and, ideally, the current matter and galaxy distribution. Identifying the right tools to analyze the distribution of density fluctuations, their initial conditions, and their evolution is a major challenge in the study of cosmic structures. In the following sections, we will discuss some of these tools in detail.

For a cosmic random field, such as matter density, the infinite volume condition of the ergodic hypothesis is not satisfied, due to the limited size of the observable Universe. In this case, the ensemble average of an observable overall realization is equal to the spatial average. This means that the ensemble average of any cosmological quantity will be an estimator of its true value. The expectation value of the n-point correlation function (=i.e. the n^{th} central moment for a multivariate joint probability density distribution of the density field $\rho(\mathbf{x})$ is defined as [Gnedenko, 1998].

$$\begin{aligned}\langle(\rho(\mathbf{x}_1)-\rho_0)(\rho(\mathbf{x}_2)-\rho_0)\ldots(\rho(\mathbf{x}_n)-\rho_0)\rangle = \int_V (\rho(\mathbf{x}_1)-\rho_0)(\rho(\mathbf{x}_2)-\rho_0)\ldots(\rho(\mathbf{x}_n)-\rho_0) \\ \times\, p[\rho(\mathbf{x}_1),\rho(\mathbf{x}_2),\ldots,\rho(\mathbf{x}_n)] \qquad (4.1)\end{aligned}$$

Here ρ is the expectation value of the average density field and the integration is over the infinite volume. Then the Fourier transformation of the joint probability density function is given by the characteristic function $M(t)$ as:

$$M(t) = \langle \exp(t\rho)\rangle = \int P(\rho)e^{t\rho}\,d\rho \qquad (4.1)$$

Using the series expansion of the exponential function, we get:

$$e^{t\rho} = \sum_{n=0}^{\infty}\frac{(t\rho)^n}{n!} \qquad (4.2)$$

Substituting this series expansion into the integral, we obtain:

$$M(t) = \int P(\rho)\left(\sum_{n=0}^{\infty}\frac{(t\rho)^n}{n!}\right)d\rho \qquad (4.3)$$

Interchanging the order of summation and integration, we have:

$$M(t) = \sum_{n=0}^{\infty}\frac{t^n}{n!}\int P(\rho)\rho^n\,d\rho \qquad (4.4)$$

The integral $\int P(\rho)\rho^n\,d\rho$ represents the n-th moment of ρ, denoted as $\langle\rho^n\rangle$. Therefore, the characteristic function can be written as:

$$M(t) = \sum_{n=0}^{\infty}\frac{t^n}{n!}\langle\rho^n\rangle \qquad (4.5)$$

where $\rho = \rho_1, \rho_2, \ldots, \rho_n$ is a vector. The expectation value of $\langle\rho^n\rangle$ is the raw moment, i.e. the expectation value for $\rho_0 = 0$, and it is related to the central moments through the binomial transformation. Taking logarithm of Eq. 3.2 we get the n-th cumulants as:

$$\ln M(t) = \sum_{n=1}\frac{t^n}{n!}\langle\rho^n\rangle_c \qquad (4.2)$$

Equating the two, after expanding the Maclaurin series of $\ln M(t)$, we can get the relationship between cumulants and central moments of the density random field. Lets us see the the one-point results for the first four:

$$\langle\rho\rangle_c = \langle\rho\rangle \tag{4.6}$$
$$\left\langle\rho^2\right\rangle_c = \left\langle\rho^2\right\rangle - \langle\rho\rangle_c^2 \tag{4.7}$$
$$\left\langle\rho^3\right\rangle_c = \left\langle\rho^3\right\rangle - 3\left\langle\rho^2\right\rangle_c\langle\rho\rangle_c - \langle\rho\rangle_c^3 \tag{4.3}$$
$$\left\langle\rho^4\right\rangle_c = \left\langle\rho^4\right\rangle - 4\left\langle\rho^3\right\rangle_c\langle\rho\rangle_c - 3\left\langle\rho^2\right\rangle_c^2 - 6\left\langle\rho^2\right\rangle_c\langle\rho\rangle_c^2 - \langle\rho\rangle_c^4, \tag{4.8}$$

For the overdensity field, $\delta(\mathbf{x})$, these relations can be derived by normalizing with the average density. Since the first central moment of the random overdensity fluctuations is zero (i.e. $\langle\delta\rangle = 0$), the above equations are significantly simplified.

When working with multipoint correlation functions, it is important to consider all permutations of the random field between points. The cumulant is the fundamental statistical measure of interest since it represents a set of independent quantities that fully characterize the perturbations' probability distribution function (PDF). Cumulants are commonly referred to in the literature as connected correlation functions, a term derived from quantum field theory and Feynman diagrams eg.Fig. 3.1. The covariance is the second-order cumulant, or two-point cumulant, while the variance is the first-order cumulant, or one-point cumulant (i.e., the diagonal component).

Now, most inflationary models predict a Gaussian distribution for the initial density fluctuations due to the quadratic nature of the action. In the multi-variant case, we have

$$P\left[\delta\left(\mathbf{x}_1\right), \delta\left(\mathbf{x}_2\right), \ldots, \delta\left(\mathbf{x}_n\right)\right] = \frac{1}{\sqrt{2\pi}\det(C)} \exp\left[\frac{1}{2}\delta_i C_{ij}^{-1}\delta_j\right] \tag{4.4}$$

where $Cij = \langle\delta_i\delta_j\rangle_c$ is the covariance and $\delta_i \equiv \delta\left(\mathbf{x}_i\right)$. Substituting the Gaussian PDF in the characteristic function, all the odd cumulants vanish while the even are obtained by the sum of the product of the ensemble averages of two-point correlators, with all possible combinations between the different positions. This is encoded in the Wick's theorem of quantum and classical field theories as

$$\langle\delta_1\delta_2\ldots\rangle = \sum_{\text{pairings}}\prod_{\text{pairs}\ (i,j)}\langle\delta_i\delta_j\rangle_c \tag{4.5}$$

However, in the case of a primordial non-Gaussian(PNG) perturbation field, the even-order cumulants are non-zero which implies that these higher-order correlators are a direct indication of the departure from Gaussianity. However, in the case of LSS, non-zero higher-order cumulants are seen due to the non-linear nature of gravity, which induces couplings between different modes and one of the challenging tasks is to disentangle the PNG signals from the late time evolution. In contrast, CMB probes see fluctuations immediately, following decoupling, making them free of such nonlinear influences.

4.2 Two-point correlation function and power spectrum

As we saw in section 2.7.5 the two-point correlation function of the density field is the ensemble average of δ at two different points. It is given by

$$\langle\delta\left(\mathbf{x}_1\right)\delta\left(\mathbf{x}_2\right)\rangle = \langle\delta\left(\mathbf{x}_1\right)\delta\left(\mathbf{x}_2\right)\rangle_c = \xi(r), \tag{4.6}$$

and taking the Fourier transform of $\xi(r)$ we get to Eq. 2.179 where the quantity $P(k)$ is the Fourier coefficient of the two-point correlation function, called the power spectrum. We saw strictly positive for a continuous random field $\delta(\mathbf{x})$ and it is non-zero only for equal and opposite wavenumbers.

Now, let us try to understand what the power spectrum actually measures and look into Eq. 2.180, which is the one-point second cumulant ($\xi(0)$) of the linear overdensity field and we write it here again,

$$\sigma^2 = \left\langle\delta^2(\mathbf{k})\right\rangle_c = \int\frac{dk}{2\pi^2}k^2P(k) \tag{4.7}$$

From the equation, we see that the power spectrum characterizes the amplitude of the density fluctuations around the mean background value.

Now, we will define the smoothed linear matter density contrast δ_R, over a radius $R(M) = (3M/4\pi\bar{\rho}_m)^{1/3}$, as[1]

$$\delta_R(\mathbf{k}, z) = W_R(k)\delta(\mathbf{k}, z). \tag{4.8}$$

. A popular choice for the filter is $W_R(k) = 3(\sin kR - kR\cos(kR))/(kR)^3$, which is the Fourier transform of the spherical top-hat window function [2]. Hence, we can write a smoothed linear power spectrum $P_R(k)$ as in Eq. 3.29, where now $M(k, z)$ is replaced by $M_R(k, z) = W_R(k)M(k, z)$. The smoothed mass variance of the density field, at mass scale M, is defined as

$$\sigma_R^2(z) = \left\langle \delta_R^2(\mathbf{k}) \right\rangle_c = \frac{1}{2\pi^2} \int k^2 P_R(k, z) dk \tag{4.9}$$

For a density random field with a well-defined average ($\rho > 0$), it is evident that the mass variance must diminish towards zero as the radius R extends indefinitely (i.e., $\lim_{R\to\infty} \sigma^2(R) = 0$). Consequently, the two-point correlation function must also approach zero for large separations (i.e., $\xi(r \to \infty) \to 0$). This leads to the condition:

$$\int d^3r\, \xi(\mathbf{r}) = 0 \tag{4.10}$$

where the integral spans across all space. As a consequence, there exist distances r where $\xi(r) < 0$. Another significant property of the two-point correlation function, arising from the ergodicity of the density random field, is its maximum value at zero separation (i.e., $\xi(0) > |\xi(\mathbf{r})|$). Finally, Eq. (3.85) defines the correlation length as:

$$r_c^2 = \frac{\int dr\, r^2 |\xi(r)|}{\int dr\, |\xi(r)|} \tag{4.11}$$

This quantity characterizes how far correlations persist in the density fluctuation field, indicating the extent to which a localized perturbation influences the surrounding system.

4.2.1 Perturbative Expansion: up to one-loop

In the previous chapter, we saw that the perturbative solution to the density and the velocity depended on the PT scheme we used, for the first-order (linear) fields. Now we will write the linear matter power spectrum using the linear part of the expansion. Using Eq 3.29 we get:

$$\begin{aligned} &\left\langle \delta_m^{(1)}(\mathbf{k}_1)\, \delta_m^{(1)}(\mathbf{k}_2) \right\rangle_c = M(k_1, z)\, M(k_2, z) \left\langle \Phi(\mathbf{k}_1)\, \Phi(\mathbf{k}_2) \right\rangle_c \quad \Rightarrow \\ &P_m^L(k, z) = M^2(k, z) P_\Phi(k) \end{aligned} \tag{4.12}$$

where Φ is the primordial Bardeen gravitational potential, with a power spectrum given by,

$$P_\phi(k) = \frac{9}{25} \frac{2\pi^2}{k^3} \mathscr{A}_s^2 \left(\frac{k}{k_0}\right)^{n_s - 1} \tag{4.13}$$

where we used Eqs. (2.181, 2.184 and 3.29). Here we have used in the perturbative solution the transfer function, incorporated through the Poisson relation of the linear density field with the primordial gravitational potential Eq. 3.31.

Now, we proceed to derive higher-order corrections to the linear power spectrum, from the perturbative solution of the matter and velocity fields which is done by using the perturbative

[1] In the above equation, M is the mass originating from the matter inside the region of size $R(M)$ and $\bar{\rho}_m$ is the average density of the Universe at present time.

[2] A

expansion of a field and substituting it in the two-point correlation function Eq. 2.174. Keeping only the non-zero combinations, after taking into account the cumulant relation, results in a series of terms with an increasing power of the linear field. The terms that have the power of the linear solution to the n-th order denote the n th-loop correction to the linear order, which is also called tree-level order. The above process leads to:

$$\begin{aligned}
&\left\langle X(\mathbf{k},z)X\left(\mathbf{k}',z\right)\right\rangle_c = \left\langle X^{(1)}_{\mathbf{k}}(z)X^{(1)}_{\mathbf{k}'}(z)\right\rangle_c + 2\left\langle X^{(1)}_{\mathbf{k}}(z)X^{(2)}_{\mathbf{k}'}(z)\right\rangle_c \\
&+ \left\langle X^{(2)}_{\mathbf{k}}(z)X^{(2)}_{\mathbf{k}'}(z)\right\rangle_c + 2\left\langle X^{(1)}_{\mathbf{k}}(z)X^{(3)}_{\mathbf{k}'}(z)\right\rangle_c + \ldots \Rightarrow \\
&P(k,z) = P^{(0)}(k,z) + P^{(1)}(k,z) + \ldots
\end{aligned} \tag{4.14}$$

where the zero loop term is just the linear power spectrum (i.e. $P^{(0)}(k,z) \equiv P^L_m(k,z)\big)$ and $P^{(1)}(k,z)$ is the 1-loop correction. The field X denotes a quantity that can have a perturbative solution, e.g. matter overdensity and velocity fields. In the case of Gaussian initial conditions, all the odd terms in each loop order are zero[3], according to Wick's theorem. In the Eulerian PT, the higher-order perturbative terms are given in Eq. 3.35 and Eq. 3.36 for the density and velocity fields respectively. Gaussian initial conditions, the 1-loop results in standard PT are [Bernardeau et al., 2002]

$$P^{(1)}(k,z) = P_{22}(k,z) + P_{13}(k,z) \tag{4.15}$$

where

$$P_{22}(k,z) \equiv 2\int \left[F^{(s)}_2(\mathbf{k}-\mathbf{q},\mathbf{q})\right]^2 P^L_m(|\mathbf{k}-\mathbf{q}|,z)P^L_m(q,z)\mathrm{d}^3\mathbf{q} \tag{4.16}$$

$$P_{13}(k,z) \equiv 6P^L_m(k,z)\int F^{(s)}_3(\mathbf{k},\mathbf{q},-\mathbf{q})P^L_m(q,z)\mathrm{d}^3\mathbf{q} \tag{4.17}$$

One-loop corrections describe the primal effects of mode coupling and can give a quantitative estimation of the breakdown scales of standard PT. The first part of the 1-loop contribution (i.e. $P_{22}(k,z)$) is positive definite and describes the mode coupling between waves with wave vectors$\mathbf{k}-\mathbf{q}$ and $\mathbf{q}$, coming from the presence of the second order standard PT kernel. On the other hand, the P_{13} term is negative and does not exhibit any mode coupling, i.e. it is just a term proportional to the linear power spectrum. This term can be interpreted as the nonlinear correction to the standard $\alpha(\tau)$l linear growth.

[Jeong and Komatsu, 2006] demonstrated that the power spectrum, including up to 1-loop corrections, agrees with simulations to within 1 % accuracy for redshifts in the range $1 < z < 6$ and on quasi-nonlinear scales. However, beyond a certain scale, 1-loop corrections become insufficient to accurately describe the power spectrum in simulations, necessitating the use of higher-order corrections, such as 2-loop corrections.

To characterize the degree of non-linear evolution when including one loop correction corrections, it is convenient to define scales in real space, R_0, as the scale where the smoothed linear variance is unity $\sigma^2(R_0) = 1$. Additionally, one can use the dimensionless power spectrum and define the non-linear scales k_{NL}, as those where $\mathscr{P}(k_N L) = 4\pi k^3_{NL}P(k_{NL}) = 1$ [Bernardeau et al., 2002].

4.3 Three-point correlation function and bispectrum

The lowest higher-order correlator, beyond the two-point correlation function we can take into account is the three-point correlation function. It is defined as

$$\langle\delta(\mathbf{x}_1)\delta(\mathbf{x}_2)\delta(\mathbf{x}_3)\rangle = \langle\delta(\mathbf{x}_1)\delta(\mathbf{x}_2)\delta(\mathbf{x}_3)\rangle_c = \xi(\mathbf{x}_1,\mathbf{x}_2,\mathbf{x}_3) \tag{4.18}$$

[3]eg. $\left\langle X^{(1)}_{\mathbf{k}}(z)X^{(2)}_{\mathbf{k}'}(z)\right\rangle_c = 0$

The Fourier transform of the three-point correlation function is known as the bispectrum which is written as:

$$\langle \delta(\mathbf{k}_1)\,\delta(\mathbf{k}_2)\,\delta(\mathbf{k}_3)\rangle\, c = (2\pi)^3 \delta_D(\mathbf{k}_1+\mathbf{k}_2+\mathbf{k}_3)\, B(\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3) \tag{4.19}$$

The Dirac delta function in the above equation enforces the condition of translational invariance, implying that the wave vectors $\mathbf{k}_i$ form the sides of a closed triangle (i.e., $\mathbf{k}1+\mathbf{k}_2+\mathbf{k}_3=0$). Due to statistical isotropy, the bispectrum $B(\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3)$ depends only on the magnitudes of the wave vectors, $k_i=|\mathbf{k}_i|$, and the distances between points, $r_{ij}=|\mathbf{r}_{ij}|$. The bispectrum provides crucial insights into the nonlinear evolution of the density and velocity fields, as it captures information from the nonlinear regime

The relationship between the three-point correlation function in real space and its Fourier transform, the bispectrum, is given by:

$$\xi(\mathbf{x}_1,\mathbf{x}_2,\mathbf{x}_3) = \int \frac{d^3k1}{(2\pi)^3}\int \frac{d^3k_2}{(2\pi)^3}\frac{d^3k_3}{(2\pi)^3}\delta_D(\mathbf{k}_1+\mathbf{k}_2+\mathbf{k}_3)\, B(\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3)\, e^{i(\mathbf{k}_1\cdot\mathbf{x}_1+\mathbf{k}_2\cdot\mathbf{x}_2+\mathbf{k}_3\cdot\mathbf{x}_3)} \tag{4.20}$$

Now let us consider the perturbative expansion of the quantity X we get:

$$\begin{aligned}\langle X(k_1,z)\,X(k_2,z)\,X(k_3,z)\rangle_c &= \left\langle X^{(1)}(k_1,z)\,X^{(1)}(k_2,z)\,X^{(1)}(k_3,z)\right\rangle_c \\ &+ \left\langle X^{(1)}(k_1,z)\,X^{(1)}(k_2,z)\,X^{(2)}(k_3,z)\right\rangle_c + \left\langle X^{(1)}(k_1,z)\,X^{(2)}(k_2,z)\,X^{(1)}(k_3,z)\right\rangle_c \\ &+ \left\langle X^{(2)}(k_1,z)\,X^{(1)}(k_2,z)\,X^{(1)}(k_3,z)\right\rangle_c + \ldots \Rightarrow \end{aligned} \tag{4.21}$$

$$B(k_1,k_2,k_3,z) = B_{111}(k_1,k_2,k_3,z) + B_{112}(k_1,k_2,k_3,z) + 2 \text{ perm } + \ldots \tag{4.22}$$

Let's recall the application of Wick's theorem for a Gaussian random field, where the first-order term vanishes. Moving to the second-order density field, we find $\left\langle X^{(1)}X^{(1)}X^{(2)}\right\rangle_c \propto \left\langle X^{(1)}X^{(1)}X^{(1)}X^{(1)}\right\rangle_c$, leading us to the four-point correlator. In the Gaussian case, this is expressed through pairs of the product of two-point correlators. Now, let's outline the tree-level bispectrum of the matter field:

$$\begin{aligned} B_G(k_1,k_2,k_3,z) \equiv B_{112} = 2\big[&F_2(\mathbf{k}_1,\mathbf{k}_2)\,P_m^L(k_1,z)\,P_m^L(k_2,z) \\ &+ F_2(\mathbf{k}_2,\mathbf{k}_3)\,P_m^L(k_2,z)\,P_m^L(k_3,z) \\ &+ F_2(\mathbf{k}_3,\mathbf{k}_1)\,P_m^L(k_3,z)\,P_m^L(k_1,z)\big] \end{aligned} \tag{4.23}$$

The configuration dependence of B_G stems from the kernel F_2, which includes terms like $\alpha(\mathbf{k}_1,\mathbf{k}_2)$ arising from the gradient of the density with the velocity field (represented by the $\mathbf{u}\nabla\delta$ term in the continuity equation) and contributions from the $(\mathbf{u}\nabla)\cdot\mathbf{u}$ term in the Euler equation which is gravity-induced. Consequently, the tree-level bispectrum exhibits high sensitivity to the nonlinear effects of gravity, enhancing the amplitude of the three-point correlator. Its scale dependence originates from the linear power spectrum, which amplifies the anisotropic characteristics of the bispectrum on large scales (refer to [Liguori et al., 2010] and the references therein). To be able to study the shape of the tree-level bispectrum we introduce the reduced bispectrum quantity as [Fry, 1984]:

$$Q \equiv \frac{B(k_1,k_2,k_3,z)}{P_m^L(k_1,z)\,P_m^L(k_2,z)+P_m^L(k_2,z)\,P_m^L(k_3,z)+P_m^L(k_1,z)\,P_m^L(k_3,z)}, \tag{4.24}$$

which at tree-level is independent of redshift [Fry, 1984] and scale dependencies of the gravity contribution [Fry, 1994]. Now, varying the angle θ that satisfies the translation invariance $k_1+k_2+k_3=0$ and generating triangle configurations by fixing *k1*, *k2* and z it can be shown that it is easier to find galaxies in a row than in an equilateral configuration physically which means that filaments are the preferred structures in the Universe [Liguori et al., 2010].

Now, let's consider the case of primordial non-Gaussianity (PNG), where additional terms appear in the bispectrum even at the tree-level. The bispectrum is an ideal tool for studying PNG due to its high sensitivity to non-linearities. However, this sensitivity also presents a challenge: we must carefully remove any contributions arising from gravitational evolution to draw meaningful conclusions about the primordial Universe. The first non-zero term in this context is B_{111}, which represents the linearly extrapolated primordial bispectrum of the gravitational potential as predicted by inflationary models. It is given by:

$$B_I\left(k_1,k_2,k_3,z\right)\equiv B_{111}\left(k_1,k_2,k_3,z\right)=M\left(k_1,z\right)M\left(k_2,z\right)M\left(k_3,z\right)B_\Phi\left(k_1,k_2,k_3\right) \tag{4.25}$$

Let us see in the case of the local primordial non-Gaussianity, the primordial gravitational field is written as a Taylor expansion around the Gaussian part as,

$$\Phi(\mathbf{x})=\Phi_G(\mathbf{x})+f_{\mathrm{NL}}^{\mathrm{local}}\left(\Phi_G^2(\mathbf{x})-\left\langle\Phi_G^2(\mathbf{x})\right\rangle\right)+\ldots \tag{4.26}$$

Plugging this expansion in Eq. (4.20), where now $X=\Phi$, we get a similar expression to that of the tree-level bispectrum. Note that the kernel F_2 will not be present, due to the primordial nature of Φ. taking into account the Poisson equation, as well as by using the local expansion of Φ, we get up to second order:

$$\begin{aligned}
&\delta_{\mathrm{lin}}\left(\mathbf{k},z\right)=M(k,z)\Phi(\mathbf{k})\\
&=M(k,z)\Phi^{(1)}(\mathbf{k})+f_{\mathrm{NL}}^{\mathrm{loc}}\,M(k,z)\Phi^{(2)}(\mathbf{k})\\
&=M(k,z)\Phi_G(\mathbf{k})+M(k,z)f_{\mathrm{NL}}^{\mathrm{loc}}\int\frac{d^3q_1}{(2\pi)^3}\frac{d^3q_2}{(2\pi)^3}\delta_D\left(\mathbf{k}-\mathbf{q}_{12}\right)\Phi_G\left(\mathbf{q}_1\right)\Phi_G\left(\mathbf{q}_2\right)\\
&=\delta_{\mathrm{lin}}^{(1)}\left(\mathbf{k},z\right)+f_{\mathrm{NL}}^{\mathrm{loc}}\delta_{\mathrm{lin}}^{(2)}\left(\mathbf{k},z\right)
\end{aligned} \tag{4.27}$$

The first-order results, which coincide with the Gaussian case, are given by $\delta^{(1)}(\mathbf{k},z)\equiv$ $\delta_{\mathrm{lin}}^{(1)}\left(\mathbf{k},z\right)=M(k,z)\Phi_G(\mathbf{k})$. Plugging this in B_{111}, we get Eq, (2.217) [Creminelli et al., 2006]. The signal coming from the primordial bispectrum is much smaller than the tree-level gravitational bispectrum [Sefusatti and Komatsu, 2007].

4.3.1 One-loop matter bispectrum

The 1-loop corrections to the tree-level bispectrum is given in an analogous way to the power spectrum $\left(B=B^{(0)}+B^{(1)}\right)$, i.e. by adding higher order solutions of the density field in the connected correlator and keeping terms up to some power of the linear density field. For Gaussian initial conditions, we get four terms that constitute the bispectrum 1-loop, involving up to fourth order perturbative solution, i.e. $B^{(1)}=B_{222}+B_{321}^I+B_{321}^I+B_{411}$.

Each term is given as follows [Scoccimarro, 1997, Scoccimarro et al., 1999, Bernardeau et al., 2002]

$$\begin{aligned}
B_{222}\equiv&8\int\frac{d^3q}{(2\pi)^3}P_m^L(q,z)F_2^{(s)}\left(-\mathbf{q},\mathbf{q}+\mathbf{k}_1\right)P_m^L\left(\left|\mathbf{q}+\mathbf{k}_1\right|,z\right)\\
&\times F_2^{(s)}\left(-\mathbf{q}-\mathbf{k}_1,\mathbf{q}-\mathbf{k}_2\right)P_m^L\left(\left|\mathbf{q}-\mathbf{k}_2\right|,z\right)F_2^{(s)}\left(\mathbf{k}_2-\mathbf{q},\mathbf{q}\right)
\end{aligned} \tag{4.28}$$

$$\begin{aligned}
B_{321}^I\equiv&6P_m^L\left(k_3,z\right)\int\frac{d^3q}{(2\pi)^3}P_m^L(q,z)F_3^{(s)}\left(-\mathbf{q},\mathbf{q}-\mathbf{k}_2,-\mathbf{k}_3\right)P_m^L\left(\left|\mathbf{q}-\mathbf{k}_2\right|,z\right)\\
&\times F_2^{(s)}\left(\mathbf{q},\mathbf{k}_2-\mathbf{q}\right)+5\text{ perm}
\end{aligned} \tag{4.29}$$

$$\begin{aligned}
B_{321}^{II}\equiv&6P_m^L\left(k_2,z\right)P_m^L\left(k_3,z\right)F_2^{(s)}\left(\mathbf{k}_2,\mathbf{k}_3\right)\int\frac{d^3q}{(2\pi)^3}P_m^L(q,z)F_3^{(s)}\left(\mathbf{k}_3,\mathbf{q},-\mathbf{q}\right)\\
&+5\text{ perm},
\end{aligned} \tag{4.30}$$

$$\begin{aligned}
B_{411}\equiv&12P_m^L\left(k_2,z\right)P_m^L\left(k_3,z\right)\int\frac{d^3q}{(2\pi)^3}P_m^L(q,z)F_4^{(s)}\left(\mathbf{q},-\mathbf{q},-\mathbf{k}_2,-\mathbf{k}_3\right)\\
&+2\text{ perm}.
\end{aligned} \tag{4.31}$$

For non-Gaussian initial conditions, there are additional terms introduced in the above, contributing at each order with up to $\mathscr{O}\left(\delta^6\right)$. They are given by [Sefusatti, 2009]:

$$B_{NG}^{(1)} = B_{112}^{II} + B_{122}^{I} + B_{122}^{II} + B_{113}^{I} + B_{113}^{II} \tag{4.32}$$

The first non-trivial term, involving up to second-order solutions of the density field, is

$$B_{112}^{II} \equiv= \int \frac{d^3q}{(2\pi)^3} F_2^{(s)}\left(\mathbf{q}, \mathbf{k}_3 - \mathbf{q}\right) T_I\left(\mathbf{k}_1, \mathbf{k}_2, \mathbf{q}, \mathbf{k}_3 - \mathbf{q}\right) + 2 \text{ perm} \tag{4.33}$$

This correction to the $\mathscr{O}\left(\delta^4\right)$ terms of the matter bispectrum Eq. (4.22) is negligibly small, due to the kernel suppression [Scoccimarro et al., 2004]. The remaining terms of the 1-loop correction are listed below [Sefusatti, 2009]:

$$\begin{aligned}
B_{122}^{I} &= 2P_m^L\left(k_1, z\right)\left[F_2^{(s)}\left(\mathbf{k}_1, \mathbf{k}_3\right) \int \frac{d^3q}{(2\pi)^3} F_2^{(s)}\left(\mathbf{q}, \mathbf{k}_3 - \mathbf{q}\right) B_I\left(k_3, q, |\mathbf{k}_3 - \mathbf{q}|\right) + \left(k_3 \leftrightarrow k_2\right)\right] + 2 \text{ perm.} \\
&= F_2^{(s)}\left(\mathbf{k}_1, \mathbf{k}_2\right)\left[P_m^L\left(k_1, z\right) P_{12}\left(k_2\right) + P_m^L\left(k_2, z\right) P_{12}\left(k_1\right)\right] + 2 \text{ perm.} \qquad (4.34) \\
B_{122}^{II} &= 4 \int \frac{d^3q}{(2\pi)^3} F_2^{(s)}\left(\mathbf{q}, \mathbf{k}_2 - \mathbf{q}\right) F_2^{(s)}\left(\mathbf{k}_1 + \mathbf{q}, \mathbf{k}_2 - \mathbf{q}\right) B_I\left(k_1, q, |\mathbf{k}_1 + \mathbf{q}|\right) \\
&\quad \times P_m^L\left(|\mathbf{k}_2 - \mathbf{q}|, z\right) + 2 \text{ perm.} \qquad (4.35) \\
B_{113}^{I} &= 3B_I\left(k_1, k_2, k_3\right) \int \frac{d^3q}{(2\pi)^3} F_3^{(s)}\left(\mathbf{k}_3, \mathbf{q}, -\mathbf{q}\right) P_m^L(q, z) + 2 \text{ perm.}, \qquad (4.36) \\
B_{113}^{II} &= 3P_m^L\left(k_1, z\right) \int \frac{d^3q}{(2\pi)^3} F_3^{(s)}\left(\mathbf{k}_1, \mathbf{q}, \mathbf{k}_2 - \mathbf{q}\right) B_I\left(k_2, q, |\mathbf{k}_2 - \mathbf{q}|\right) + \left(k_1 \leftrightarrow k_2\right) + 2 \text{ perm.} \qquad (4.37)
\end{aligned}$$

4.3.2 Marked Statistics

The commonly used statistics to extract cosmological information for eg. the power spectrum is significantly affected by the mass of the most massive objects, and is expected to be suboptimal when extracting information embedded in low-density regions [Rimes and Hamilton, 2005]. The idea of marked statistics is to give more weight to low-density regions and extract information from there.

Marked density fields are weighted density fields where their weight can represent galaxy properties, halo merger histories or environmental densities defined as a weighted sum over particle positions [Skibba et al., 2006, Balaguera-Antolínez, 2014, Beisbart and Kerscher, 2000]. So, The key statistic is the marked density field, defined as a weighted sum over particle positions[Philcox et al., 2020];

$$\rho_M(\mathbf{x}) = \sum_i \delta_D\left(\mathbf{x} - \mathbf{x}_i\right) m\left(\mathbf{x}_i\right) = \int d\mathbf{x}' \left[\sum_i \delta_D\left(\mathbf{x}' - \mathbf{x}_i\right) m\left(\mathbf{x}'\right)\right] \delta_D\left(\mathbf{x} - \mathbf{x}'\right), \tag{4.38}$$

where δ_D is a Dirac delta and i runs over all matter particles. In the above expression, $m(\mathbf{x})$ is the mark, defined as a local overdensity as in [Massara et al., 2021]:

$$m(\mathbf{x}) = \left(\frac{1 + \delta_s}{1 + \delta_s + \delta_R(\mathbf{x})}\right)^p \equiv \left(1 + \frac{\delta_R(\mathbf{x})}{1 + \delta_s}\right)^{-p} \tag{4.39}$$

where $\delta_R(\mathbf{x})$ is the matter overdensity filtered on scale R using Top-Hat filter of radius R, the bias δ_s is a density parameter and the exponent p being user-defined parameters. Defining the sample density field $n\left(\mathbf{x}'\right) = \sum_i \delta_D\left(\mathbf{x}' - \mathbf{x}_i\right)$, Eq. (4.38) can be rewritten as:

$$\rho_M(\mathbf{x}) = m(\mathbf{x}) n(\mathbf{x}) = m(\mathbf{x}) \bar{n}[1 + \delta(\mathbf{x})], \tag{4.40}$$

where $\bar{n} = \langle n(\mathbf{x}) \rangle$ is the average density. In order to convert Eq. (4.40) into an overdensity field, we require the mean density:

$$\langle \rho_M(\mathbf{x}) \rangle = \langle n(\mathbf{x}) m(\mathbf{x}) \rangle = \bar{n}\bar{m}, \tag{4.41}$$

where we have defined $\bar{m}$ as $\langle n(\mathbf{x})m(\mathbf{x})\rangle/\langle n(\mathbf{x})\rangle$, i.e., the average of $m(\mathbf{x})$ weighted by the number density field. The marked overdensity field is thus

$$\delta_M(\mathbf{x}) \equiv \frac{\rho_M(\mathbf{x}) - \langle\rho_M\rangle}{\langle\rho_M\rangle} = \frac{1}{\bar{m}} m(\mathbf{x})[1+\delta(\mathbf{x})] - 1. \tag{4.42}$$

So, we see that depending on the parameters in Eq. (4.39) i.e R, δ_s,p we can define our marks. The exponent p and the bias δ_s dictate how much the value of the local density δ_R impacts the final mark. When p is close to zero or δ_s is very large, the mark $m(\mathbf{x})$ tends to unity. Moreover, when p is positive, galaxies in low-density environments are up-weighted compared to those in high-density regions. This means that mark power spectra with $p > 0$ are more sensitive to low-density regions, while the opposite happens when p is negative.

Recent studies have demonstrated that measuring the power spectrum [Massara et al., 2021, Massara et al., 2023] and bispectrum [Jung et al., 2024a] of the marked density field reveals new cosmological information beyond what is captured by the standard power spectrum. This additional information originates from higher-order statistics embedded in the marked field.

Halo Mass function

Halos are assumed to form on the peaks of the smoothed underlying matter overdensity field when its value exceeds some threshold value δ_c. Therefore, the number of created objects depends on the distribution of points that exceed such a threshold. In the early work of [Press and Schechter, 1974], it was found that the mass function can be expressed in terms of the height of the peaks, by using the spherical collapse model. It states that in a smoothed linear contrast field in Lagrangian space $\delta_R^{(1)}(\mathbf{q})$, a spherical region of radius R with uniform density and enclosed mass $M = (4\pi/3)\bar{\rho}_m R^3$, where $\bar{\rho}_m$ is the mean co-moving density at time t, will collapse to form a bound object when $\delta_R^{(1)}$ exceeds a threshold δ_c (spherical collapse threshold). Assuming Gaussian statistics for the smoothed overdensity field, one can write the probability of having a halo with mass greater than M as a fraction of a Lagrangian volume by

$$p_G\left(\delta_R^{(1)} > \delta_c\right) = \frac{1}{\sqrt{2\pi}\sigma(R)} \int_{\delta_c}^{\infty} d\delta \exp\left[-\frac{1}{2}\frac{\delta^2}{\sigma^2(R)}\right] \tag{4.43}$$

where $\sigma(R)$ is the smoothed variance of the density field over a radius R [Eq.(3.84)]. The Lagrangian volume fraction that encloses the halo with mass greater than M is given by:

$$F(>M) = \frac{1}{\bar{\rho}_m} \int_M^{\infty} d\ln M' M' n_h\left(M'\right) = p_G\left(\delta_R^{(1)} > \delta_c\right) \tag{4.44}$$

where $n_h(M)$ is the co-moving number density of halos above mass M. Differentiating over the halo mass M gives

$$f(M) \equiv -\frac{dF(>M)}{dM} \tag{4.45}$$

which leads to the mass function of halos

$$\begin{aligned} \frac{dF}{dM} &= -n_h(M)M = \frac{1}{\sqrt{2\pi}\sigma(R)} \exp\left[-\frac{1}{2}\frac{\delta^2}{\sigma^2(R)}\right] \frac{d}{dM}\left(\frac{\delta_c}{\sigma_R}\right) \Rightarrow \\ n_h(M) &= -\frac{2\bar{\rho}_m}{\sqrt{2\pi}} \frac{\delta_c}{\sigma_R^2} \frac{1}{M} \frac{d\sigma_R}{dM} \exp\left[-\frac{1}{2}\frac{\delta^2}{\sigma^2(R)}\right] \end{aligned} \tag{4.46}$$

The factor 2 in front of the mean density is introduced to recover the proper normalisation and to get the total mass after the integration over the whole range of M, i.e. $\int_0^{\infty} dM M n_h(M) = \bar{\rho}_m/2$. It is convenient to parametrize the mass function with a multiplicity function $f(\nu)$ as:

$$n_h(M,z) = \frac{d\mathcal{N}}{\mathrm{d}\ln M} = \frac{\bar{\rho}_m}{M} f(\nu) \left| \frac{\mathrm{d}\ln \nu}{\mathrm{d}\ln M} \right| \tag{4.47}$$

where $\nu(M,z) = \delta_c/\sigma_R(M,z)$ is the height of the peak. In the Press-Schechter (PS) formalism this function is simply:

$$f_{PS}(\nu) = \sqrt{\frac{2}{\pi}} \nu e^{-\nu^2/2} \tag{4.48}$$

The PS mass function has a dependence on redshift and cosmological parameters, as well as on the primordial power spectrum, therefore it has a universal character [Sheth and Tormen, 1999] (ST hereafter). The mass function shows that an increasing halo mass leads to high-peaks ($\nu \gg 1$) and therefore to more rare objects. The opposite happens for low-mass halos, which seem to be a common case during the process of halo formation. In ST [Sheth and Tormen, 1999] a modified version of the PS mass function is proposed to improve the agreement with simulations

$$f_{ST}(\nu) = A\sqrt{\frac{2q}{\pi}} \left(1 + \frac{1}{(q\nu^2)^p}\right) \nu e^{-\frac{q\nu^2}{2}} \tag{4.49}$$

where $A = 1/\left(1 + 2^{-p}\Gamma(0.5-p)/\sqrt{\pi}\right) \approx 0.322184, q = 0.707$ and $p = 0.3$.

Although the PS formalism predictions on the form of the mass function are in agreement with simulations, it does not treat properly the small overdensities that might exist inside the Lagrangian radius. This is due to the fact that the PS formalism considers the whole smoothed region as one halo. In other words, all the points inside the halo exceed the threshold value, which is not generally true for realistic cases. This is known as the cloud-in-cloud problem and excursion set formalism was introduced [Bond et al., 1991] to solve it. The latter approach adds the first-crossing condition, where it states that a region belongs to a bound structure only if the smoothing radius R has the maximum value, in order for $\delta_R^{(1)}$ to reach the threshold δ_c.

HMF is known to be sensitive to non-Gaussian initial conditions. Depending on these initial conditions, it skews the distribution by changing the abundance of massive halos. It was proposed as an interesting complementary probe of primordial non-Gaussianity (PNG) to the bispectrum in several studies [Matarrese et al., 2000, Desjacques et al., 2009, Sefusatti and Komatsu, 2007, Sefusatti, 2009]. A major advantage of the HMF is that it directly depends on the PNG amplitude parameter $f_{\rm NL}$. Therefore, it does not exhibit the b_ϕ–$f_{\rm NL}$ degeneracy that affects the scale-dependent power spectrum signature.

In the case of massive neutrinos, the free streaming of the non-relativistic neutrinos counteracts gravitational collapse, the fundamental process of cosmic structure formation. This leads to a significant suppression in the average number density of massive structures which is particularly evident in the high-mass end of the HMF. For a fixed amplitude of primordial curvature perturbations, the degree of suppression in the number density of dark matter halos depends on the total neutrino mass M_ν. As the neutrino mass increases, the suppression in the comoving number density of dark matter halos becomes more pronounced so HMF is also sensitive to the total neutrino mass M_ν [Marulli et al., 2011, Brandbyge et al., 2010].

Chapter 5

Methodology and Data Analysis

The traditional way to put constraints on cosmological parameters from observational data is to compare summary statistics to theoretical predictions. However, an important question that arises is which statistic or combination of statistics should be used to gain the maximum information.

We previously saw that for Gaussian density fields, the power spectrum or correlation function fully describes the field, making them the primary tools in cosmological data analysis. These statistics are effective at high redshifts or on large scales at low redshifts, where the universe resembles a Gaussian density field. The CMB measurements from [Akrami et al., 2019] have limited the primordial universe to be weakly non-Gaussian. However, CMB is limited to high-redshift observations and is insensitive to low-redshift phenomena like the transition from the matter-dominated epoch to the Dark Energy-dominated epoch.

Currently active and upcoming three-dimensional (3D) surveys such as the Euclid mission[EUC,] and Vera C. Rubin Observatory's Legacy Survey of Space and Time (LSST) [Rub,] will allow us to observe a large fraction of the sky and provide very detailed maps of the large-scale structure of the universe. The vast amounts of data collected from these surveys, i.e. all the small-scale modes will offer greater constraining power. In three-dimensional surveys, a significant portion of the modes exists on scales ranging from mildly to non-linear, where the density field exhibits non-Gaussian behaviour. This non-Gaussianity causes information to propagate from the two-point function to higher-order statistics. It is currently an open question as to what is the statistic or set of statistics that will help in improving the constraints on cosmological parameters.

Numerical simulations are essential for tackling this challenge. They provide powerful means to obtain theoretical predictions in fully non-linear regimes (section 3.3) and in both real- and redshift space for any considered statistic. This capability is crucial for extracting cosmological information from non-linear modes, thereby tightening the constraints on cosmological parameters and enhancing our understanding of fundamental physics. The Quijote simulations help in quantifying the information content of various statistics in the fully non-linear regime.

Moreover, advanced statistical techniques such as machine learning and deep learning to perform LFI are being explored to identify new and optimal statistics for extracting cosmological information [Jung et al., 2024b, Alsing et al., 2019, Charnock et al., 2018]. These methods require extensive datasets for training, which the Quijote simulations provide, thereby supporting the analysis of data from new and upcoming surveys.
It was shown in [Bayer et al., 2021] and [Coulton et al., 2023] using the Quijote simulation that the information we can extract on the cosmological parameters using solely the power spectrum saturates at scales $k = 0.2\,h\,\mathrm{Mpc}^{-1}$ and that adding higher-order correlation functions helps in improving constraints on higher scales, greater than $k = 0.2\,h\,\mathrm{Mpc}^{-1}$.

The main aim of this thesis is to develop a machine learning pipeline to perform LFI and identify different summary statistics and combinations of summary statistics that can be used to constrain the sum of massive neutrino (M_ν), f_{NL}^{equil} and other cosmological parameters by extending the analysis to non-linear scale upto to $k_{\max} = 0.5\,h\,\mathrm{Mpc}^{-1}$ leveraging on the Quijote simulation and Quijote-PNG simulations.

In this chapter, we will discuss our approach to constraint M_ν, $f_{NL}^{\rm equil}$ and the cosmological parameters using the Quijote simulation by using Likelihood-free inference using Moment Network [Jeffrey and Wandelt, 2020]. Firstly, we describe the Quijote Simulation used for analysis, followed by the summary statistics used for constraining the parameters. Then we define the method implemented i.e. Likelihood-free inference using moment neural network followed by a summary of machine learning and the neural network model we implemented. Finally, we discuss the indicators we use for evaluating the moment prediction

5.1 Quijote Simulation

The analyses presented in this paper are based on the Quijote [Villaescusa-Navarro et al., 2020a] sets of simulations for M_ν, and Quijote-PNG [Coulton et al., 2023] sets of simulations for $f_{NL}^{\rm equil}$ respectively.
These are dark matter only N-body simulations of volume $1\left(h^{-1}{\rm Gpc}\right)^3$, containing 512^3 CDM particles each for PNG and M_ν, and an extra 512^3 neutrino particles for M_ν . The simulation is run using the TreePM code GadGet-III [Springel et al., 2005] from initial conditions generated at $z = 127$ using 2LPTPNG [Coulton et al., 2023] and Zel'dovich approximation [Zel'dovich, 1970] for the simulations with PNG and M_v respectively. [Villaescusa-Navarro et al., 2020a] uses Zel'dovich approximation and not the second-order Lagrangian perturbation theory (2LPT) for simulation with massive neutrinos as it is currently unknown how to predict 2LPT in massive neutrino models. In particular, there is no estimate for the two significant quantities, the second-order scale-dependent growth factor, and growth rate, necessary to calculate the 2LPT in the presence of massive neutrinos We consider dark matter halos with $\mathrm{M} > 3.2 \times 10^{13}\, h^{-1} M_\odot$ corresponding to a mean tracer density[1] of $\bar{n}(z = 0) = 1.55 \times 10^{-4}\, h^3{\rm Mpc}^{-3}$, which are identified in each simulation by the standard Friends-of-Friends (FoF) algorithm [Davis et al., 1985] by setting the linking length parameter to $b = 0.2$ and considering halos with more than 20 dark matter particles.

- For PNG we mainly use a set of 1000 simulations with varying amounts of equilateral PNG, with $f_{\rm NL}^{\rm equil} \in [-600, 600]$, and the cosmological parameters are changed within $\Omega_m \in [0.1, 0.5]$,$\Omega_b = [0.049]$, $\sigma_8 \in [0.6, 1.0]$,$n_s \in [0.8, 1.2]$ and $h \in [0.5, 0.9]$ where the parameters are distributed in a Latin hypercube (LH $f_{NL}^{\rm equil}$)
- For M_v we use a set of 2000 simulations with varying amounts of $M_\nu({\rm eV}) \in [0, 1]$, $\Omega_m \in [0.1, 0.5]$, $\Omega_b = \in [0.03, 0.07]$, $\sigma_8 \in [0.6, 1.0]$, $n_s \in [0.8, 1.2]$ and $h \in [0.5, 0.9]$ where the parameters are distributed in a Latin hypercube with massive neutrinos (nwLH).

5.2 Summary Statistics Estimators

In this section, we describe the different statistical probes we use from the halo catalogues generated by the Quijote and Quijote PNG simulations in red-shift space, specifically looking at non-linear scales up to $k_{\rm max} = 0.5\, h\, {\rm Mpc}^{-1}$.

We use halo summary statistics as halos are a biased tracer of the underlying matter density field and galaxies form within these halos. So, studying the statistical properties of halos to gain inference on cosmological parameters is a crucial first step in drawing conclusions from real observational data.

5.2.1 Power Spectrum

To measure the halo power spectrum, we use the standard estimator [Feldman et al., 1994] as described in [Jung et al., 2024a] given by:

[1] It is the average number density of tracers in a given volume.

$$\hat{P}(k_i) = \frac{1}{VN_i} \sum_{k \in \Delta_i} \delta(k)\delta^*(k), \tag{5.1}$$

where $\delta(k)$ is the density field in Fourier space defined on a grid, the k-range has been divided into bins Δ_i of width equal to the fundamental mode, V is the survey volume, and N_i is the number of vectors k in each bin.

5.2.2 Bispectrum

To extract the halo bispectrum information we use modal estimators as shown in [Jung et al., 2023]. They can be efficiently extracted from data by measuring the following modal coefficients:

$$\hat{\beta}_n = \frac{1}{V} \int d^3x\, M_p(\mathbf{x}) M_q(\mathbf{x}) M_r(\mathbf{x}), \tag{5.2}$$

where

$$M_p(\mathbf{x}) \equiv \int \frac{d^3k}{(2\pi)^3} \frac{q_p(k)\delta(\mathbf{k})}{\sqrt{kP(k)}} e^{i\mathbf{k}\cdot\mathbf{x}}, \tag{5.3}$$

for a well-chosen basis of one-dimensional functions $q_p(k)$ and mode triplets $n \leftrightarrow (p, q, r)$.

Modal estimators [Fergusson et al., 2012] are based on constructing complete, orthogonal bases of separable bispectrum templates ("bispectrum modes") and finding their amplitudes by fitting them to the data (For a review refer to [Liguori et al., 2010] [Jung et al., 2022]). This procedure can be made fast, using a KSW type of approach [Komatsu et al., 2005] which takes advantage of the fact that certain theoretical bispectrum templates can be written in a separable form. This means the bispectrum can be broken down into products of functions of individual wave vectors. By leveraging the separability, the computational complexity is significantly reduced. Instead of summing over all possible triplets of modes (which is computationally expensive), the problem is reduced to summing over individual modes multiple times, which is much faster.
The vector of estimated mode amplitudes is referred to as the "mode spectrum." This mode spectrum is theory-independent and it contains all the information that needs to be extracted from the data. It is also possible to obtain theoretical mode spectra, by expanding primordial shapes in the same modal basis used to analyse the data. This allows us to measure f_{NL} for any given primordial bispectrum template, by correlating the theoretical mode vectors, which can be quickly computed for any shape, with the data mode spectrum. This feature makes modal techniques ideal for analyses of a large number of competing models. Also important is that non-separable bispectra are expanded with arbitrary precision into separable basis modes. Therefore the treatment of nonseparable shapes is always numerically efficient in the modal approach. Finally, the data mode spectrum can be used, in combination with measured mode amplitudes, to build linear combinations of basis templates, which provide a model-independent reconstruction of the full data bispectrum.

5.2.3 Halo Mass Function

In addition to the halo power spectrum and bispectrum, we extract the halo mass function (HMF) defined using:

$$\frac{\mathrm{d}n}{\mathrm{d}\log M} = \frac{1}{V} \sum_{i=1}^{N_h} \frac{1}{\Delta \log M}, \tag{5.4}$$

where:

- $\frac{\mathrm{d}n}{\mathrm{d}\log M}$ is the number density of halos per unit comoving volume per logarithmic mass bin,
- V is the comoving volume of the simulation or survey,
- N_h is the number of halos within the mass bin $[M, M + \Delta M]$,

- $\Delta \log M$ is the width of the logarithmic mass bin.

We measure it in the Quijote simulations using 15 logarithmic bins corresponding to halo masses M between approximately 2.0×10^{13} and $4.6 \times 10^{15}\, M_\odot/h$ which is the same binning used in [Bayer et al., 2021] where the counted halos each contain between 30 and 7000 dark matter particles.

5.2.4 Marked Statistics

In our analysis, we extend the standard power spectrum and bispectrum calculations to include marked statistics as discussed in section 4.3.2. We saw that these marked statistics provide additional cosmological insights by weighting the density field with a mark function before computing its spectral properties. Let us write down the equation again.

$$m(\mathbf{x}; R, p, \delta_s) = \left[\frac{1+\delta_s}{1+\delta_s+\delta_R(\mathbf{x})}\right]^p, \tag{5.5}$$

where $\delta_R(\mathbf{x})$ was the local density field, smoothed with a top-hat filter at scale R.δ_s and p were the parameters that adjusted the sensitivity and enhancement of low and high-density regions, respectively.

Recent studies [Massara et al., 2021, Massara et al., 2023, Jung et al., 2024a] have demonstrated that measuring the power spectrum of the marked density field reveals new cosmological information beyond what is captured by the standard power spectrum. This additional information originates from higher-order statistics embedded in the marked field.

In our analysis, we use both the power spectrum and the bispectrum of the marked density field. The bispectrum of the marked field requires more computational resources compared to its power spectrum counterpart due to its higher-order nature.

To explore the cosmological implications effectively, we focus on four distinct marks characterized by the following parameter choices. These specific parameter configurations were identified in [Massara et al., 2023] as providing the most stringent constraints on cosmological parameters when analyzing galaxy catalogues derived from the Quijote N-body simulations.:

- Scale R: [30, 25, 20, 30] h^{-1}Mpc
- Power p: [1, 1, 1, 1]
- Sensitivity δ_s: [0.10, 0.25, 0.50, 0.50]

In the case of massive neutrinos marked bispectrum, we only use the following values for R,p, δ_s : [20,1,0.50] i.e the smallest smoothing scale due to the large computational time required for calculating the combination.

5.3 Method

5.3.1 Likelihood Free Inference

In cosmological data analysis, we frequently encounter situations where we can generate mock data through sophisticated forward simulations, yet are unable to formulate a tractable likelihood function.[2] For instance, the physics associated with non-linear structure formation on small scales [Springel et al., 2005, Villaescusa-Navarro et al., 2020b, Klypin et al., 2011], baryonic feedback [Springel et al., 2018] etc., can be captured to varying extents by simulations however, developing compact and accurate models for the statistical properties of these processes remains challenging. Similarly, on the measurement side, it can be difficult to incorporate complicated noise models, selection bias etc., into the likelihood function.

[2]The likelihood function is a function of the parameters of a statistical model, given observed data. It measures the plausibility of different parameter values, based on how likely the observed data is under those parameter values.
Formally, let $\mathbf{X} = (X_1, X_2, \ldots, X_n)$ be a set of observed data points, and let θ be the parameters of the statistical model. If $f(\mathbf{X} \mid \theta)$ denotes the probability density function (or probability mass function, for discrete data) of the data given the parameters, the likelihood function $L(\theta \mid \mathbf{X})$ is defined as $L(\theta \mid X) = f(X \mid \theta)$

The standard approach is to construct an approximate likelihood that aims to capture as much of the known physics and measurement processes underlying the data as possible, hoping that the adopted approximations do not lead to biased posterior inferences. Even if the means and variances of inferences are not significantly biased, assessing tensions between data sets (Marshall et al. 2006), combining inferences, and comparing models can be strongly affected by posterior tail probabilities[Krywonos et al., 2024], which are unlikely to be accurate when using popular likelihood approximations. Moreover, traditional methods like Fisher information and Markov Chain Monte Carlo (MCMC) [Verde, 2008] introduce significant computational complexities. Evaluating likelihood functions in high-dimensional parameter spaces is computationally expensive. MCMC, while effective for exploring complex posterior distributions, often demands extensive computational time to ensure convergence, especially in scenarios involving numerous parameters or complex dependencies[Akeret et al., 2013]. These challenges underscore the need to explore alternative approaches, such as likelihood-free methods, which can mitigate computational burdens while accommodating the complexities inherent in cosmological data analysis

Likelihood-free inference (LFI) also known as simulation-based inference(SBI) allows us to perform Bayesian inference using only forward simulations, free from any likelihood assumptions or approximations. This approach has great appeal for cosmological data analysis, as including complex physical processes, instrumental effects, selection biases, etc., into a forward simulation is typically much easier than incorporating these effects into a complicated likelihood function.

There are several methods to implement likelihood-free inferences, with the simplest incarnation being Approximate Bayesian Computation (ABC). ABC involves drawing parameters from a proposal distribution, simulating mock data, and then accepting or rejecting the parameters based on whether the simulated data falls within an ϵ-ball around the observed data [Alsing et al., 2019, Alsing et al., 2018]. Another widely used method is Density-estimation likelihood-free inference (DELFI), which aims to train a flexible density estimator for the target posterior from a set of simulated data-parameter pairs. DELFI can yield high-fidelity posterior inference from orders of magnitude fewer simulations than traditional ABC-based methods [Papamakarios, 2019, Alsing et al., 2018, Alsing et al., 2019, Lueckmann et al., 2019].

In this thesis, we use likelihood-free inference using a fully connected neural network (NN) and implementing the Moment Network method introduced by [Jeffrey and Wandelt, 2020].

5.3.2 Moment Network

The primary objective of any inference problem is estimating the posterior probability density $p(\theta \mid x)$ of a set of parameters θ given some observed data x where the posterior $p(\theta \mid x)$ contains all beliefs and uncertainties about the unknown quantities θ. However, for physical inference problems involving high-dimensional parameter spaces, the whole high-dimensional posterior distribution is not required nor interpretable. Instead, the objective is frequently to find the marginal one- and two-dimensional posterior distributions of the parameters. The posterior estimations are then primarily utilized to calculate posterior moments. Moment Networks eliminate the need to estimate posterior density and instead focus on estimating parameter location, scale, and covariance (including higher-order moments).

The above model outputs two numbers per parameter, the mean (μ_i)and standard deviation (σ_i)of the marginal posterior 5.1:

$$\mu_i(\mathbf{X}) = \int_{\theta_i} p\left(\theta_i \mid \mathbf{X}\right) \theta_i \, d\theta_i, , \tag{5.6}$$

$$\sigma_i(\mathbf{X}) = \int_{\theta_i} p\left(\theta_i \mid \mathbf{X}\right) \left(\theta_i - \mu_i\right)^2 \, d\theta_i., \tag{5.7}$$

where $p\left(\theta_i \mid \mathbf{X}\right)$ is the marginal posterior over the parameter i:

$$p\left(\theta_i \mid \mathbf{X}\right) = \int_{\theta} p\left(\theta_1, \theta_2, \ldots, \theta_n \mid \mathbf{X}\right)\, d\theta_1 \ldots d\theta_{i-1} d\theta_{i+1} \ldots d\theta_n. \tag{5.8}$$

In this notation, $\mathbf{X}$ represents the summary statistics. Following the moment network method, the loss for the NN function is defined as [Villaescusa-Navarro et al., 2022]:

$$\begin{aligned} \mathscr{L} = & \sum_{i \in \text{params}} \log\left(\sum_{j \in \text{batch}} \left(\theta_{i,j} - \mu_{i,j}\right)^2\right) \\ & + \sum_{i \in \text{params}} \log\left(\sum_{j \in \text{batch}} \left(\left(\theta_{i,j} - \mu_{i,j}\right)^2 - \sigma_{i,j}^2\right)^2\right). \end{aligned} \tag{5.9}$$

which differs from the original one presented by [Jeffrey and Wandelt, 2020],

$$\mathscr{L} = \sum_{i \in \text{params}} \sum_{j \in \text{batch}} \left(\left(\theta_{i,j} - \mu_{i,j}\right)^2 + \left(\left(\theta_{i,j} - \mu_{i,j}\right)^2 - \sigma_{i,j}^2\right)^2\right). \tag{5.10}$$

In the loss defined in [Villaescusa-Navarro et al., 2020a] we see that the arithmetic sum is replaced by the sum of the logarithm of each term (either posterior mean or posterior standard deviation) which was found to provide much tighter and reliable values for $\mu_i$and σ_i than its original version. The reason behind this is that by taking the logarithm of each term before summing the losses are effectively multiplied rather than summed. This adjustment mitigates the issue mentioned above by rescaling all terms to a comparable magnitude. Consequently, the loss function assigns similar weight to all terms, preventing certain parameters from overshadowing others and allowing for more balanced parameter estimation.

5.3.3 Neural Network Model

Before going into the NN model let's briefly try to understand the very basics of machine learning.

Supervised and unsupervised learning are the two primary methods of machine learning. The distinction between the two lies in the fact that in supervised learning, the network is provided with a label in addition to the data sample; in contrast, unsupervised learning lacks this knowledge. This implies that the various strategies may be used to learn either independently in the case of unsupervised learning or dependently in the case of supervised learning, depending on the dataset and the issue. We use supervised learning, which enables the model to evaluate its current performance in comparison to the genuine values.
NN have multiple layers of nodes; one input, one or many hidden, and one output layer. Each node has a threshold which is defined by the activation function, and only if the output is above the threshold will the information be carried to the next layer. The model will fine-tune by learning the best weights to apply to each node in order to get the output which is the closest to the true value. There are different hyperparameters which are set prior to beginning the training, and these are introduced briefly in the following paragraphs.

Nodes and Layers

Neural networks consist of interconnected layers of nodes. Each node processes input data using a weighted sum, which passes through an activation function to determine if the output should propagate to the next layer. The network typically includes:

- **Input Layer**: Receives input data features.
- **Hidden Layers**: Intermediate layers that perform transformations on the input data.
- **Output Layer**: Produces the final output prediction.

Optimizers

Optimization is crucial in achieving the goal of any machine learning algorithm. The optimizer updates the model parameters iteratively to minimize the error function. Two widely used optimizers are:

- **Gradient Descent**: Calculates gradients for the entire dataset and updates weights accordingly.
- **Adam (Adaptive Moment Estimation)**: Combines momentum and squared gradients to adaptively adjust learning rates for each parameter.

Learning Rate

The learning rate is a hyperparameter that controls the step size of the optimizer during training. A higher learning rate accelerates convergence but may risk overshooting the optimal solution, while a lower learning rate requires more epochs for convergence but is less likely to overshoot. Adaptive learning rates that adjust during training are often preferred for improved convergence efficiency.

Batch Size

Batch size refers to the number of data samples processed in one iteration of the optimizer. Common choices include batch, mini-batch, or stochastic (one sample) gradient descent. Batch size affects training time per epoch and can influence model performance when combined with learning rate adjustments.

Regularization

Regularization techniques constrain a model's capacity to prevent overfitting and improve generalization to unseen data. Below mentioned are three most commonly used regularization techniques are

- **L1 Regularization**: Adds a penalty proportional to the sum of the absolute values of the model weights to the loss function:

$$\mathscr{L}_{\text{regularized}} = \mathscr{L}_{\text{original}} + \lambda \sum_i |w_i|$$

  This technique shrinks the less important feature's coefficient to zero thus, removing some features altogether. It works well for feature selection in case we have a huge number of features.

- **L2 Regularization**: Adds a penalty proportional to the sum of the squared values of the model weights to the loss function:

$$\mathscr{L}_{\text{regularized}} = \mathscr{L}_{\text{original}} + \lambda \sum_i w_i^2$$

  This method penalizes large weights, promoting smoother model responses, reducing sensitivity to noise in the training data, and avoiding overfitting.

- **Dropout**: A regularization technique where a fraction of neurons is randomly deactivated during each training iteration:

$$\text{Dropout}(x) = x \cdot \text{mask}$$

  where the mask is a binary vector sampled from a Bernoulli distribution. Dropout prevents overfitting by ensuring that the model does not rely on any specific neurons.

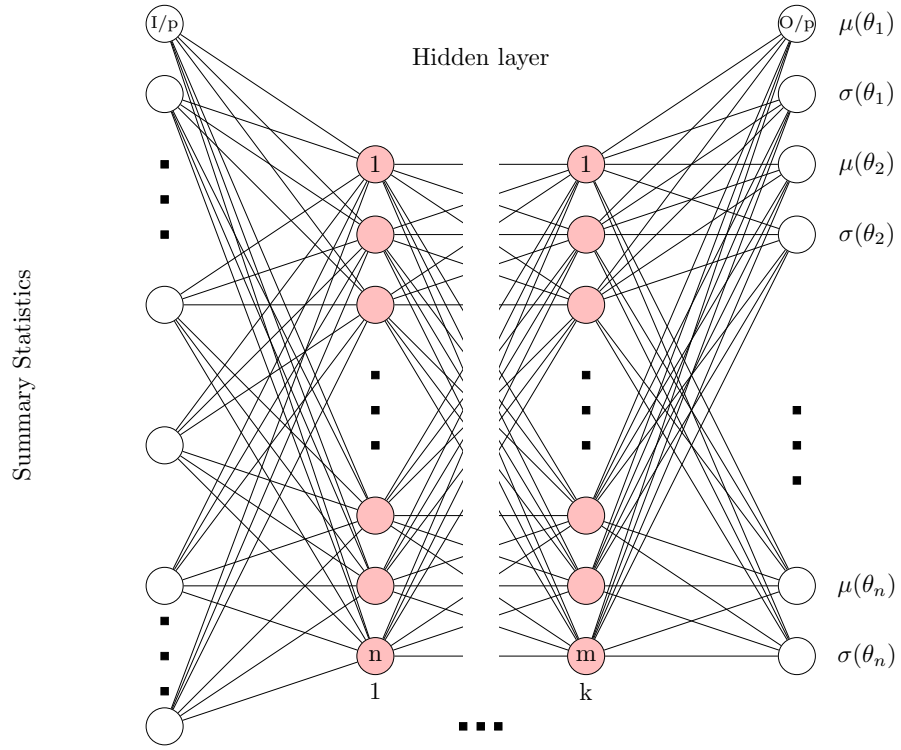

Figure 5.1: Moment Neural Network Architecture

Network Training and Architecture.

We perform our NN training in the "CloudVeneto" server using Nvidia V100 GPU. We start by splitting our data randomly into training, validation and test sets in the ratio of 60 %,20% and 20% respectively. This is done to avoid hidden correlations between the data from the same simulation that we do not want the network to learn. The input layer consists of various combinations of summary statistics and we are interested in their constraining capabilities. This is followed by a normalization layer and a variable number of hidden layers with varying numbers of nodes. The output layer concatenates two sets of variable mean and sigma described in a moment with the number of nodes equal to the number of parameters.
For the hidden layers, we use the Scaled Exponential Linear Unit (SELU) activation function [Klambauer et al., 2017]. The SELU activation function is defined as:

$$\text{SELU}(x) = \lambda \begin{cases} x & \text{if } x > 0 \\ \alpha e^x - \alpha & \text{if } x \leq 0 \end{cases} \tag{5.11}$$

where $\lambda \approx 1.0507$ and $\alpha \approx 1.67326$. The weights are initialized using the prescription in [He et al., 2015], which is commonly known as He initialization. An "L2" layer weight regularizer is used to prevent overfitting. We also apply dropout to each hidden layer [Srivastava et al., 2014] to further regularize the model.

The hidden layer is then followed by the output layer that concatenates the two sets of variables described in the moment, i.e., the mean and the sigma values of each parameter. To ensure that the sigma values are positive, we use the ELU+1 activation function [Jung et al., 2024a], which is defined as:

where ELU (Exponential Linear Unit) is defined as:

$$\text{ELU}(x) + 1 = \begin{cases} x + 1 & \text{if } x > 0 \\ e^x & \text{if } x \leq 0 \end{cases} \tag{5.12}$$

For the means, we use a linear activation function, which is simply:

$$\text{Linear}(x) = x \tag{5.13}$$

The weights are initialized similarly to the hidden layers along with the "L2" layer weight regularizer.

Now, The training is performed using the Adam optimizer [Kingma and Ba, 2017], with a cyclical learning rate [Smith, 2017]. We chose to use a cyclic learning rate as it helps the optimizer avoid falling into the local minima and saddle points by periodically varying the learning rate between a lower and upper bound. So, Instead of using a fixed learning rate or a monotonically decreasing schedule, CLR allows the learning rate to increase and decrease in a cyclical fashion.

The neural network architecture includes several hyperparameters that were optimized using Bayesian optimization [O'Malley et al., 2019]. This technique explores different values within specified ranges to find the best-performing configuration. We performed 50 trials for hypertuning the NN model with Mν and 30 trials for f_{NL}^{equil}. The decision to use fewer trials for f_{NL}^{equil} was primarily driven by the need to avoid extensive computational time. The hyperparameters and their ranges are as follows:

- **Number of Layers** : The number of hidden layers varies between 1 and 8.
- **Architecture** : The number of nodes in each hidden layer ranges from 8 to 1024, with increments of 32.
- **Dropout Rate** : The dropout rate ranges from 0.1 to 0.5, sampled logarithmically.
- **Learning Rate**: The learning rate ranges from 10^{-5} to 10^{-2}, sampled logarithmically.
- **Regularization Rate** : The "L2" regularization rate ranges from 10^{-5} to 10^{-1}, sampled logarithmically.
- **Cyclic Base Learning Rate** : The base learning rate for cyclical learning ranges from 10^{-5} to 10^{-2}, sampled logarithmically.
- **Batch Size**: The batch size can take values of $[16, 32, 64, 128, 256, 512]$.

The network is then trained and to evaluate our neural network's (NN) learning capacity, we examine the "training" vs "validation loss" plots shown in Figure 6.1. The training loss indicates how effectively the model performs on the training data, while the validation loss measures its performance on a separate dataset, known as the validation set. This validation set helps us evaluate how well the model generalizes to unseen data, beyond its training data. We stop the training when the validation loss does not decrease for 100 epochs.

5.3.4 Evaluators

To evaluate the quality of moment predictions.

- Firstly, we calculate the coefficient of determination for each parameter defined as :

$$R^2(\hat{\theta}) \equiv 1 - \frac{\sum_i \left(\theta_i - \hat{\theta}(x_i)\right)^2}{\sum_i \left(\theta_i - \bar{\theta}\right)^2}, \tag{5.14}$$

  where i runs over simulations in the test set, θ_i represents the true parameter for the i-th simulation, $\bar{\theta}$ is the average of the true parameter across the test set, and $\hat{\theta}(x_i)$ is the posterior mean estimate from the i-th simulation.
  An R^2 value close to 1 indicates that the mean posterior estimate of the parameters is close to the average of the true value of the parameter hence it indicates a close recovery of the true parameters, while $R^2 = 0$ suggests that using the average value as a prediction is as good as the estimator. A thing to note is that R^2 can be negative if the estimator performs worse than simply using the average value. The coefficient of determination assesses the quality of the posterior mean estimates.

- To evaluate the second moment, we use a different metric χ^2. For each parameter, we compute:

$$\chi^2(\hat{\theta}, \hat{\sigma}) \equiv \frac{1}{N} \sum_i \frac{\left(\theta_i - \hat{\theta}(x_i)\right)^2}{\hat{\sigma}^2(x_i)}, \tag{5.15}$$

  where $\hat{\sigma}(x_i)$ is the predicted standard deviation for the i-th simulation, and N is the number of simulations in the set. If the model overfits to get a reduced $\hat{\sigma}(x_i)$ the χ^2 shoots up and if it underestimates the sigma we get a χ^2, The closer to 1, the better the calibration. For each trained network, We calculate χ^2 using the simulations in the validation set and discard all those where $|\chi^2 - 1| > 0.5$.

Chapter 6

Results and Discussion

As we mentioned in the previous chapter one of the primary objectives of this thesis is to determine which statistics, or combinations of them, yield the greatest information about cosmological parameters at small, non-linear scales, as was indicated in the preceding chapter. To achieve this, we developed a neural network pipeline that uses the moment network approach to carry out likelihood-free inferences. As both $f_{NL}^{\rm equil}$ and M_ν show significant signals at small, non-linear scales, we tested our pipeline on two distinct datasets: Quijote-PNG, which has non-Gaussian initial conditions, and the Quijote dataset with massive neutrinos included to assess the efficacy of our pipeline.

This chapter is divided into two sections. Firstly we discuss our results of the model on $f_{\rm NL}^{\rm equil}$ LH which has non-gaussian initial conditions and then we discuss it for "nwLH" i.e latin-hypercube with massive neutrinos.

6.1 $f_{NL}^{\rm equil}$ LH

In order to evaluate our neural network's (NN) learning capacity, we examine the training vs validation loss plots shown in Figure 6.1.

We see a consistent pattern over the course of our tests with different combinations of summary statistics i.e. both the training and validation losses decrease over time. Most importantly, our model does not overfit the training data since the validation loss is constantly less than or comparable to the training loss. This analysis affirms that our neural network effectively learns from diverse combinations of summary statistics, representing varied training data, and demonstrates strong generalization to unseen data. This is evident from the consistent stabilization and convergence observed in both training and validation losses. The stability and convergence of these metrics across different sets of summary statistics underscore the model's ability to extract meaningful patterns and relationships from the data.

In Fig. 6.3 and Fig. 6.2 we present an overview of our primary finding. In Fig 6.3 we plot different evaluators for checking the accuracy of our model's ability to constrain the parameters on a test set, and in Fig. 6.2 we show the standard deviation prediction of the NN model. The power spectrum (Pk0), which is important to constrain the conventional cosmological parameters, is taken for granted in our study and we determine what additional single summary statistic may be added to it to increase further constraints on $f_{NL}^{\rm equil}$.

From Fig. 6.2 we can see that adding the marked power spectrum (MP0_0) to the power spectrum (Pk0) outperforms both Pk0 and Bk0.[1] Adding the marked power spectrum or marked bispectrum to the power spectrum and bispectrum does not result in any significant changes in the standard deviation.

As noted in [Jung et al., 2023, Jung et al., 2024a], the HMF significantly improves the constraints on cosmological parameters, especially $f_{NL}^{\rm equil}$. Based on this insight, we included HMF in our analysis. We observed that incorporating HMF into any combination enhances the accuracy of our constraints.

[1] We found that the mark with the smallest smoothing scale (R = 20, p = 1, $\delta_s = 0.5$) gives results as good as the combination of the three others, which was also noted in [Jung et al., 2024a].

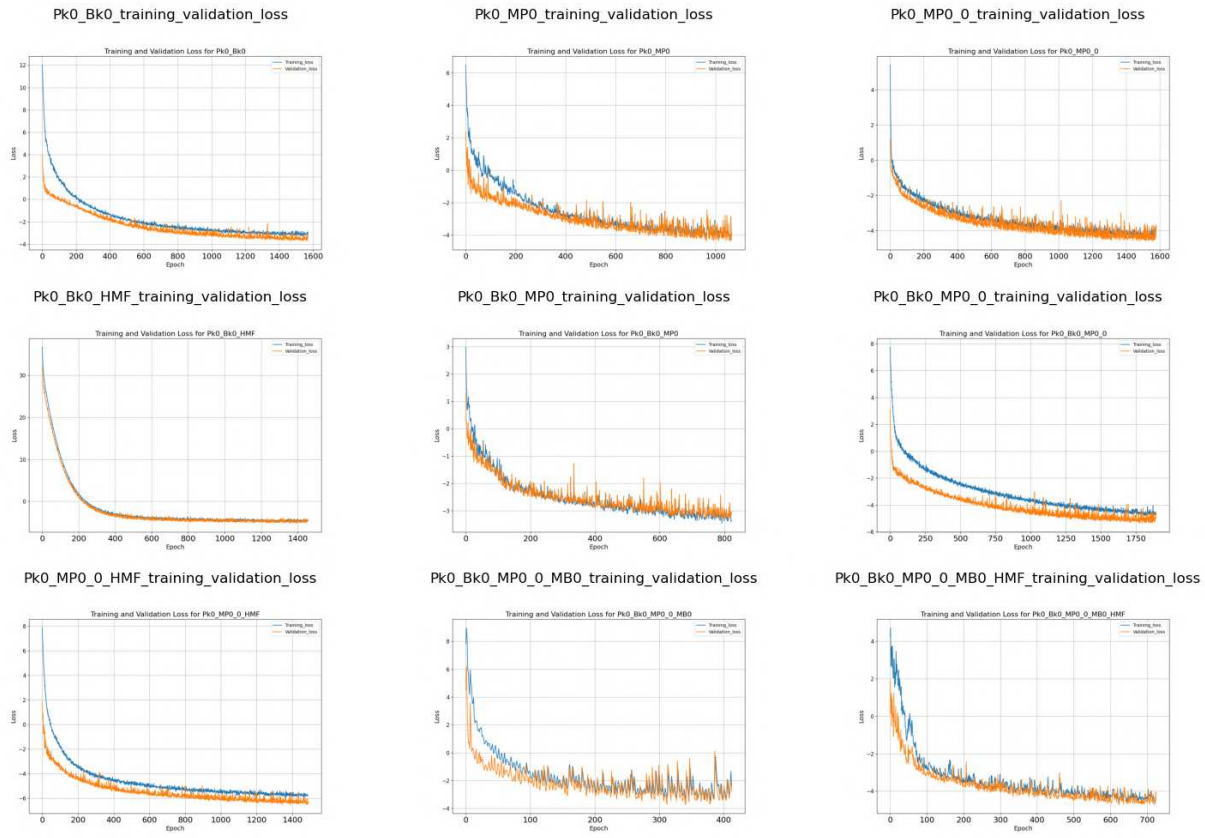

Figure 6.1: Loss plot for a combination of different summary statistics trained in f_{NL}^{equil} LH

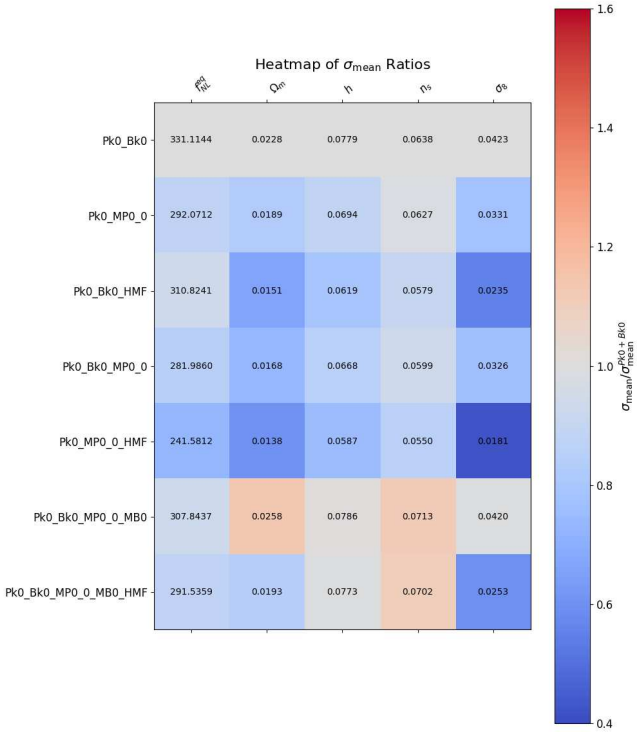

Figure 6.2: Standard deviation prediction of the NN model for different combinations of summary statistics in f_{NL}^{equil} LH. The colour scale gives the ratio of each error with respect to its equivalent computed using only the power spectrum and bispectrum.

Notably, the best-performing combination of statistics is Pk0 + MP0_0 + HMF, achieving a standard deviation of 241.58 on $f_{NL}^{\rm equil}$.

Now, let's examine Fig. 6.3, where we sometimes notice a drop in accuracy when adding more observables. To understand this, it is important to recall that better accuracy generally requires either a large set of input features for the model to learn from or highly informative features. In our case, we are limited by the available training data, and having a larger training set would improve the accuracy in constraining the parameters. Additionally, some statistics are strongly correlated, so adding new features to the existing ones may not provide additional relevant information. This redundancy can negatively impact the training, resulting in a drop in accuracy, which we observe in the case of the standard deviation predictions.

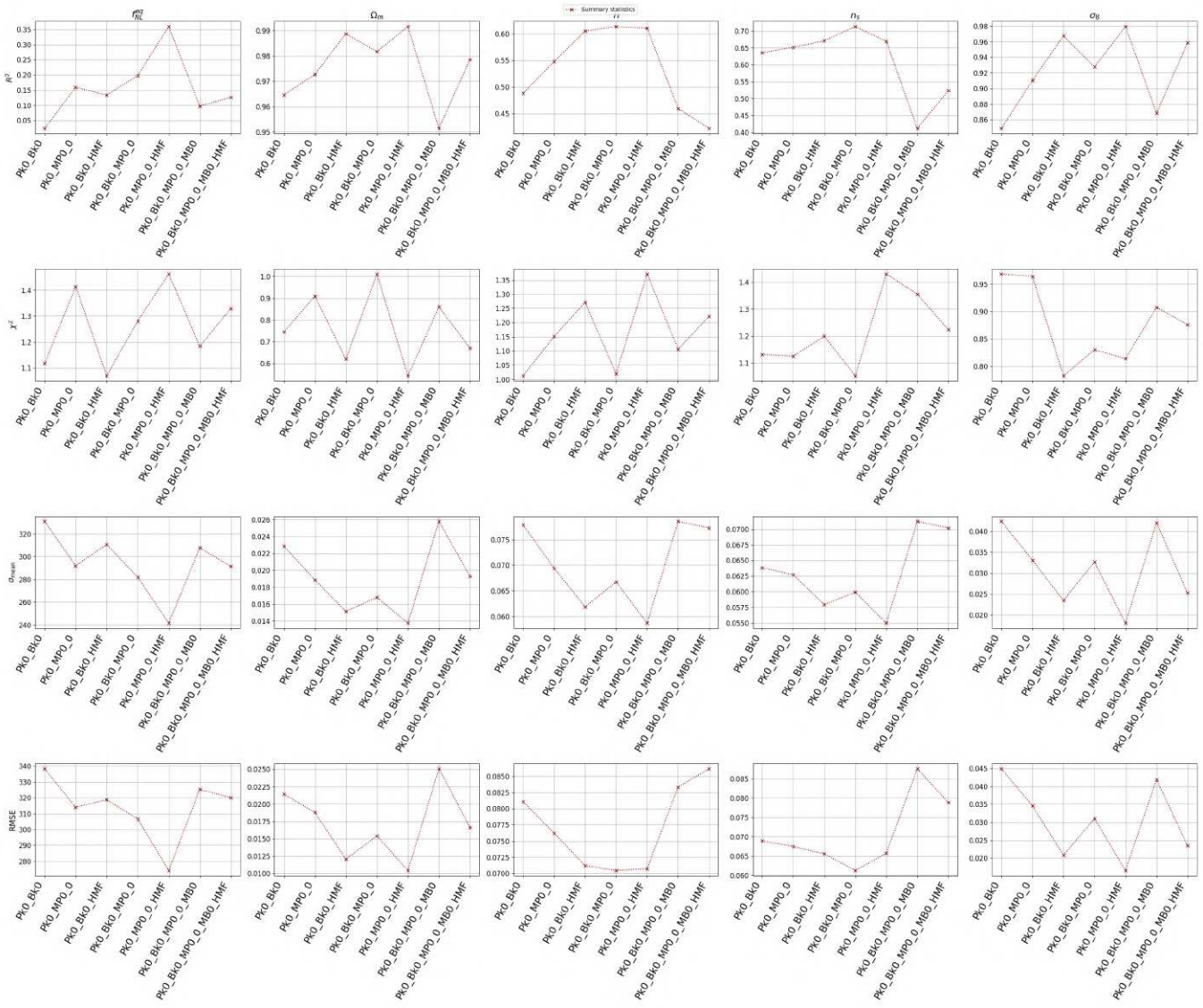

Figure 6.3: Different evaluators for assessing the model's accuracy in predicting $f_{NL}^{\rm equil}$, Ω_m, h, n_s, and σ_8 parameters using combinations of power spectrum (Pk0), bispectrum (Bk0), marked power spectrum with smallest smoothing scale (MP0_0), combination of all marked power spectra (MP0), and combination of all marked bispectra (MB0). Each column represents a specific combination.

In Fig.6.4 we show the predictions of different parameters from the NN which was trained on Pk0, MP0_0, and HMF and applied to the test $f_{NL}^{\rm equil}$ LH which shows the best constraints. We observe a very high R^2 value of 0.99 for Ω_m and 0.98 for σ_8 indicating that the model has very high accuracy in predicting these parameters. In the case of h, n_s, and $f_{NL}^{\rm equil}$ we see the scatter in the standard deviation is more considerable than Ω_m and σ_8 indicating the influence of other cosmological parameters in each simulation. The trend in the bias as a function of true value for h, n_s, and $f_{NL}^{\rm equil}$ indicates that the

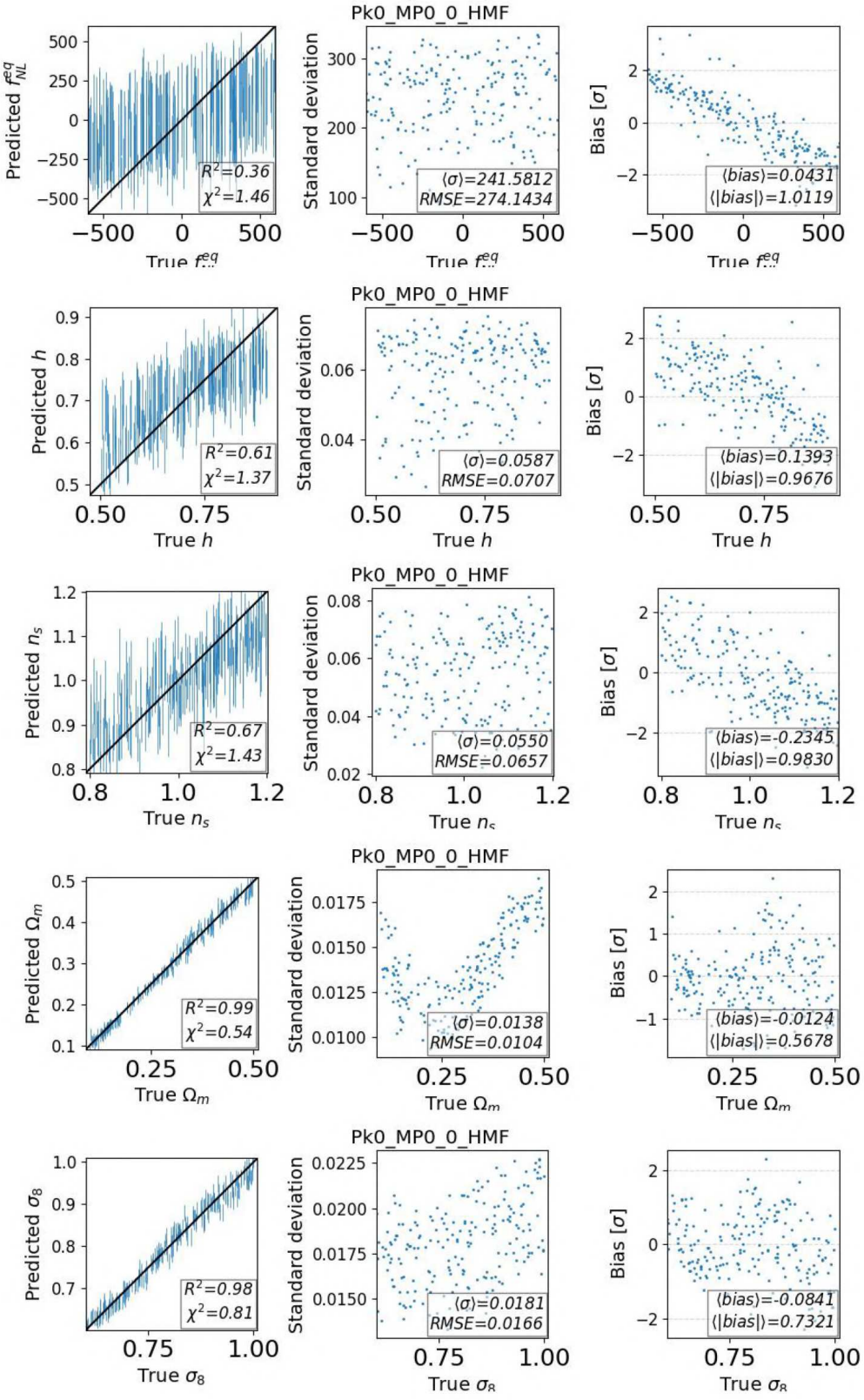

Figure 6.4: Predictions of parameters for the best-performing network trained on the combination of the Pk0, MP0_0, and HMF statistics, applied to the test set of the f_{NL} LH. Left panel — Comparison of the true and predicted values of Ω_m. The lines span the predicted 1σ about the predicted value. Centre panel — Predicted standard deviations as a function of the true parameter value. The mean standard deviation is also shown, compared with the root mean squared errors (RMSE) of the mean predictions. Right panel — Bias of the mean prediction in units of the predicted standard deviations.

predictions are bounded by the prior. If the true value is near the edge of the prior, the model struggles more with the prediction. The results for other combinations are presented in B.1.

6.2 nwLH

As we mentioned before massive neutrinos also show strong signals at small non-linear scales so we tried implementing our NN pipeline to LH data with massive neutrinos. We discuss our results below.

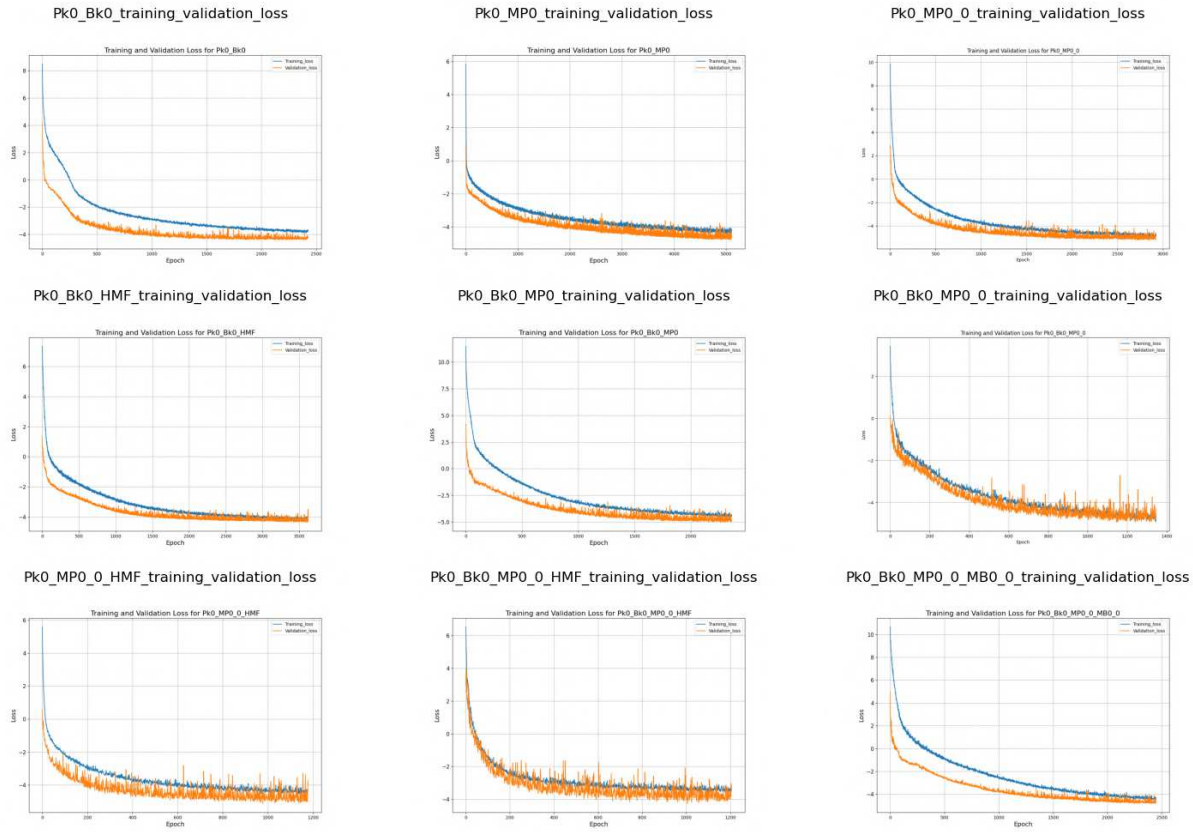

Figure 6.5: Loss plot for a combination of different summary statistics trained in nwLH

We see in Fig. 6.5 that even for the nwLH data set, our model learns well on the training set. The training and validation losses consistently decrease over time, indicating effective learning. Additionally, the validation loss remains lower than or comparable to the training loss, suggesting that the model is not overfitting and is generalizing well to unseen data. This pattern is consistent with the results observed from other data sets and combinations of summary statistics, further validating the robustness and reliability of our model.

In Fig. 6.6, we plot the standard deviation prediction of the NN for different combinations of summary statistics trained and tested on the nwLH dataset. We observe that adding marked statistics to the power spectrum, even in the case of massive neutrinos, provides better constraints than adding the bispectrum. The constraints are improved further when adding marked statistics with the smallest smoothing scale compared to the sum of all the marks, both when added to the power spectrum alone and to both the power spectrum and bispectrum. Additionally, we observe that adding the marked bispectrum does not significantly improve the constraints. However, adding the Halo Mass Function (HMF) improves the constraints on M_ν. Overall, the best constraint on M_ν is found with the combination of statistics Pk0 + Bk0 + MP0_0 + HMF, achieving a standard deviation of 0.2493. In Figure 6.7, we notice a pattern similar to what we saw with the equilateral shape of f_{NL} likelihood (LH) i.e Fig. 6.3, where accuracy decreases when certain observables are added. This trend shows that our ability to accurately predict outcomes is limited by the amount of data we have and how

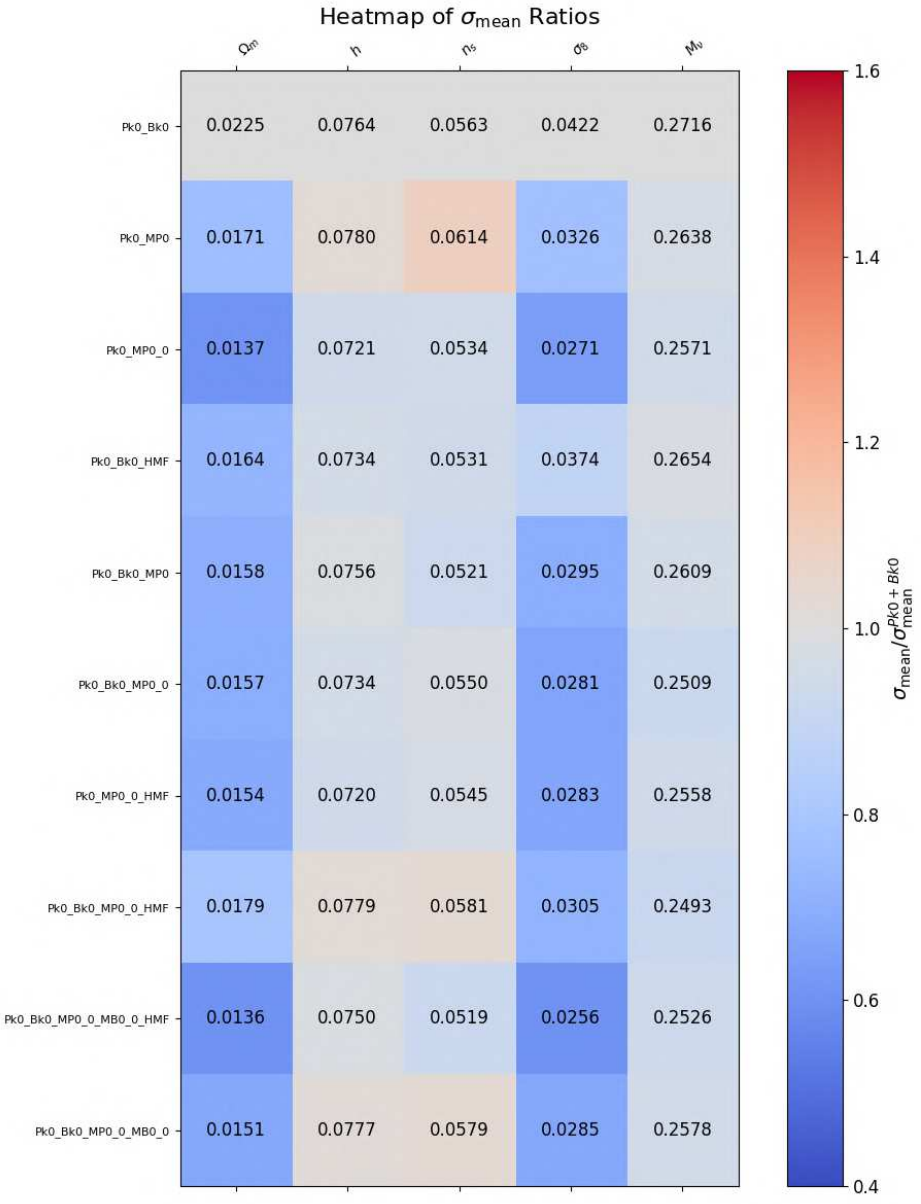

Figure 6.6: Standard deviation prediction of the NN model for different combinations of summary statistics in nwLH. The colour scale gives the ratio of each error with respect to its equivalent computed using only the power spectrum and bispectrum.

closely related the different statistics are. [Jung et al., 2024a] showed that increasing the number of samples results in improved accuracy. This suggests that if we gather more data, we might overcome the accuracy drop we observe and handle the challenges posed by statistical correlations and data limitations more effectively.

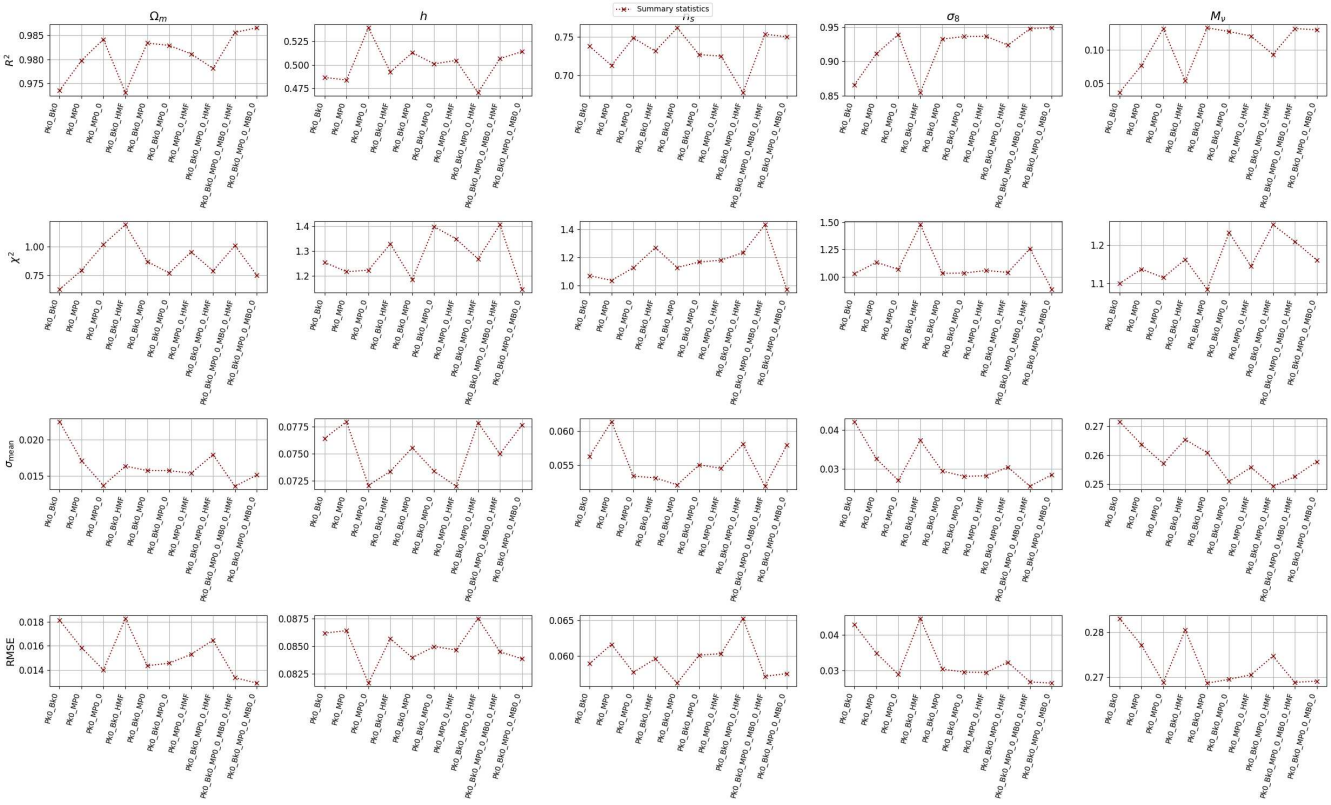

Figure 6.7: Different evaluators for assessing the model's accuracy in predicting f_{NL}^{equil}, Ω_m, h, n_s, and σ_8 parameters using combinations of power spectrum (Pk0), bispectrum (Bk0), marked power spectrum with smallest smoothing scale (MP0_0), combination of all marked power spectra (MP0), and combination of all marked bispectra (MB0). Each column represents a specific combination.

In Fig.6.8 we present the predictions made by our neural network (NN) model trained on Pk0, MP0_0, and HMF data, applied to the test case of f_{NL}^{equil} likelihood (LH). Notably, the accuracy of predicting the parameter M_ν is remarkably low. However, when we calculate the standard deviation based on the uniform prior we used for neutrino masses, which ranges from [0, 1] eV, we expect a standard deviation of ≈ 0.2887. Interestingly, our model demonstrates an improvement in the mean sigma, suggesting it is learning meaningful patterns from the data.

The poor accuracy observed could be attributed to firstly, the limited amount of available data. With insufficient data, the model may struggle to generalize effectively beyond the training set, leading to inaccurate predictions. Secondly, as discussed before for f_{NL}^{equil} we are limited by the prior which we can see from the bias trend where it is overestimating the sigma values for lower true values of M_ν and underestimation the higher true values. The results for other combinations of summary statistics trained in nwLh are presented in B.2.

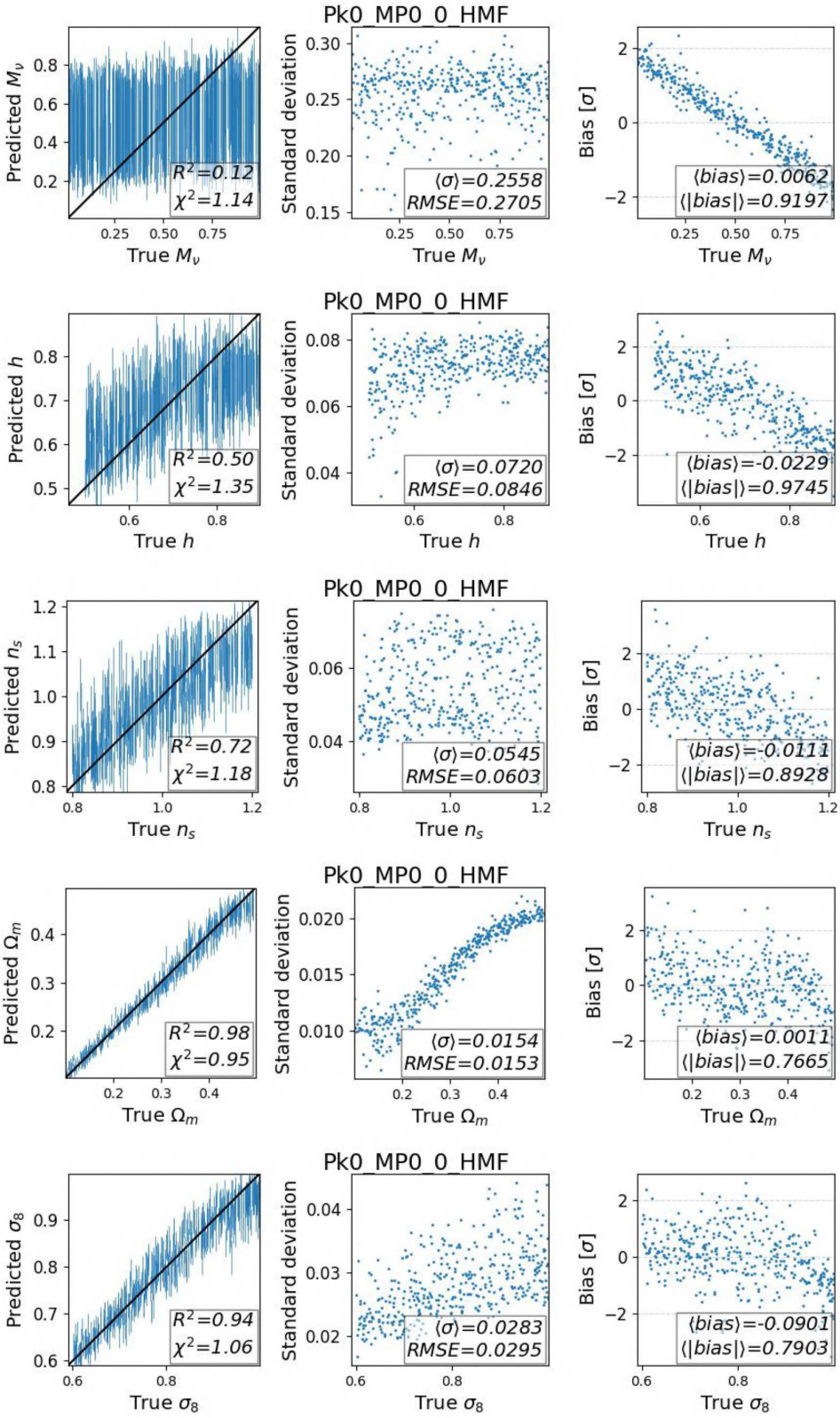

Figure 6.8: Predictions of parameters for the best-performing network trained on the combination of the Pk0, MP0_0, and HMF statistics, applied to the test set of the nwLH. Left panel — Comparison of the true and predicted values of Ω_m. The lines span the predicted 1σ about the predicted value. Centre panel — Predicted standard deviations as a function of the true parameter value. The mean standard deviation is also shown, compared with the root mean squared errors (RMSE) of the mean predictions. Right panel — Bias of the mean prediction in units of the predicted standard deviations.

Chapter 7

Conclusion

In this thesis, we developed a machine learning pipeline to perform Likelihood-Free Inference (LFI) for M_ν, $f_{NL}^{\rm equil}$, and other cosmological parameters using the moment network method. This approach avoids relying on assumptions or approximations of likelihood. Utilizing the Quijote and Quijote-PNG simulations, we identified various summary statistics and their combinations to constrain the sum of massive neutrinos M_ν, $f_{NL}^{\rm equil}$, and other cosmological parameters by extending the analysis to non-linear scales up to $k_{\rm max} = 0.5, h, {\rm Mpc}^{-1}$, where significant signals for both M_ν and $f_{NL}^{\rm equil}$ are expected.

We extracted the power spectrum, modal bispectrum, marked power spectrum, marked modal bispectrum, and halo mass function (HMF) in redshift space from both Quijote and Quijote-PNG halo catalogues. These were used to examine constraints on M_ν and $f_{NL}^{\rm equil}$, along with other cosmological parameters. Our analysis focused on the halo field, as halos are biased tracers of the underlying matter field and are sites of galaxy formation. This approach represents an initial step toward constraining cosmological parameters using actual observational data.

We found that incorporating marked statistics and the Halo Mass Function (HMF) can show promising potential for detecting $f_{NL}^{\rm equil}$ signals and provide a moderate contribution to the study of M_ν constraints from the upcoming large-scale structure (LSS) surveys. The best constraints were observed when adding the marked power spectrum and HMF to our analysis. While we experimented with different combinations of marks, we found that using the mark with the smallest smoothing scale, i.e., $R = 20, h^{-1}, {\rm Mpc}$, $p = 1$, and $\delta s = 0.50$, provided constraints equivalent to the combination of all marks. Adding a marked bispectrum, however, did not yield any additional constraints.

Although our neural network (NN) model demonstrated good learning capabilities, as evidenced by the training vs. validation loss plots in Fig. 6.1 and Fig. 6.5, the accuracy of our model predictions was low for $f_{NL}^{\rm equil}$ and M_ν due to the limited dataset and restrictive prior boundaries.

In the future, we aim to test our model on the matter field for M_ν, which has been shown to provide more constraints as indicated in [Massara et al., 2021]. Additionally, we plan to perform a Fisher analysis of the combination of different statistics on massive neutrinos for the halo field to further validate our current findings.

Appendices

Appendix A

Window Function

A window function is a mathematical tool which is introduced to treat the analytically random field as random fields are not differentiable. The smallest comoving scale of the universe is that of the typical separation between neighbouring galaxies of the order of 1 Mpc. To exclude scales smaller than $R\,(r < R \text{ or } k > R^{-1})$ we can filter the density field with a window function. This can be done in k-space or x-space. The filtering in x-space is done by convolution. We introduce a (usually spherically symmetric) window function $W(\mathbf{r}; R)$ such that

$$\int d^d r W(\mathrm{r}; R) = 1 \tag{A.1}$$

(normalization) and $W \sim 0$ for $|\mathbf{r}| \gg R$ and define the filtered density field

$$\delta(\mathrm{x}; R) \equiv (\delta * W)(\mathrm{x}) \equiv \int d^d \mathrm{x}' \delta\left(\mathrm{x}'\right) W\left(\mathrm{x} - \mathrm{x}'\right) \tag{A.2}$$

Here, the desired resolution is indicated by R, and the functions $\delta(\mathrm{x}; R)$ and $W(\mathrm{x}; R)$ are taken into consideration. We will refer to the following notation as $W(\mathrm{x}; R)$, keeping the scale R implicit. We can now write simply $W(r)$ because we also assume that W is spherically symmetric.

Denote the Fourier coefficients of $\delta(\mathrm{x}; R)$ by $\delta_{\mathrm{k}}(R)$. We use the Fourier series for $\delta(\mathrm{x})$ and $\delta(\mathrm{x}; R)$, however, since $W(r)$ vanishes for big r, we can use $\tilde{W}(k)$ instead, which is a Fourier transform. Thus, a mixed form of the convolution theorem is required. Let's state it clearly:

$$\delta_{\mathbf{k}}(R) = \frac{1}{V} \int_V d^d x \delta(\mathbf{x}; R) e^{-i\mathbf{k}\cdot\mathbf{x}} = \frac{1}{V} \int_V d^d x d^d x' \delta\left(\mathbf{x}'\right) W\left(\mathbf{x} - \mathbf{x}'\right) e^{-i\mathbf{k}\cdot\mathbf{x}} \tag{A.3}$$

$$= \frac{1}{V} \int_V d^d x' \delta\left(\mathbf{x}'\right) e^{-i\mathbf{k}\cdot\mathbf{x}'} \int d^d r W(\mathbf{r}) e^{-i\mathbf{k}\cdot\mathbf{r}} = \tilde{W}(\mathbf{k}) \delta_{\mathbf{k}}, \tag{A.4}$$

where

$$\tilde{W}(\mathbf{k}) = \int d^d r W(\mathbf{r}) e^{-i\mathbf{k}\cdot\mathbf{r}} \tag{A.5}$$

is the Fourier transform of $W(\mathbf{r})$ With our normalization, $W(\mathrm{r})$ has dimension $1/V$ and $\tilde{W}(\mathbf{k})$ is dimensionless with $\tilde{W}(\mathbf{k} = 0) = 1$. Since $W(\mathbf{r}) = W(r)$ is spherically symmetric, so is $\tilde{W}(\mathbf{k}) = \tilde{W}(k)$. Since $W(-\mathbf{r}) = W(\mathbf{r}), \tilde{W}(\mathbf{k})$ is real. For the correlations of these filtered Fourier coefficients, we get

$$\langle \delta^*_{\mathbf{k}}(R) \delta_{\mathbf{k}'}(R) \rangle = \tilde{W}(\mathbf{k})^* \tilde{W}\left(\mathbf{k}'\right) \langle \delta_{\mathbf{k}^*} \delta_{\mathbf{k}'} \rangle = \frac{1}{V} \delta_{\mathbf{k}\mathbf{k}'} \tilde{W}(k)^2 P(k) \tag{A.6}$$

so the filtered power spectra are

$$\tilde{W}(k)^2 P(k) \quad \text{and} \quad \tilde{W}(k)^2 \mathscr{P}(k) \tag{A.7}$$

The filtered correlation function is

$$\xi(\mathbf{r};R) \equiv \langle \delta(\mathbf{x};R)\delta(\mathbf{x}-\mathbf{r};R)\rangle = \frac{1}{(2\pi)^d}\int d^d k e^{i\mathbf{k}\cdot\mathbf{r}}\tilde{W}(k)^2 P(k) \tag{A.8}$$

and the variance of the filtered density field is

$$\sigma^2(R) \equiv \left\langle \delta(\mathrm{x};R)^2 \right\rangle = \xi(0;\mid R) = \int_0^\infty \tilde{W}(k)^2 \mathscr{P}(k)\frac{dk}{k} \tag{A.9}$$

The simplest window function is the top-hat window function

$$W_T(\mathrm{r}) \equiv \frac{1}{V(R)}\Theta\left(1-\frac{r}{R}\right) \tag{A.10}$$

and $W_T(\mathbf{r}) = 0$ elsewhere, i.e., $\delta(\mathrm{x})$ is filtered by replacing it with its mean value within the distance R.Its Fourier transformation is

$$\tilde{W}_R^{\mathrm{T}}(k) = 3\frac{\sin(kR)-(kR)\cos(kR)}{k^3R^3}, \tag{A.11}$$

Mathematically more convenient is the Gaussian window function

$$W_G(r) \equiv \frac{1}{V_G(R)}e^{-\frac{1}{2}r^2/R^2} \tag{A.12}$$

Where

$$V_G(R) \equiv \int d^d r e^{-\frac{1}{2}|\mathbf{r}|^2/R^2} \tag{A.13}$$

is the volume of W_G. The volume of a window function is defined as what $\int d^d r W(\mathrm{r})$ would be if W were normalized so that $W(0) = 1$, instead of the normalization we chose in A.1. For the top hat, this is the volume over which the filter averages; for others, a generalization of this. The volume of W_G is

$$V_G(R) = (2\pi)^{d/2}R^d \tag{A.14}$$

$$\tilde{W}_R^{\mathrm{G}}(k) = e^{-\frac{1}{2}(kR)^2} \tag{A.15}$$

There are also other choices available apart from these two.

In principle, there could be some ambiguity on which field has to be smoothed out if the radiation field δ or the curvature field ζ. However, in linear order, the two possibilities are equivalent, in fact, if we apply the smoothing procedure to Eq. (**??**) we obtain

$$\begin{aligned}
\delta_R(\mathbf{x}) =& \int d^3y W_R(|\mathbf{x}-\mathbf{y}|)\delta(\mathbf{y}) && \text{(A.16)}\\
=& \frac{2(1+w)}{5+3w}\frac{1}{a^2H^2}\int d^3y W_R(|\mathbf{x}-\mathbf{y}|)\nabla^2_{\mathbf{y}}\zeta(\mathbf{y}) && \text{(A.17)}\\
=& \frac{2(1+w)}{5+3w}\frac{1}{a^2H^2}\int d^3y \left\{\zeta(\mathbf{y})\nabla^2_{\mathbf{y}}W_R(|\mathbf{x}-\mathbf{y}|)\right. && \text{(A.18)}\\
& \left. +\nabla_{\mathbf{y}}\cdot[W_R(|\mathbf{x}-\mathbf{y}|)\nabla_{\mathbf{y}}\zeta(\mathbf{y}) - \zeta(\mathbf{y})\nabla_{\mathbf{y}}W_R(|\mathbf{x}-\mathbf{y}|)]\right\}, && \text{(A.19)}
\end{aligned}$$

where we have used the relation $\psi\nabla^2\phi = \phi\nabla^2\psi + \nabla\cdot(\psi\nabla\phi - \phi\nabla\psi)$ between two generic scalar fields $\psi(\mathbf{x})$ and $\phi(\mathbf{x})$. The second term in the integrand is a surface contribution and vanishes using the divergence theorem, under the fairly general assumption that W_R and its derivative vanish at large

scales. Furthermore, at least for the two window functions in equations A.10 and A.12, we have that $\nabla^2_{\mathbf{y}} W_R = \nabla^2_{\mathbf{x}} W_R$, therefore the above equation reads as

$$\delta_R(\mathbf{x}) = \frac{2(1+w)}{5+3w}\frac{1}{a^2H^2}\int d^3y \zeta(\mathbf{y})\nabla^2_{\mathbf{x}} W_R(|\mathbf{x}-\mathbf{y}|) = \frac{2(1+w)}{5+3w}\frac{1}{a^2H^2}\nabla^2_{\mathbf{x}}\zeta_R(\mathbf{x}), \tag{A.20}$$

which is the smoothed version of Eq.(3.30), as we expected. This proves the equivalence between smoothing out the density or the curvature field at a linear level.

In the end, the complete relation between the smoothed density field and curvature perturbation reads as

$$\delta_R(\mathbf{k},\tau) = -\frac{2(1+w)}{5+3w}\frac{k^2}{a^2H^2}T(k,\tau)\tilde{W}_R(k)D_1^{(+)}(\tau)\zeta(\mathbf{k}) = \frac{2}{3}T(k,\tau)\tilde{W}_R(k)D_1^{(+)}(\tau)\Phi(\mathbf{k}) \tag{A.21}$$

Appendix B

Additional Figures

B.1 f_{NL}^{equil} LH summary statistics combination plots

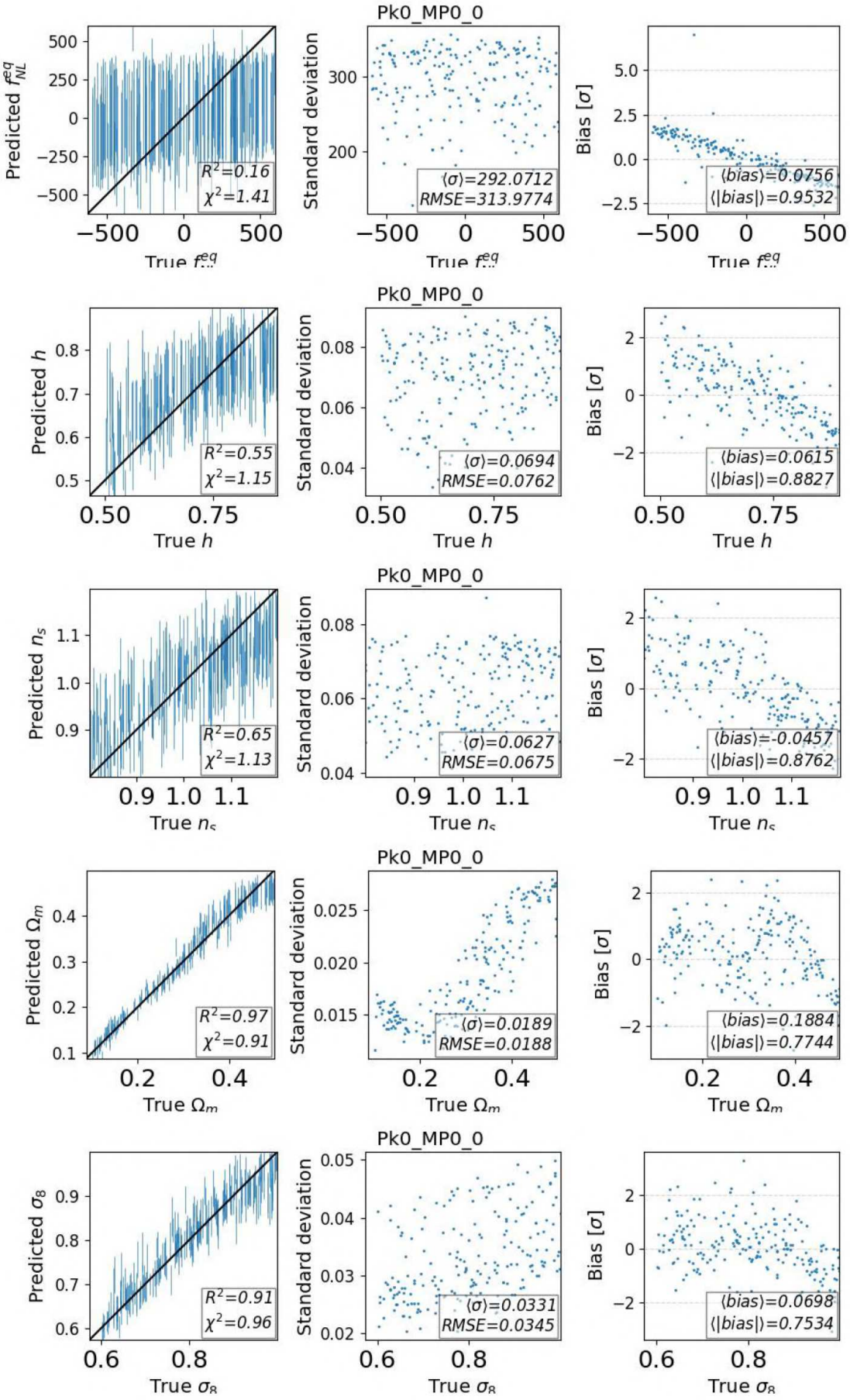

Figure B.1: Predictions of parameters for the best-performing network trained on the combination of Pk0˙MP0˙0 statistics, applied to the test set of the f_{NL}^{equil} LH.

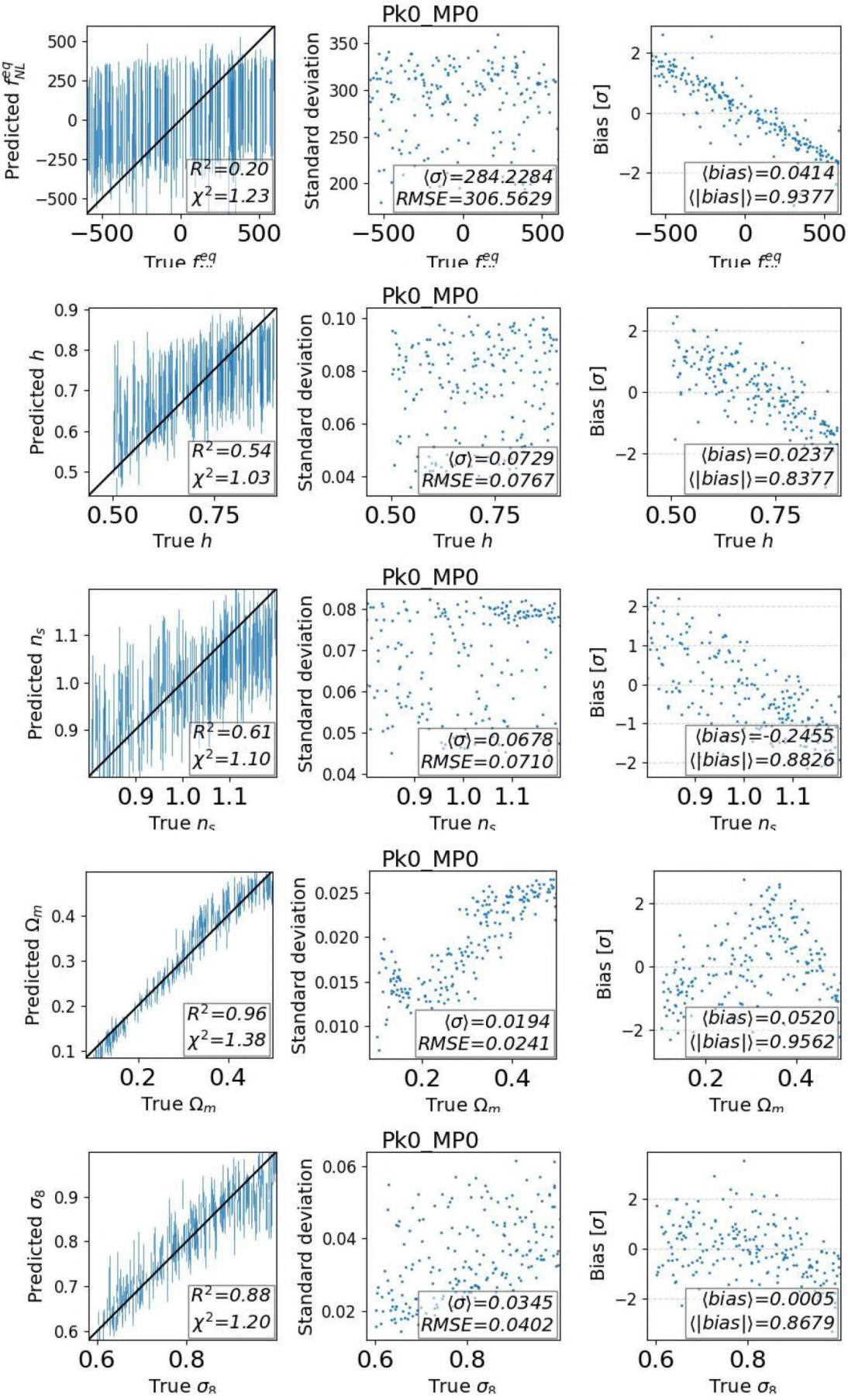

Figure B.2: Predictions of parameters for the best-performing network trained on the combination of Pk0˙MP0 statistics, applied to the test set of the f_{NL}^{equil} LH.

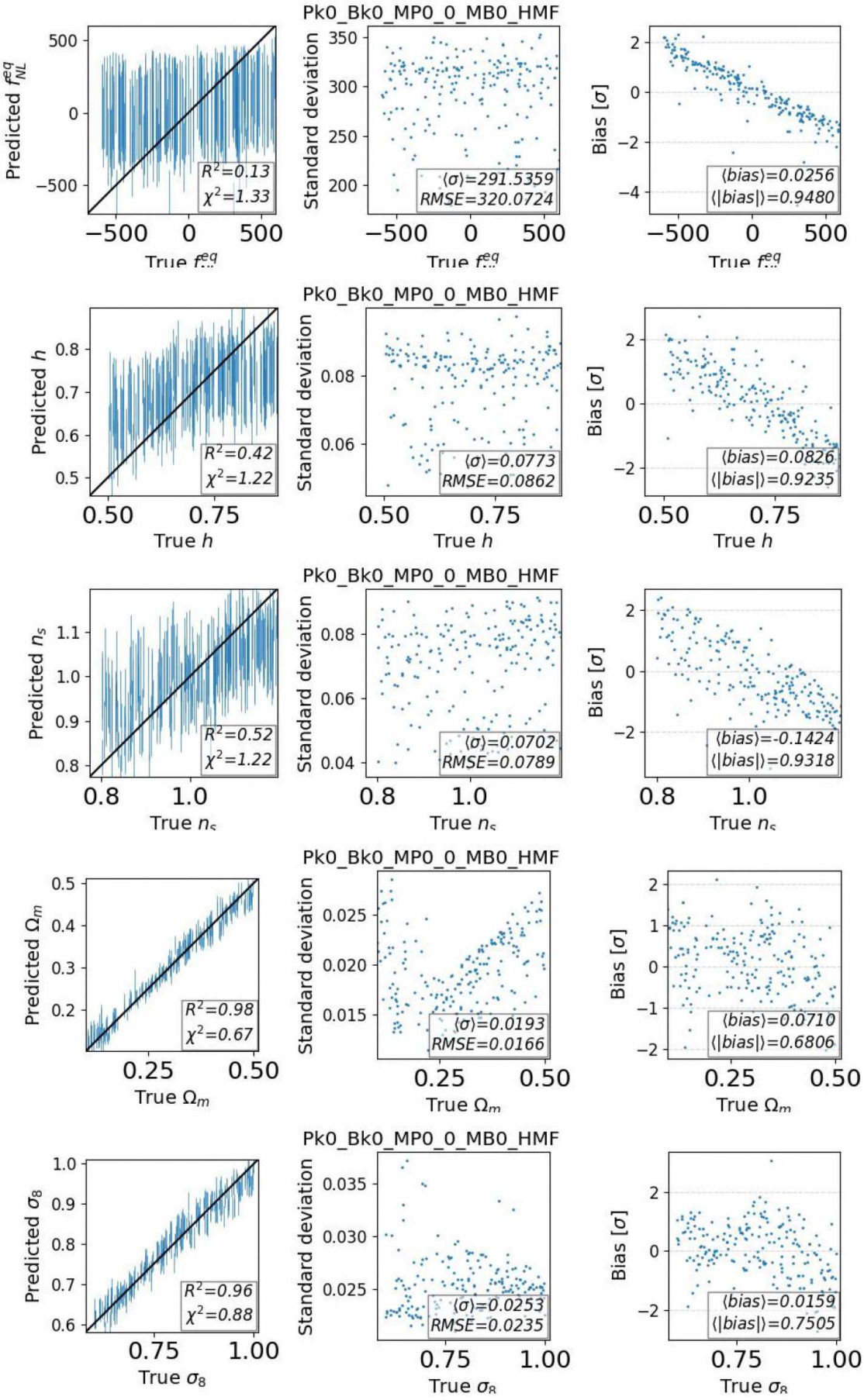

Figure B.3: Predictions of parameters for the best-performing network trained on the combination of Pk0˙Bk0˙MP0˙0˙MB0˙HMF statistics, applied to the test set of the f_{NL}^{equil} LH.

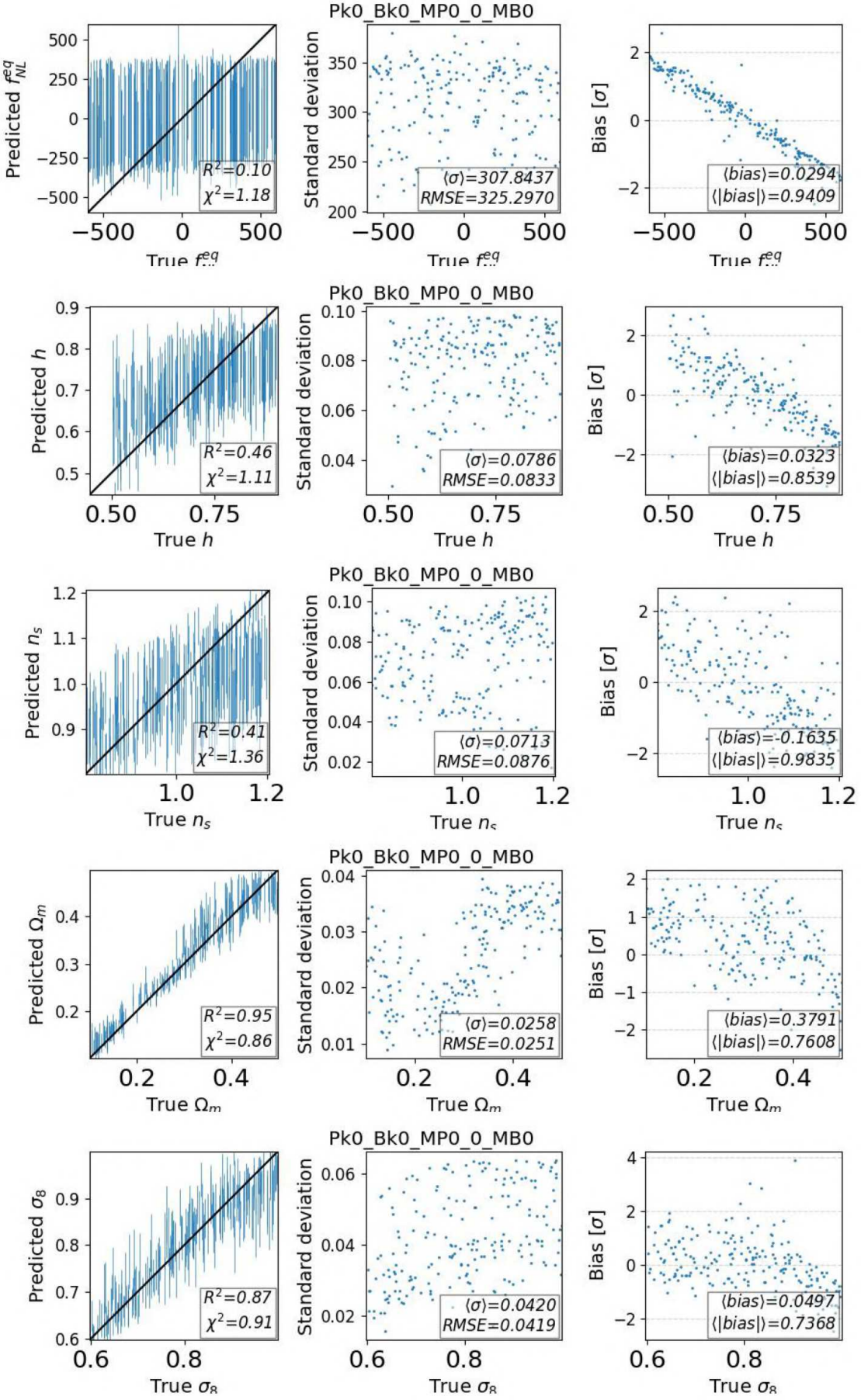

Figure B.4: Predictions of parameters for the best-performing network trained on the combination of Pk0˙Bk0˙MP0˙0˙MB0 statistics, applied to the test set of the f_{NL}^{equil} LH.

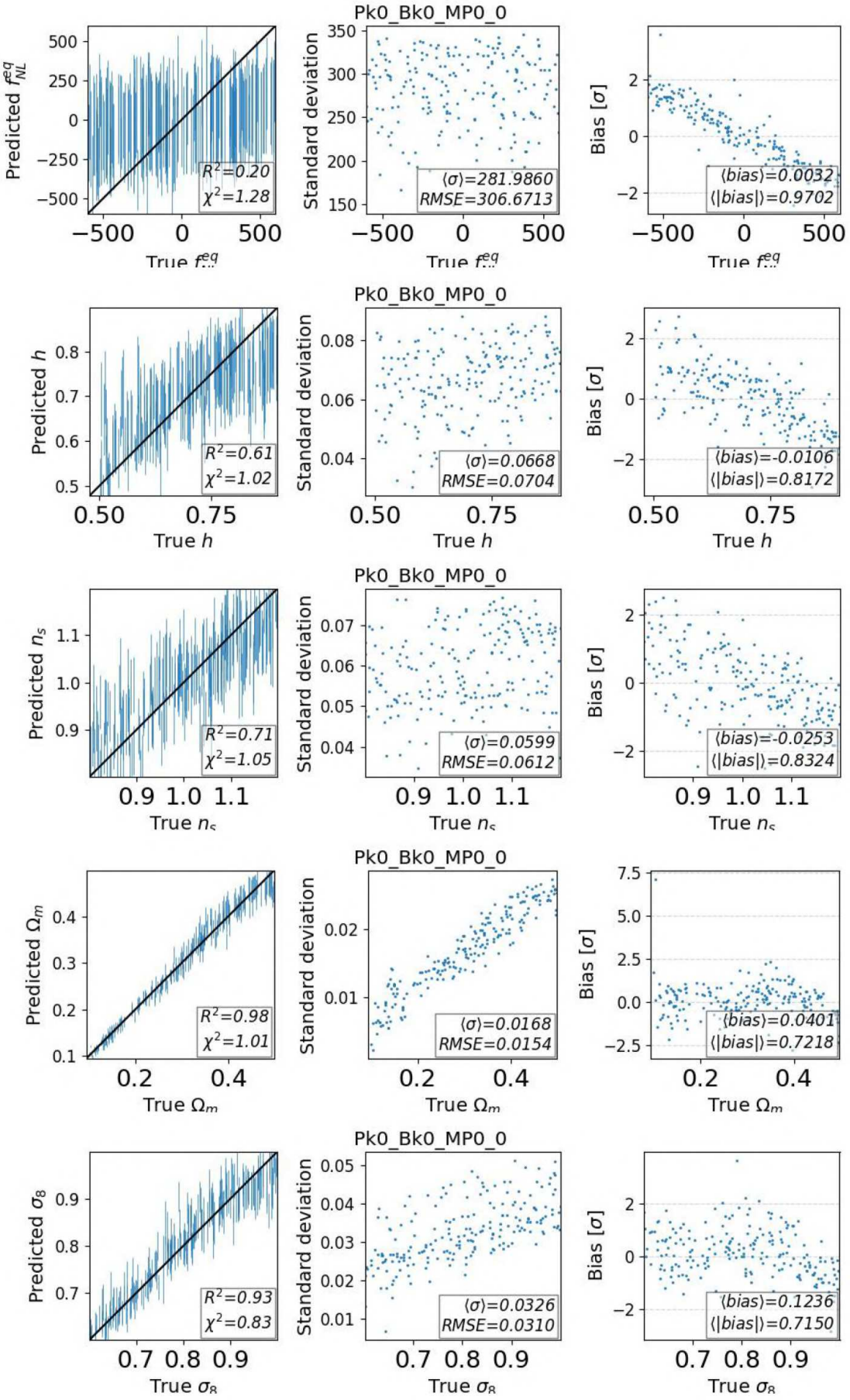

Figure B.5: Predictions of parameters for the best-performing network trained on the combination of Pk0˙Bk0˙MP0˙0 statistics, applied to the test set of the f_{NL}^{equil} LH.

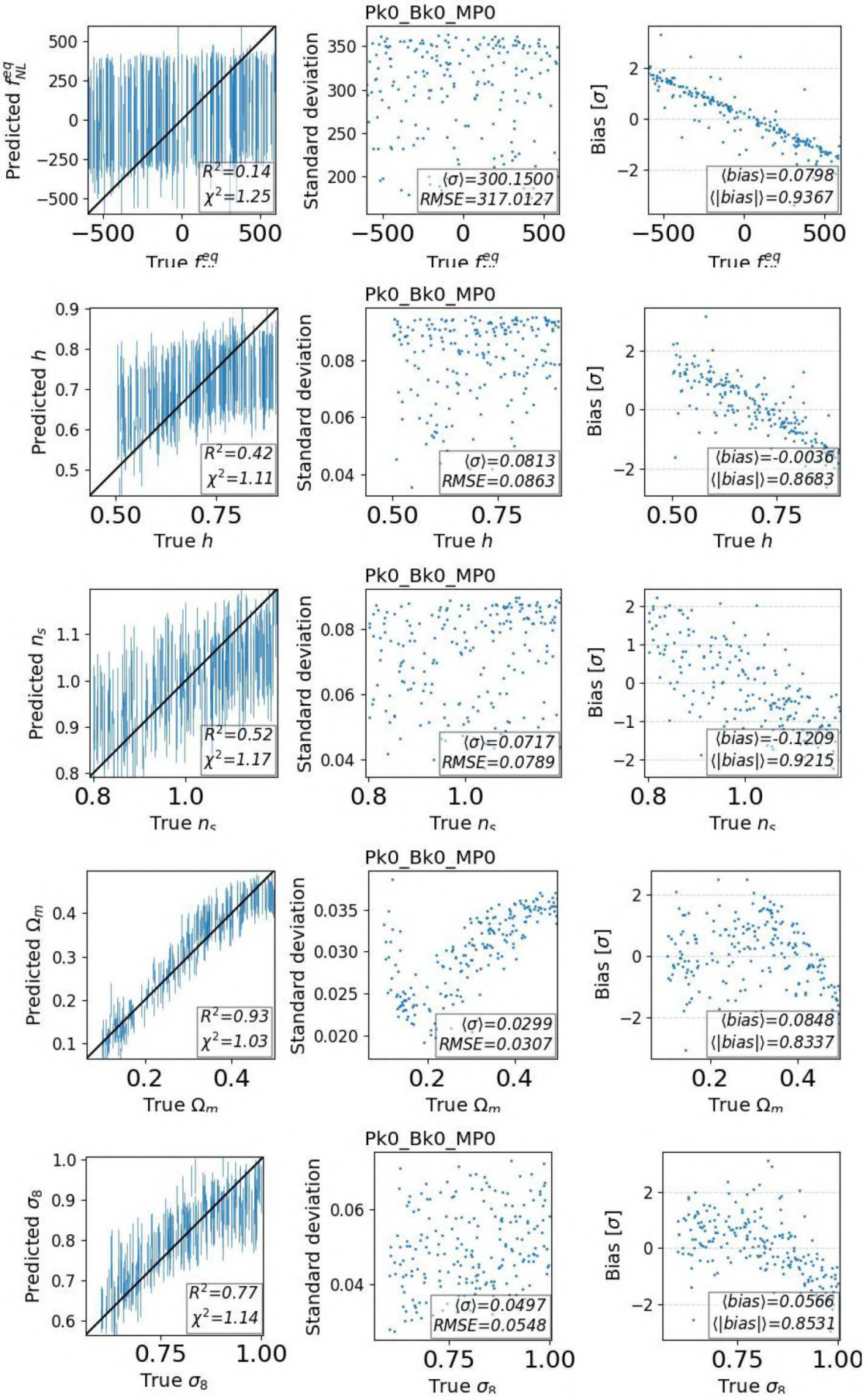

Figure B.6: Predictions of parameters for the best-performing network trained on the combination of Pk0˙Bk0˙MP0 statistics, applied to the test set of the f_{NL}^{equil} LH.

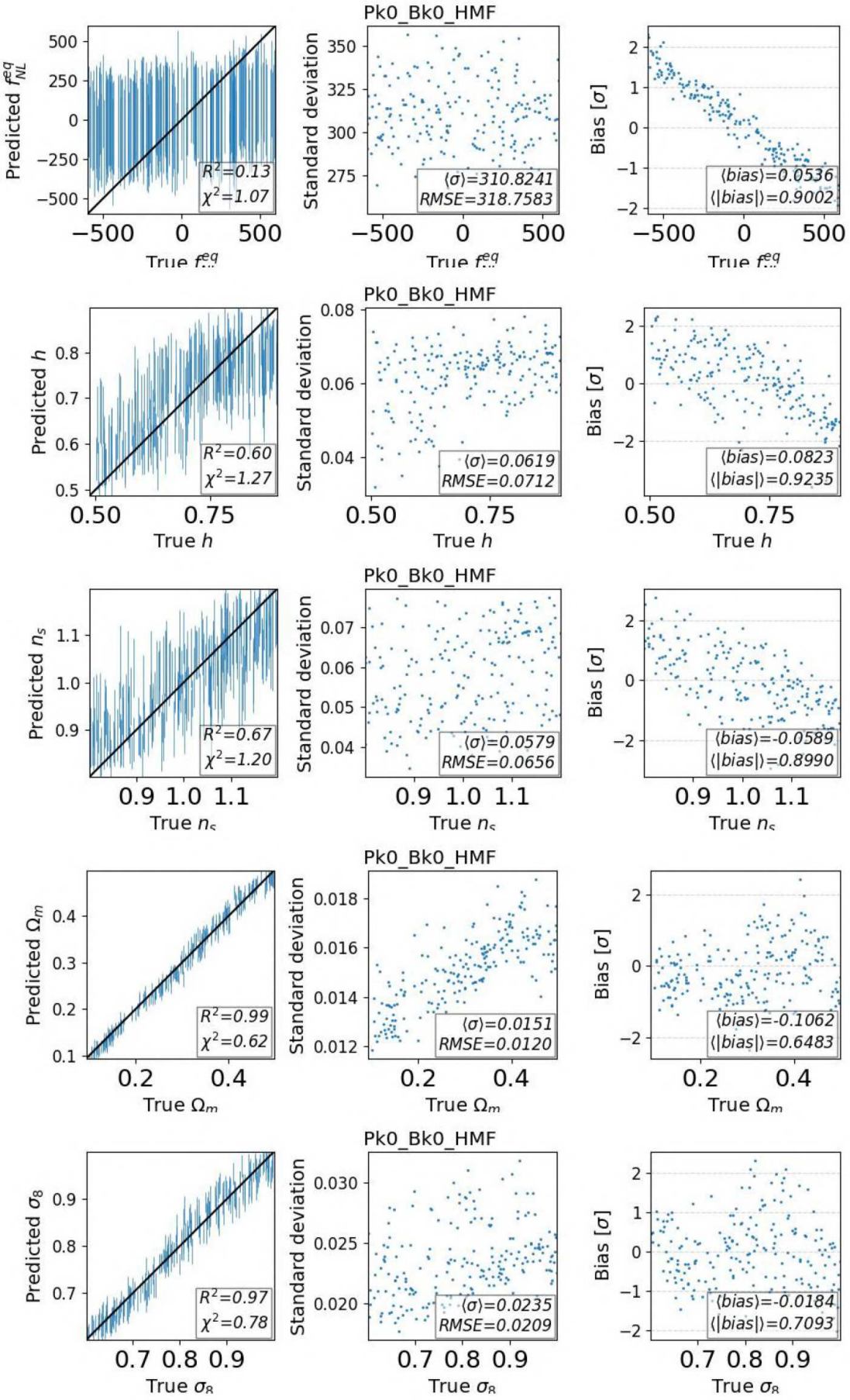

Figure B.7: Predictions of parameters for the best-performing network trained on the combination of Pk0˙Bk0˙HMF statistics, applied to the test set of the f_{NL}^{equil} LH.

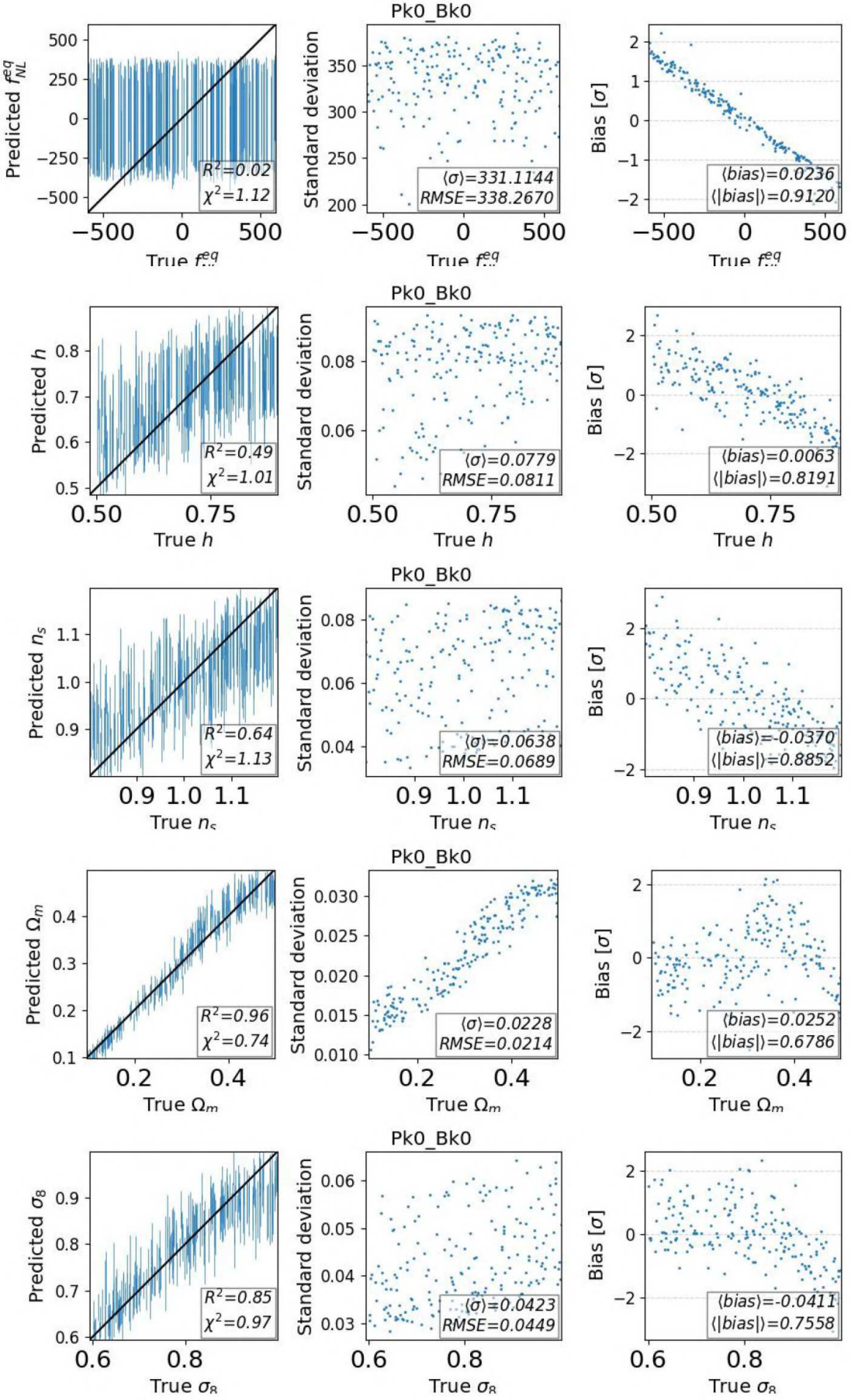

Figure B.8: Predictions of parameters for the best-performing network trained on the combination of Pk0˙Bk0 statistics, applied to the test set of the f_{NL}^{equil} LH.

B.2 nwLH summary statistics combination plots

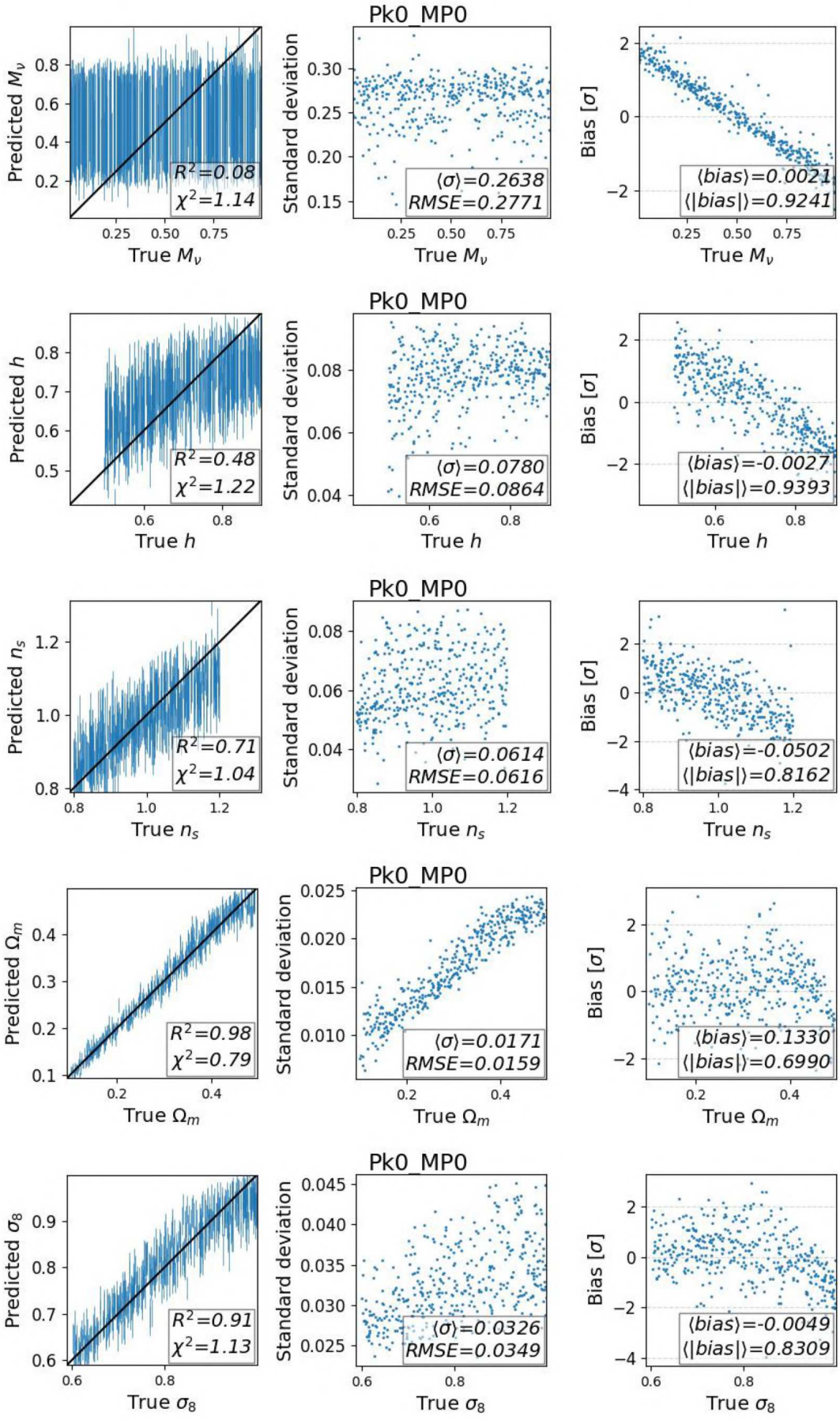

Figure B.9: Predictions of parameters for the best-performing network trained on the combination of Pk0˙MP0 statistics, applied to the test set of the nwLH.

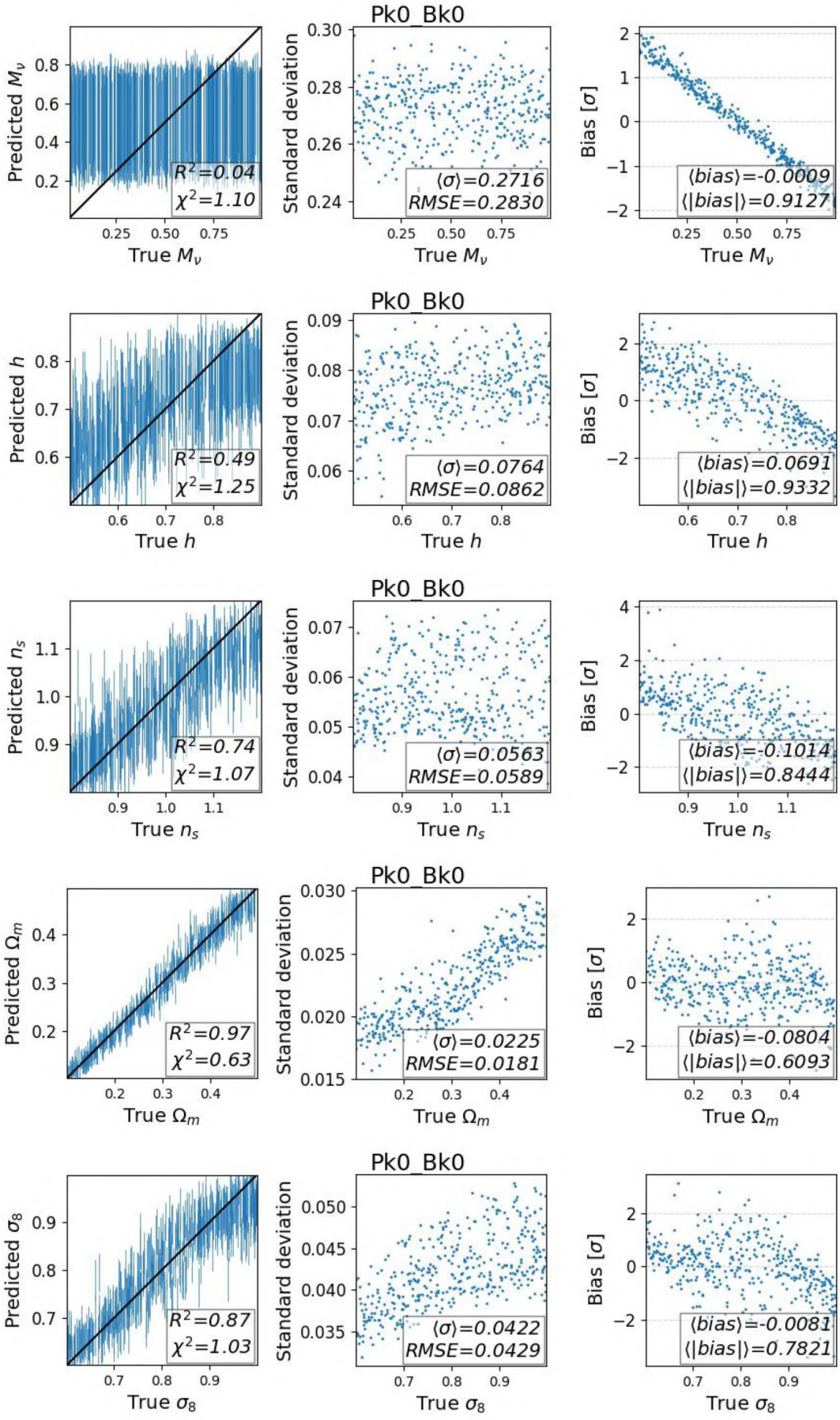

Figure B.10: Predictions of parameters for the best-performing network trained on the combination of Pk0˙Bk0 statistics, applied to the test set of the nwLH.

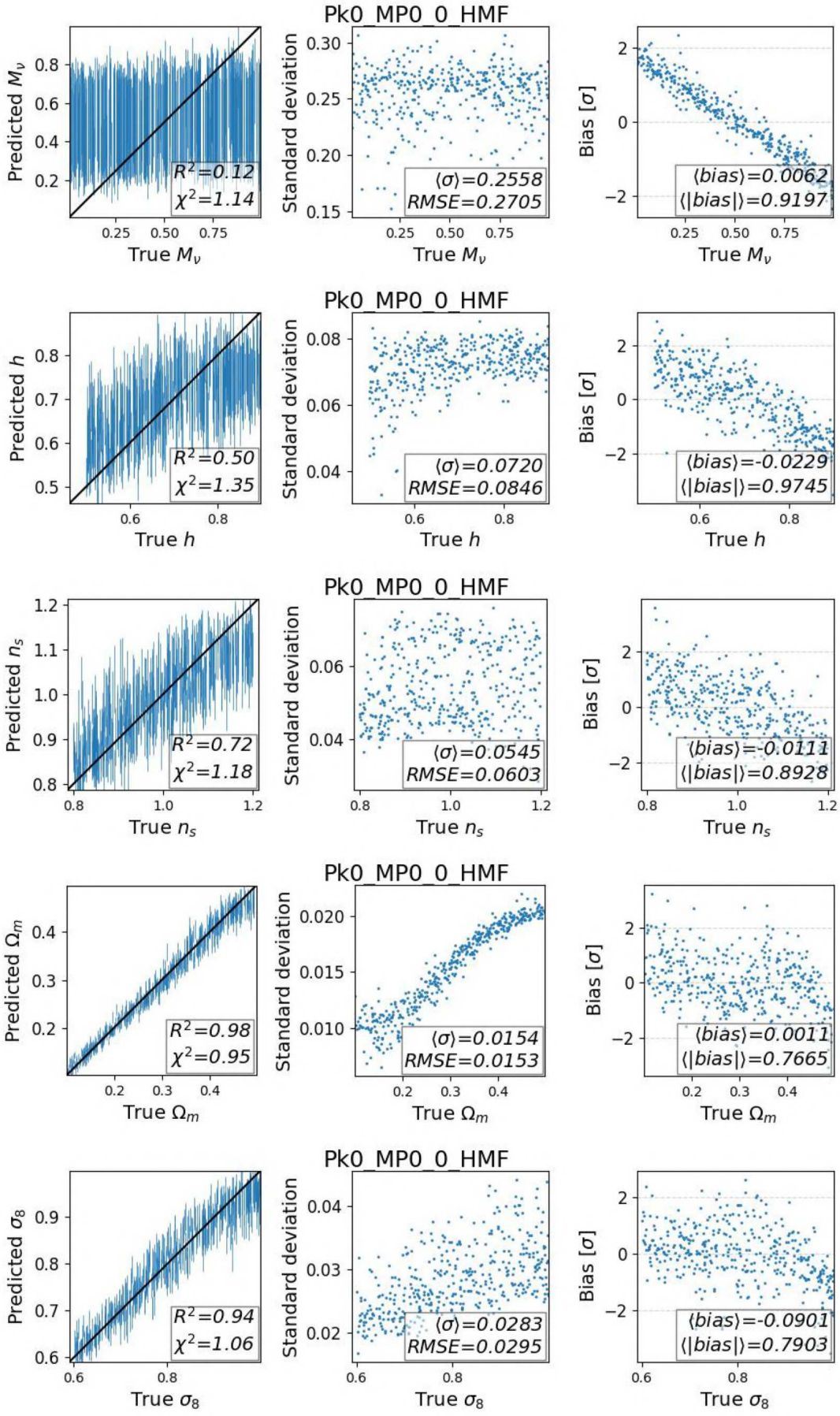

Figure B.11: Predictions of parameters for the best-performing network trained on the combination of Pk0˙MP0˙0˙HMF statistics, applied to the test set of the nwLH.

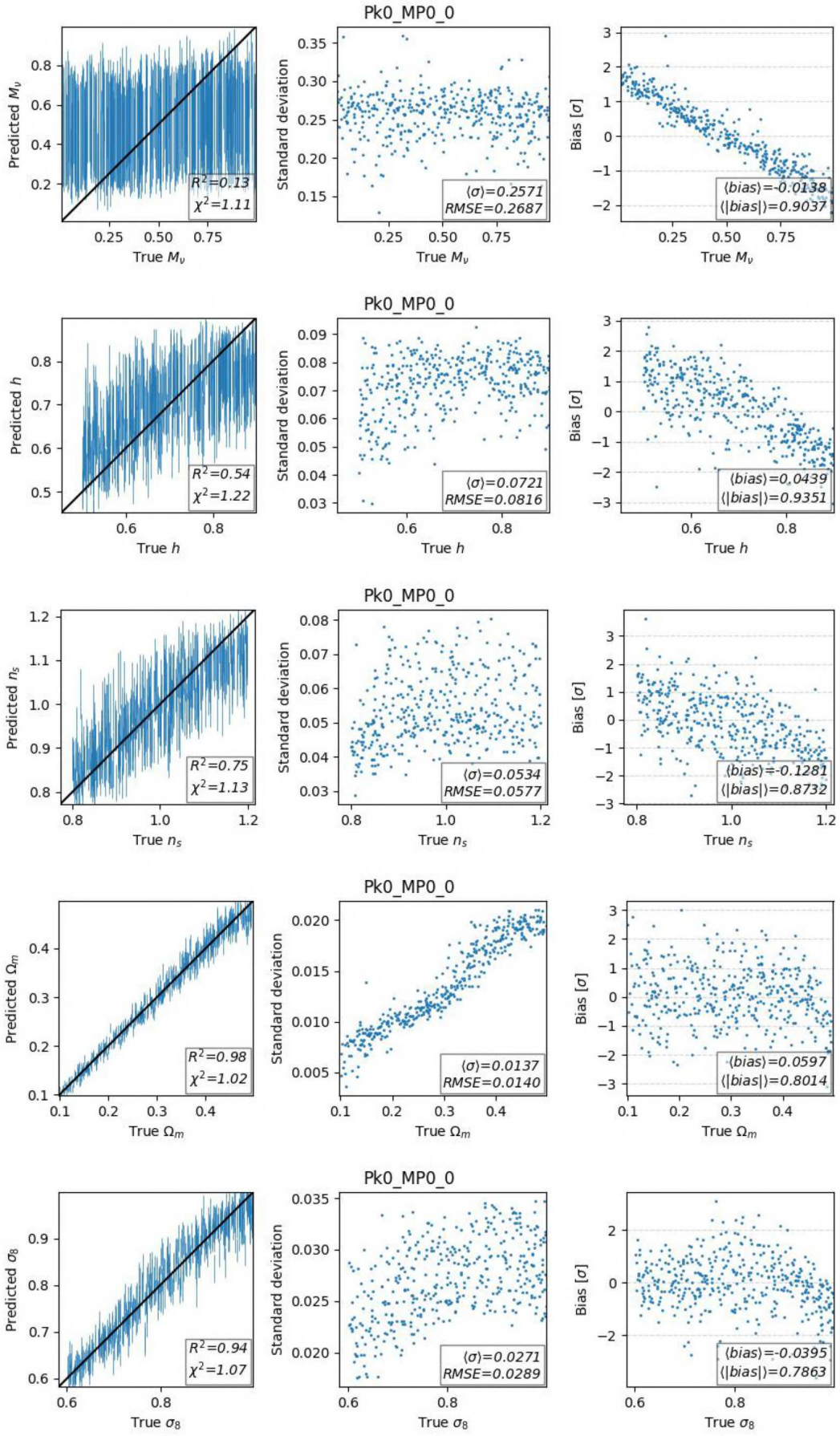

Figure B.12: Predictions of parameters for the best-performing network trained on the combination of Pk0˙MP0˙0 statistics, applied to the test set of the nwLH.

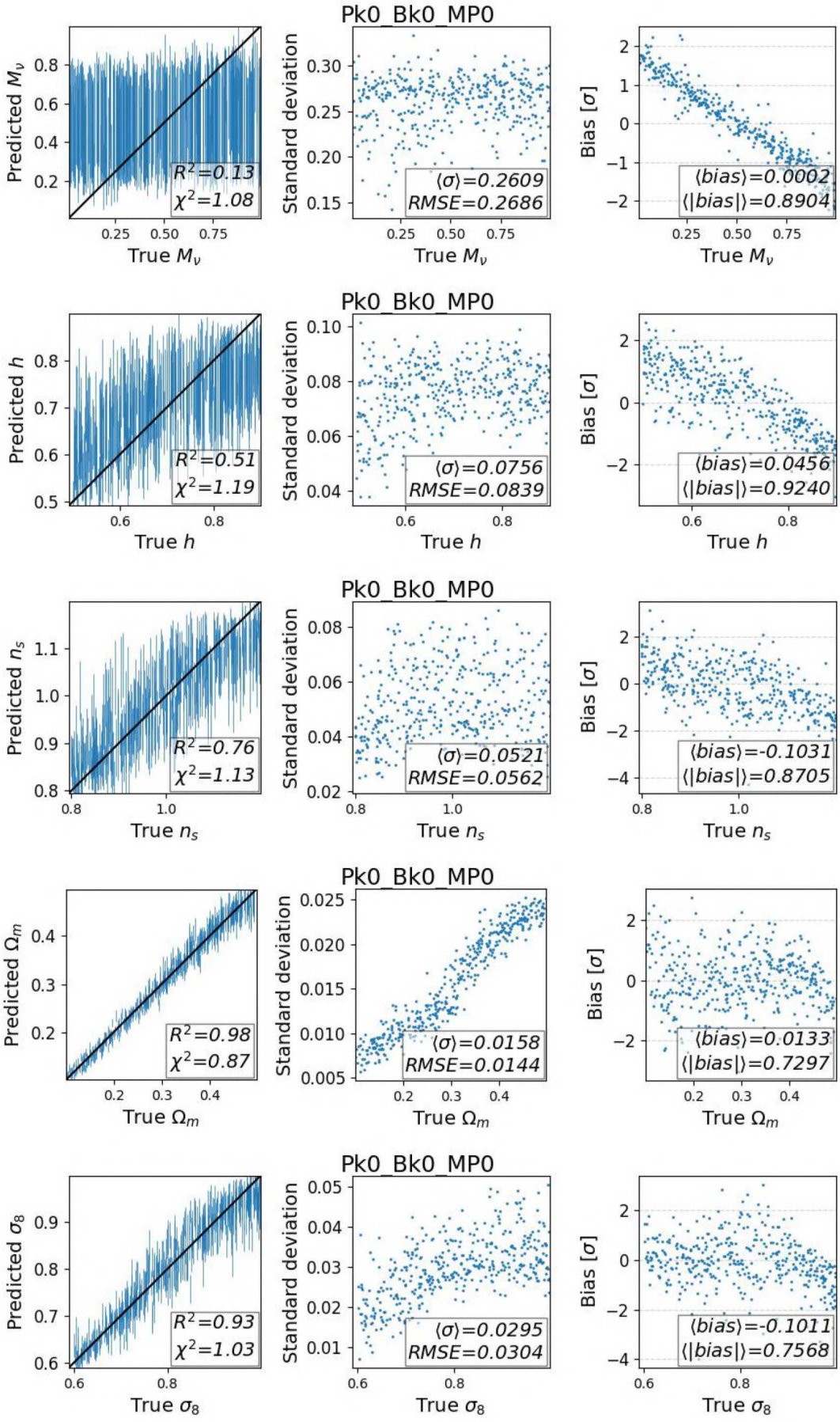

Figure B.13: Predictions of parameters for the best-performing network trained on the combination of Pk0˙Bk0˙MP0 statistics, applied to the test set of the nwLH.

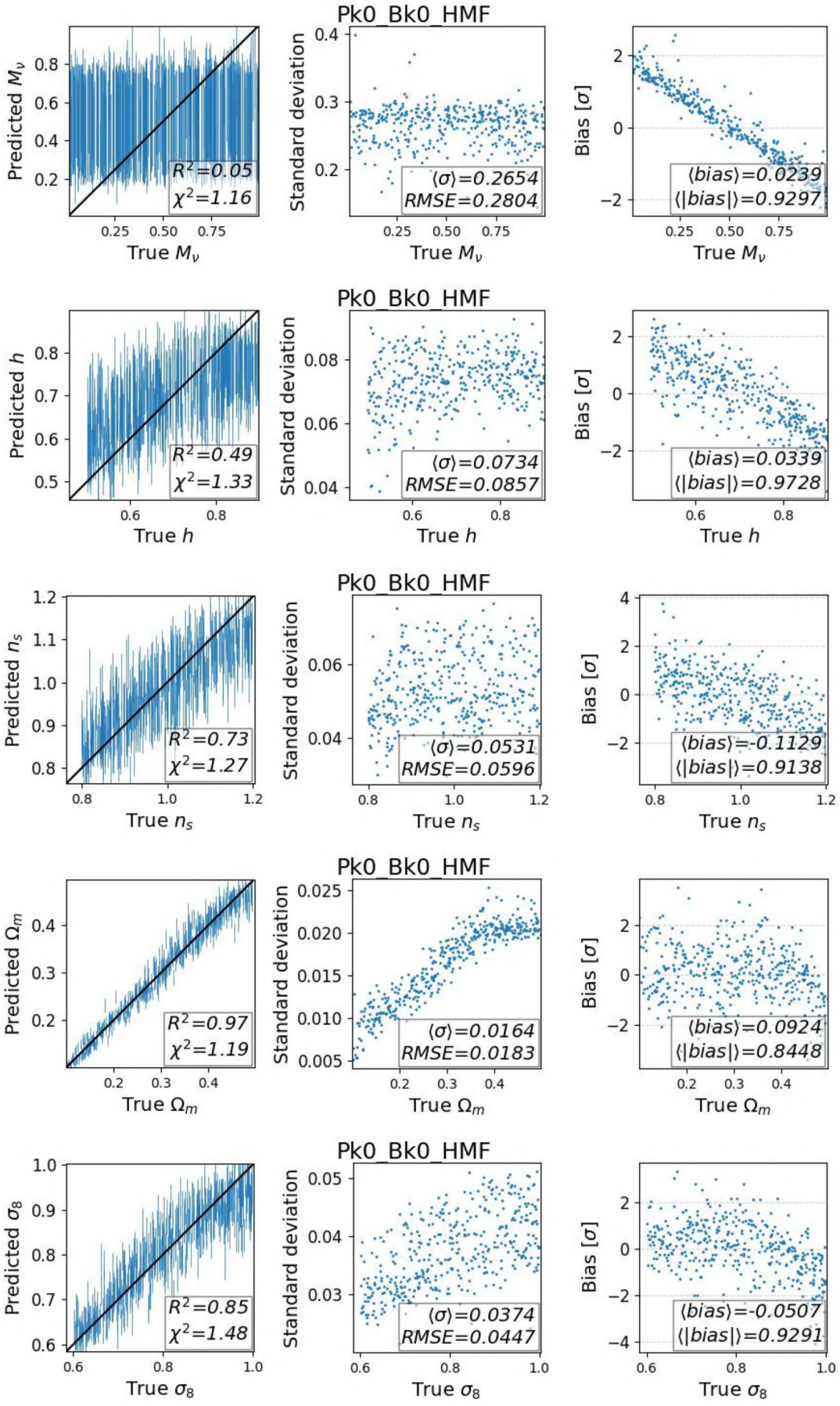

Figure B.14: Predictions of parameters for the best-performing network trained on the combination of Pk0˙Bk0˙HMF statistics, applied to the test set of the nwLH.

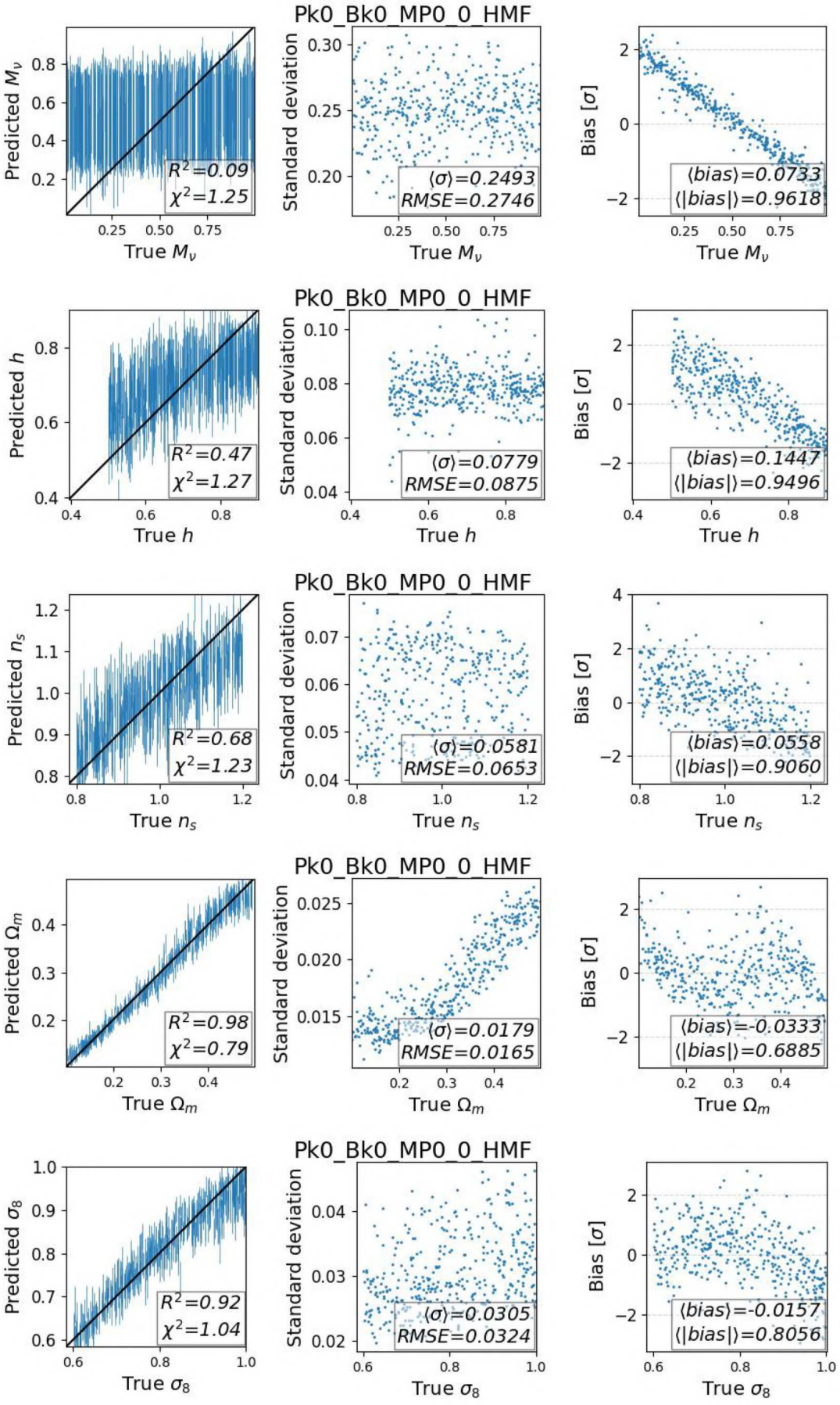

Figure B.15: Predictions of parameters for the best-performing network trained on the combination of Pk0˙Bk0˙MP0˙0˙HMF statistics, applied to the test set of the nwLH.

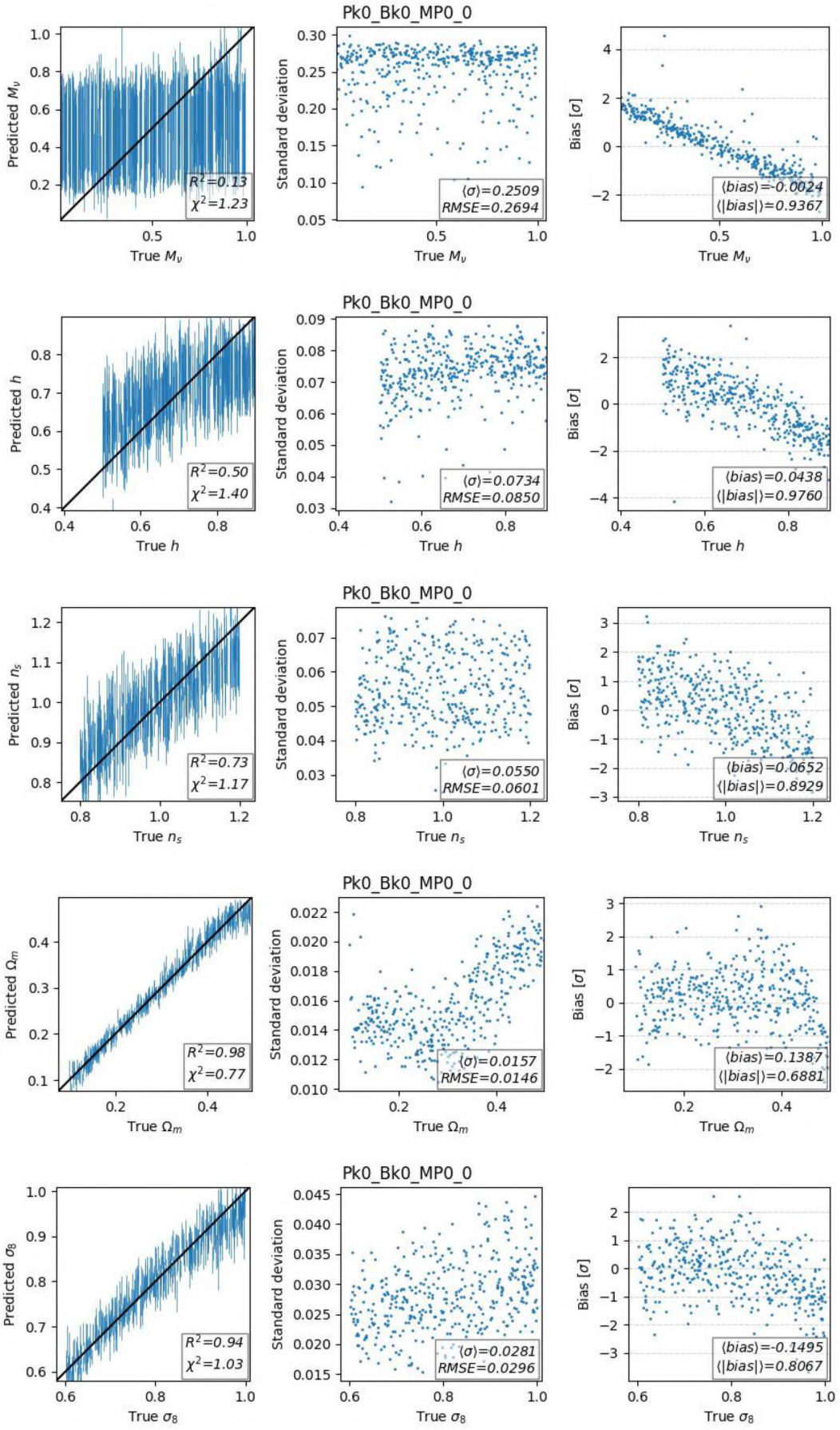

Figure B.16: Predictions of parameters for the best-performing network trained on the combination of Pk0˙Bk0˙MP0˙0 statistics, applied to the test set of the nwLH.

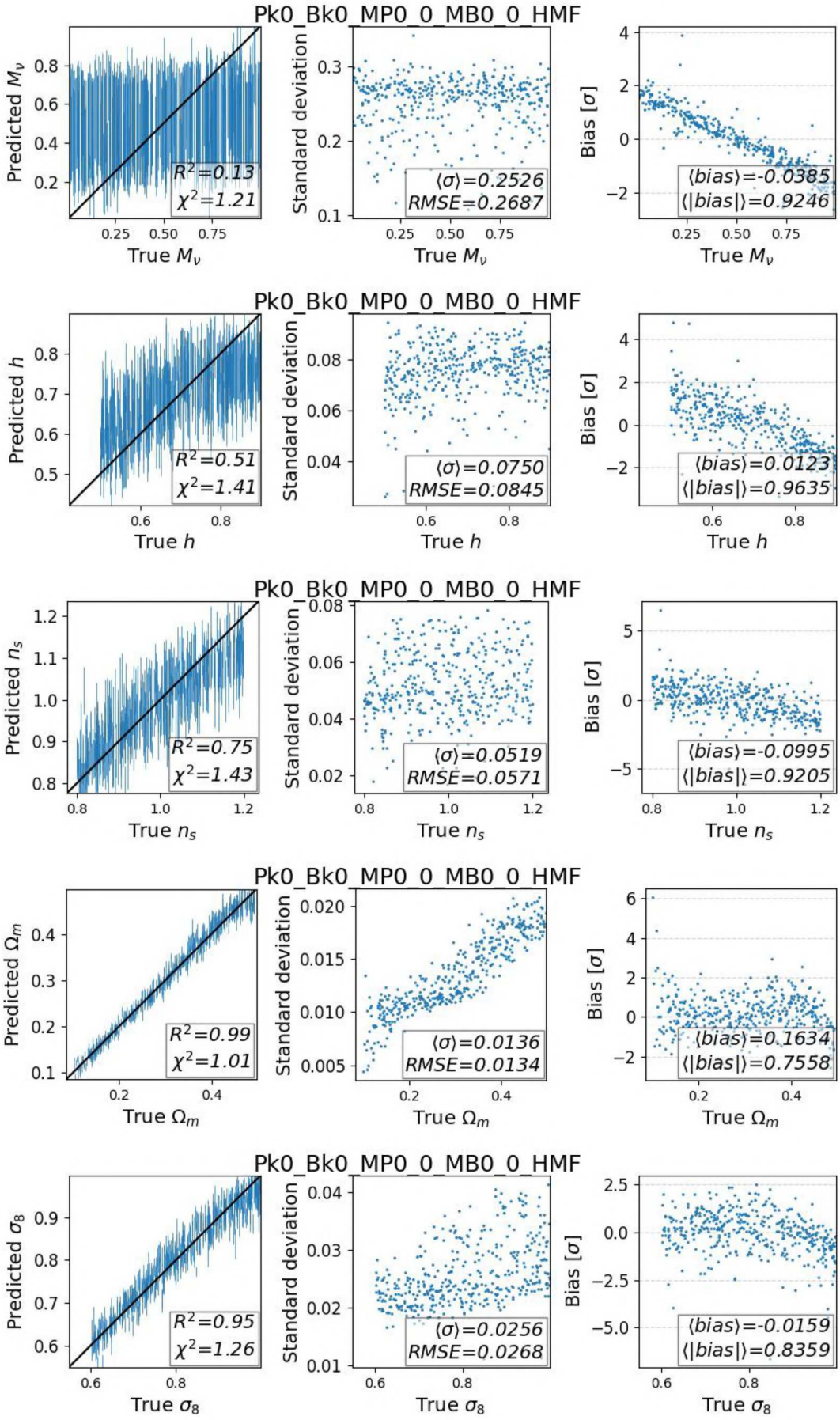

Figure B.17: Predictions of parameters for the best-performing network trained on the combination of Pk0˙Bk0˙MP0˙0˙MB0˙0˙HMF statistics, applied to the test set of the nwLH.

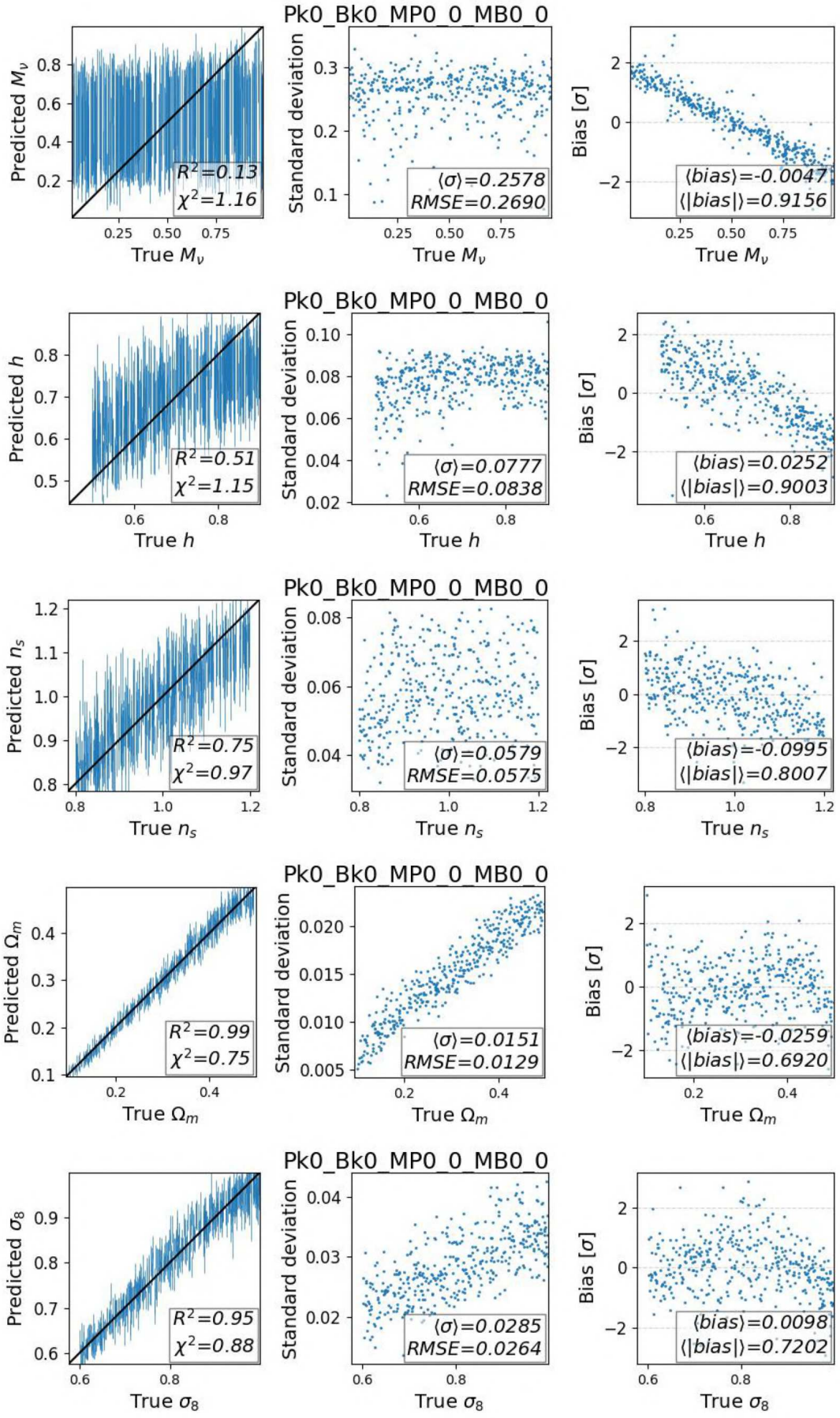

Figure B.18: Predictions of parameters for the best-performing network trained on the combination of Pk0˙Bk0˙MP0˙0˙MB0˙0 statistics, applied to the test set of the nwLH.

Bibliography

[EUC,] Euclid. https://www.euclid-ec.org/.

[Rub,] LSST. https://rubinobservatory.org/.

[Abazajian et al., 2016] Abazajian, K. N., Adshead, P., Ahmed, Z., Allen, S. W., Alonso, D., Arnold, K. S., Baccigalupi, C., Bartlett, J. G., Battaglia, N., Benson, B. A., Bischoff, C. A., Borrill, J., Buza, V., Calabrese, E., Caldwell, R., Carlstrom, J. E., Chang, C. L., Crawford, T. M., Cyr-Racine, F.-Y., Bernardis, F. D., de Haan, T., di Serego Alighieri, S., Dunkley, J., Dvorkin, C., Errard, J., Fabbian, G., Feeney, S., Ferraro, S., Filippini, J. P., Flauger, R., Fuller, G. M., Gluscevic, V., Green, D., Grin, D., Grohs, E., Henning, J. W., Hill, J. C., Hlozek, R., Holder, G., Holzapfel, W., Hu, W., Huffenberger, K. M., Keskitalo, R., Knox, L., Kosowsky, A., Kovac, J., Kovetz, E. D., Kuo, C.-L., Kusaka, A., Jeune, M. L., Lee, A. T., Lilley, M., Loverde, M., Madhavacheril, M. S., Mantz, A., Marsh, D. J. E., McMahon, J., Meerburg, P. D., Meyers, J., Miller, A. D., Munoz, J. B., Nguyen, H. N., Niemack, M. D., Peloso, M., Peloton, J., Pogosian, L., Pryke, C., Raveri, M., Reichardt, C. L., Rocha, G., Rotti, A., Schaan, E., Schmittfull, M. M., Scott, D., Sehgal, N., Shandera, S., Sherwin, B. D., Smith, T. L., Sorbo, L., Starkman, G. D., Story, K. T., van Engelen, A., Vieira, J. D., Watson, S., Whitehorn, N., and Wu, W. L. K. (2016). Cmb-s4 science book, first edition.

[Aghanim et al., 2020] Aghanim, N., Akrami, Y., Ashdown, M., Aumont, J., Baccigalupi, C., Ballardini, M., Banday, A. J., Barreiro, R. B., Bartolo, N., Basak, S., Battye, R., Benabed, K., Bernard, J.-P., Bersanelli, M., Bielewicz, P., Bock, J. J., Bond, J. R., Borrill, J., Bouchet, F. R., Boulanger, F., Bucher, M., Burigana, C., Butler, R. C., Calabrese, E., Cardoso, J.-F., Carron, J., Challinor, A., Chiang, H. C., Chluba, J., Colombo, L. P. L., Combet, C., Contreras, D., Crill, B. P., Cuttaia, F., de Bernardis, P., de Zotti, G., Delabrouille, J., Delouis, J.-M., Di Valentino, E., Diego, J. M., Doré, O., Douspis, M., Ducout, A., Dupac, X., Dusini, S., Efstathiou, G., Elsner, F., Enßlin, T. A., Eriksen, H. K., Fantaye, Y., Farhang, M., Fergusson, J., Fernandez-Cobos, R., Finelli, F., Forastieri, F., Frailis, M., Fraisse, A. A., Franceschi, E., Frolov, A., Galeotta, S., Galli, S., Ganga, K., Génova-Santos, R. T., Gerbino, M., Ghosh, T., González-Nuevo, J., Górski, K. M., Gratton, S., Gruppuso, A., Gudmundsson, J. E., Hamann, J., Handley, W., Hansen, F. K., Herranz, D., Hildebrandt, S. R., Hivon, E., Huang, Z., Jaffe, A. H., Jones, W. C., Karakci, A., Keihänen, E., Keskitalo, R., Kiiveri, K., Kim, J., Kisner, T. S., Knox, L., Krachmalnicoff, N., Kunz, M., Kurki-Suonio, H., Lagache, G., Lamarre, J.-M., Lasenby, A., Lattanzi, M., Lawrence, C. R., Le Jeune, M., Lemos, P., Lesgourgues, J., Levrier, F., Lewis, A., Liguori, M., Lilje, P. B., Lilley, M., Lindholm, V., López-Caniego, M., Lubin, P. M., Ma, Y.-Z., Macías-Pérez, J. F., Maggio, G., Maino, D., Mandolesi, N., Mangilli, A., Marcos-Caballero, A., Maris, M., Martin, P. G., Martinelli, M., Martínez-González, E., Matarrese, S., Mauri, N., McEwen, J. D., Meinhold, P. R., Melchiorri, A., Mennella, A., Migliaccio, M., Millea, M., Mitra, S., Miville-Deschênes, M.-A., Molinari, D., Montier, L., Morgante, G., Moss, A., Natoli, P., Nørgaard-Nielsen, H. U., Pagano, L., Paoletti, D., Partridge, B., Patanchon, G., Peiris, H. V., Perrotta, F., Pettorino, V., Piacentini, F., Polastri, L., Polenta, G., Puget, J.-L., Rachen, J. P., Reinecke, M., Remazeilles, M., Renzi, A., Rocha, G., Rosset, C., Roudier, G., Rubiño-Martín, J. A., Ruiz-Granados, B., Salvati, L., Sandri, M., Savelainen, M., Scott, D., Shellard, E. P. S., Sirignano, C., Sirri, G., Spencer, L. D., Sunyaev, R., Suur-Uski, A.-S., Tauber, J. A., Tavagnacco, D., Tenti, M., Toffolatti, L., Tomasi, M., Trombetti, T., Valenziano, L., Valiviita, J., Van Tent, B., Vibert, L., Vielva, P., Villa, F., Vittorio, N., Wandelt, B. D., Wehus,

I. K., White, M., White, S. D. M., Zacchei, A., and Zonca, A. (2020). Planck2018 results: Vi. cosmological parameters. *Astronomy amp; Astrophysics*, 641:A6.

[Akeret et al., 2013] Akeret, J., Seehars, S., Amara, A., Refregier, A., and Csillaghy, A. (2013). Cosmohammer: Cosmological parameter estimation with the mcmc hammer. *Astronomy and Computing*, 2:27–39.

[Akrami et al., 2019] Akrami, P., Arroja, F., Ashdown, M., Aumont, J., Baccigalupi, C., Ballardini, M., Banday, A. J., Barreiro, R. B., Bartolo, N., Basak, S., Benabed, K., Bernard, J. P., Bersanelli, M., Bielewicz, P., Bond, J. R., Borrill, J., Bouchet, F. R., Bucher, M., Burigana, C., Butler, R. C., Calabrese, E., Cardoso, J. F., Casaponsa, B., Challinor, A., Chiang, H. C., Colombo, L. P. L., Combet, C., Crill, B. P., Cuttaia, F., de Bernardis, P., de Rosa, A., de Zotti, G., Delabrouille, J., Delouis, J. M., Valentino, E. D., Diego, J. M., Doré, O., Douspis, M., Ducout, A., Dupac, X., Dusini, S., Efstathiou, G., Elsner, F., Enßlin, T. A., Eriksen, H. K., Fantaye, Y., Fergusson, J., Fernandez-Cobos, R., Finelli, F., Frailis, M., Fraisse, A. A., Franceschi, E., Frolov, A., Galeotta, S., Ganga, K., Génova-Santos, R. T., Gerbino, M., González-Nuevo, J., Górski, K. M., Gratton, S., Gruppuso, A., Gudmundsson, J. E., Hamann, J., Handley, W., Hansen, F. K., Herranz, D., Hivon, E., Huang, Z., Jaffe, A. H., Jones, W. C., Jung, G., Keihänen, E., Keskitalo, R., Kiiveri, K., Kim, J., Krachmalnicoff, N., Kunz, M., Kurki-Suonio, H., Lamarre, J. M., Lasenby, A., Lattanzi, M., Lawrence, C. R., Jeune, M. L., Levrier, F., Lewis, A., Liguori, M., Lilje, P. B., Lindholm, V., López-Caniego, M., Ma, Y. Z., Macías-Pérez, J. F., Maggio, G., Maino, D., Mandolesi, N., Marcos-Caballero, A., Maris, M., Martin, P. G., Martínez-González, E., Matarrese, S., Mauri, N., McEwen, J. D., Meerburg, P. D., Meinhold, P. R., Melchiorri, A., Mennella, A., Migliaccio, M., Miville-Deschênes, M. A., Molinari, D., Moneti, A., Montier, L., Morgante, G., Moss, A., Münchmeyer, M., Natoli, P., Oppizzi, F., Pagano, L., Paoletti, D., Partridge, B., Patanchon, G., Perrotta, F., Pettorino, V., Piacentini, F., Polenta, G., Puget, J. L., Rachen, J. P., Racine, B., Reinecke, M., Remazeilles, M., Renzi, A., Rocha, G., Rubiño-Martín, J. A., Ruiz-Granados, B., Salvati, L., Savelainen, M., Scott, D., Shellard, E. P. S., Shiraishi, M., Sirignano, C., Sirri, G., Smith, K., Spencer, L. D., Stanco, L., Sunyaev, R., Suur-Uski, A. S., Tauber, J. A., Tavagnacco, D., Tenti, M., Toffolatti, L., Tomasi, M., Trombetti, T., Valiviita, J., Tent, B. V., Vielva, P., Villa, F., Vittorio, N., Wandelt, B. D., Wehus, I. K., Zacchei, A., and Zonca, A. (2019). Planck 2018 results. ix. constraints on primordial non-gaussianity.

[Alsing et al., 2019] Alsing, J., Charnock, T., Feeney, S., and Wandelt, B. (2019). Fast likelihood-free cosmology with neural density estimators and active learning. *Monthly Notices of the Royal Astronomical Society*.

[Alsing et al., 2018] Alsing, J., Wandelt, B., and Feeney, S. (2018). Massive optimal data compression and density estimation for scalable, likelihood-free inference in cosmology. *Monthly Notices of the Royal Astronomical Society*, 477(3):2874–2885.

[Balaguera-Antolínez, 2014] Balaguera-Antolínez, A. (2014). What can the spatial distribution of galaxy clusters tell about their scaling relations? *Astronomy amp; Astrophysics*, 563:A141.

[Barnes and Hut, 1986] Barnes, J. and Hut, P. (1986). A hierarchical O(N log N) force-calculation algorithm. , 324(6096):446–449.

[Bartolo et al., 2004] Bartolo, N., Komatsu, E., Matarrese, S., and Riotto, A. (2004). Non-gaussianity from inflation: theory and observations. *Physics Reports*, 402(3-4):103–266.

[Baumann, 2012] Baumann, D. (2012). Tasi lectures on inflation.

[Baumann, 2022] Baumann, D. (2022). *Cosmology*. Cambridge University Press.

[Bayer et al., 2021] Bayer, A. E., Villaescusa-Navarro, F., Massara, E., Liu, J., Spergel, D. N., Verde, L., Wandelt, B. D., Viel, M., and Ho, S. (2021). Detecting neutrino mass by combining matter clustering, halos, and voids. *The Astrophysical Journal*, 919(1):24.

[Beisbart and Kerscher, 2000] Beisbart, C. and Kerscher, M. (2000). Luminosity- and morphology-dependent clustering of galaxies. *The Astrophysical Journal*, 545(1):6–25.

[Bernardeau et al., 2002] Bernardeau, F., Colombi, S., Gaztañaga, E., and Scoccimarro, R. (2002). Large-scale structure of the universe and cosmological perturbation theory. *Physics Reports*, 367(1–3):1–248.

[Birdsall and Fuss, 1997] Birdsall, C. K. and Fuss, D. (1997). Clouds-in-clouds, clouds-in-cells physics for many-body plasma simulation. *Journal of Computational Physics*, 135(2):141–148.

[Bond et al., 1991] Bond, J. R., Cole, S., Efstathiou, G., and Kaiser, N. (1991). Excursion Set Mass Functions for Hierarchical Gaussian Fluctuations. , 379:440.

[Bouchet, 1996] Bouchet, F. R. (1996). Introductory overview of eulerian and lagrangian perturbation theories.

[Brandbyge et al., 2010] Brandbyge, J., Hannestad, S., Haugbølle, T., and Wong, Y. Y. (2010). Neutrinos in non-linear structure formation — the effect on halo properties. *Journal of Cosmology and Astroparticle Physics*, 2010(09):014–014.

[Buchert, 1989] Buchert, T. (1989). A class of solutions in Newtonian cosmology and the pancake theory. , 223(1-2):9–24.

[Buchert et al., 1994] Buchert, T., Melott, A. L., and Weiss, A. G. (1994). Testing higher-order Lagrangian perturbation theory against numerical simulations I. Pancake models. , 288:349–364.

[Byrnes and Choi, 2010] Byrnes, C. T. and Choi, K.-Y. (2010). Review of local non-gaussianity from multifield inflation. *Advances in Astronomy*, 2010(1).

[Charnock et al., 2018] Charnock, T., Lavaux, G., and Wandelt, B. D. (2018). Automatic physical inference with information maximizing neural networks. *Physical Review D*, 97(8).

[Chen et al., 2007] Chen, X., Huang, M.-x., Kachru, S., and Shiu, G. (2007). Observational signatures and non-gaussianities of general single-field inflation. *Journal of Cosmology and Astroparticle Physics*, 2007(01):002–002.

[Coulton et al., 2023] Coulton, W. R., Villaescusa-Navarro, F., Jamieson, D., Baldi, M., Jung, G., Karagiannis, D., Liguori, M., Verde, L., and Wandelt, B. D. (2023). Quijote-png: Simulations of primordial non-gaussianity and the information content of the matter field power spectrum and bispectrum. *The Astrophysical Journal*, 943(1):64.

[Creminelli et al., 2006] Creminelli, P., Nicolis, A., Senatore, L., Tegmark, M., and Zaldarriaga, M. (2006). Limits on non-Gaussianities from WMAP data. , 2006(5):004.

[Davis et al., 1985] Davis, M., Efstathiou, G., Frenk, C. S., and White, S. D. M. (1985). The evolution of large-scale structure in a universe dominated by cold dark matter. , 292:371–394.

[Dehnen, 2001] Dehnen, W. (2001). Towards optimal softening in three-dimensionaln-body codes - i. minimizing the force error. *Monthly Notices of the Royal Astronomical Society*, 324(2):273–291.

[Desjacques et al., 2009] Desjacques, V., Seljak, U., and Iliev, I. T. (2009). Scale-dependent bias induced by local non-gaussianity: a comparison ton-body simulations. *Monthly Notices of the Royal Astronomical Society*, 396(1):85–96.

[Dodelson and Schmidt, 2020] Dodelson, S. and Schmidt, F. (2020). *Modern Cosmology*.

[Feldman et al., 1994] Feldman, H. A., Kaiser, N., and Peacock, J. A. (1994). Power-Spectrum Analysis of Three-dimensional Redshift Surveys. , 426:23.

[Fergusson et al., 2012] Fergusson, J. R., Regan, D. M., and Shellard, E. P. S. (2012). Rapid separable analysis of higher order correlators in large-scale structure. *Physical Review D*, 86(6).

[Fry, 1984] Fry, J. N. (1984). The Galaxy correlation hierarchy in perturbation theory. , 279:499–510.

[Fry, 1994] Fry, J. N. (1994). The Minimal Power Spectrum: Higher Order Contributions. , 421:21.

[Gnedenko, 1998] Gnedenko, B. (1998). *Theory of Probability*. Routledge, 6th edition.

[Goroff et al., 1986] Goroff, M. H., Grinstein, B., Rey, S. J., and Wise, M. B. (1986). Coupling of modes of cosmological mass density fluctuations. *The Astrophysical Journal*, 311:6–14.

[He et al., 2015] He, K., Zhang, X., Ren, S., and Sun, J. (2015). Deep residual learning for image recognition.

[Hockney and Eastwood, 1981] Hockney, R. W. and Eastwood, J. W. (1981). *Computer Simulation Using Particles*.

[Jain and Bertschinger, 1994] Jain, B. and Bertschinger, E. (1994). Second-order power spectrum and nonlinear evolution at high redshift. *The Astrophysical Journal*, 431:495.

[Jeffrey and Wandelt, 2020] Jeffrey, N. and Wandelt, B. D. (2020). Solving high-dimensional parameter inference: marginal posterior densities moment networks.

[Jeong and Komatsu, 2006] Jeong, D. and Komatsu, E. (2006). Perturbation theory reloaded: Analytical calculation of nonlinearity in baryonic oscillations in the real-space matter power spectrum. *The Astrophysical Journal*, 651(2):619–626.

[Jung et al., 2022] Jung, G., Karagiannis, D., Liguori, M., Baldi, M., Coulton, W. R., Jamieson, D., Verde, L., Villaescusa-Navarro, F., and Wandelt, B. D. (2022). Quijote-png: Quasi-maximum likelihood estimation of primordial non-gaussianity in the nonlinear dark matter density field. *The Astrophysical Journal*, 940(1):71.

[Jung et al., 2023] Jung, G., Karagiannis, D., Liguori, M., Baldi, M., Coulton, W. R., Jamieson, D., Verde, L., Villaescusa-Navarro, F., and Wandelt, B. D. (2023). Quijote-png: Quasi-maximum likelihood estimation of primordial non-gaussianity in the nonlinear halo density field. *The Astrophysical Journal*, 948(2):135.

[Jung et al., 2024a] Jung, G., Ravenni, A., Liguori, M., Baldi, M., Coulton, W. R., Villaescusa-Navarro, F., and Wandelt, B. D. (2024a). Quijote-png: Optimizing the summary statistics to measure primordial non-gaussianity.

[Jung et al., 2024b] Jung, G., Ravenni, A., Liguori, M., Baldi, M., Coulton, W. R., Villaescusa-Navarro, F., and Wandelt, B. D. (2024b). Quijote-PNG: Optimizing the summary statistics to measure Primordial non-Gaussianity.

[Kingma and Ba, 2017] Kingma, D. P. and Ba, J. (2017). Adam: A method for stochastic optimization.

[Klambauer et al., 2017] Klambauer, G., Unterthiner, T., Mayr, A., and Hochreiter, S. (2017). Self-normalizing neural networks.

[Klypin et al., 2011] Klypin, A. A., Trujillo-Gomez, S., and Primack, J. (2011). Dark matter halos in the standard cosmological model: Results from the bolshoi simulation. *The Astrophysical Journal*, 740(2):102.

[Komatsu et al., 2005] Komatsu, E., Spergel, D. N., and Wandelt, B. D. (2005). Measuring primordial non-gaussianity in the cosmic microwave background. *The Astrophysical Journal*, 634(1):14–19.

[Krywonos et al., 2024] Krywonos, J., Paradiso, S., Krolewski, A., Joudaki, S., and Percival, W. (2024). Cosmological measurements from the cmb and bao are insensitive to the tail probability in the assumed likelihood.

[LESGOURGUES and PASTOR, 2006] LESGOURGUES, J. and PASTOR, S. (2006). Massive neutrinos and cosmology. *Physics Reports*, 429(6):307–379.

[Liguori et al., 2010] Liguori, M., Sefusatti, E., Fergusson, J. R., and Shellard, E. P. S. (2010). Primordial non-Gaussianity and Bispectrum Measurements in the Cosmic Microwave Background and Large-Scale Structure. *Adv. Astron.*, 2010:980523.

[Lueckmann et al., 2019] Lueckmann, J.-M., Bassetto, G., Karaletsos, T., and Macke, J. H. (2019). Likelihood-free inference with emulator networks.

[Lyth and Wands, 2002] Lyth, D. and Wands, D. (2002). Generating the curvature perturbation without an inflaton. *Physics Letters B*, 524(1–2):5–14.

[Maksimova et al., 2021] Maksimova, N. A., Garrison, L. H., Eisenstein, D. J., Hadzhiyska, B., Bose, S., and Satterthwaite, T. P. (2021). ¡scp¿abacussummit¡/scp¿: a massive set of high-accuracy, high-resolution n-body simulations. *Monthly Notices of the Royal Astronomical Society*, 508(3):4017–4037.

[Maldacena, 2003] Maldacena, J. (2003). Non-gaussian features of primordial fluctuations in single field inflationary models. *Journal of High Energy Physics*, 2003(05):013–013.

[Mangano et al., 2005] Mangano, G., Miele, G., Pastor, S., Pinto, T., Pisanti, O., and Serpico, P. D. (2005). Relic neutrino decoupling including flavour oscillations. *Nuclear Physics B*, 729(1–2):221–234.

[Marulli et al., 2011] Marulli, F., Carbone, C., Viel, M., Moscardini, L., and Cimatti, A. (2011). Effects of massive neutrinos on the large-scale structure of the universe: Neutrino effects on lss. *Monthly Notices of the Royal Astronomical Society*, 418(1):346–356.

[Massara et al., 2023] Massara, E., Villaescusa-Navarro, F., Hahn, C., Abidi, M. M., Eickenberg, M., Ho, S., Lemos, P., Dizgah, A. M., and Blancard, B. R.-S. (2023). Cosmological information in the marked power spectrum of the galaxy field. *The Astrophysical Journal*, 951(1):70.

[Massara et al., 2021] Massara, E., Villaescusa-Navarro, F., Ho, S., Dalal, N., and Spergel, D. N. (2021). Using the marked power spectrum to detect the signature of neutrinos in large-scale structure. *Physical Review Letters*, 126(1).

[Matarrese et al., 2000] Matarrese, S., Verde, L., and Jimenez, R. (2000). The abundance of high-redshift objects as a probe of non-gaussian initial conditions. *The Astrophysical Journal*, 541(1):10–24.

[Meerburg et al., 2019] Meerburg, P. D., Green, D., Abidi, M., Amin, M. A., Adshead, P., Ahmed, Z., Alonso, D., Ansarinejad, B., Armstrong, R., Avila, S., Baccigalupi, C., Baldauf, T., Ballardini, M., Bandura, K., Bartolo, N., Battaglia, N., Baumann, D., Bavdhankar, C., Bernal, J. L., Beutler, F., Biagetti, M., Bischoff, C., Blazek, J., Bond, J. R., Borrill, J., Bouchet, F. R., Bull, P., Burgess, C., Byrnes, C., Calabrese, E., Carlstrom, J. E., Castorina, E., Challinor, A., Chang, T.-C., Chaves-Montero, J., Chen, X., Yeche, C., Cooray, A., Coulton, W., Crawford, T., Chisari, E., Cyr-Racine, F.-Y., D'Amico, G., de Bernardis, P., de la Macorra, A., Doré, O., Duivenvoorden, A., Dunkley, J., Dvorkin, C., Eggemeier, A., Escoffier, S., Essinger-Hileman, T., Fasiello, M., Ferraro, S., Flauger, R., Font-Ribera, A., Foreman, S., Friedrich, O., Garcia-Bellido, J., Gerbino, M., Gluscevic, V., Goon, G., Gorski, K. M., Gudmundsson, J. E., Gupta, N., Hanany, S., Handley, W., Hawken, A. J., Hill, J. C., Hirata, C. M., Hložek, R., Holder, G., Huterer, D., Kamionkowski, M., Karkare, K. S., Keeley, R. E., Kinney, W., Kisner, T., Kneib, J.-P., Knox, L., Koushiappas, S. M., Kovetz, E. D., Koyama, K., L'Huillier, B., Lahav, O., Lattanzi, M., Lee, H., Liguori, M., Loverde,

M., Madhavacheril, M., Maldacena, J., Marsh, M. C. D., Masui, K., Matarrese, S., McAllister, L., McMahon, J., McQuinn, M., Meyers, J., Mirbabayi, M., Dizgah, A. M., Motloch, P., Mukherjee, S., Muñoz, J. B., Myers, A. D., Nagy, J., Naselsky, P., Nati, F., Newburgh, Nicolis, A., Niemack, M. D., Niz, G., Nomerotski, A., Page, L., Pajer, E., Padmanabhan, H., Palma, G. A., Peiris, H. V., Percival, W. J., Piacentni, F., Pimentel, G. L., Pogosian, L., Prescod-Weinstein, C., Pryke, C., Puglisi, G., Racine, B., Stompor, R., Raveri, M., Remazeilles, M., Rocha, G., Ross, A. J., Rossi, G., Ruhl, J., Sasaki, M., Schaan, E., Schillaci, A., Schmittfull, M., Sehgal, N., Senatore, L., Seo, H.-J., Shan, H., Shandera, S., Sherwin, B. D., Silverstein, E., Simon, S., Slosar, A., Staggs, S., Starkman, G., Stebbins, A., Suzuki, A., Switzer, E. R., Timbie, P., Tolley, A. J., Tomasi, M., Tristram, M., Trodden, M., Tsai, Y.-D., Uhlemann, C., Umilta, C., van Engelen, A., Vargas-Magaña, M., Vieregg, A., Wallisch, B., Wands, D., Wandelt, B., Wang, Y., Watson, S., Wise, M., Wu, W. L. K., Xianyu, Z.-Z., Xu, W., Yasini, S., Young, S., Yutong, D., Zaldarriaga, M., Zemcov, M., Zhao, G.-B., Zheng, Y., and Zhu, N. (2019). Primordial non-gaussianity.

[Melott et al., 1995] Melott, A. L., Buchert, T., and Weiss, A. G. (1995). Testing higher-order Lagrangian perturbation theory against numerical simulations. II. Hierarchical models. , 294:345–365.

[Michaux et al., 2020] Michaux, M., Hahn, O., Rampf, C., and Angulo, R. E. (2020). Accurate initial conditions for cosmological n-body simulations: minimizing truncation and discreteness errors. *Monthly Notices of the Royal Astronomical Society*, 500(1):663–683.

[Norman and Bryan,] Norman, M. L. and Bryan, G. L. *Cluster turbulence*, page 106–115. Springer Berlin Heidelberg.

[Olive et al., 2014] Olive, K. A. et al. (2014). Review of Particle Physics. *Chin. Phys. C*, 38:090001.

[O'Malley et al., 2019] O'Malley, T., Bursztein, E., Long, J., Chollet, F., Jin, H., Invernizzi, L., et al. (2019). Kerastuner. `https://github.com/keras-team/keras-tuner`.

[Papamakarios, 2019] Papamakarios, G. (2019). *Neural Density Estimation and Likelihood-free Inference*. PhD thesis, Edinburgh U.

[Peebles, 1976] Peebles, P. J. E. (1976). The Peculiar Velocity Field in the Local Supercluster. *The Astrophysical Journal*, 205:318–328.

[Philcox et al., 2020] Philcox, O. H., Massara, E., and Spergel, D. N. (2020). What does the marked power spectrum measure? insights from perturbation theory. *Physical Review D*, 102(4).

[Polarski and Starobinsky, 1994] Polarski, D. and Starobinsky, A. A. (1994). Isocurvature perturbations in multiple inflationary models. *Physical Review D*, 50(10):6123–6129.

[Press and Schechter, 1974] Press, W. H. and Schechter, P. (1974). Formation of Galaxies and Clusters of Galaxies by Self-Similar Gravitational Condensation. , 187:425–438.

[Press et al., 2007] Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P. (2007). *Numerical Recipes: The Art of Scientific Computing*. Cambridge University Press, Cambridge, UK, 3rd edition.

[Rimes and Hamilton, 2005] Rimes, C. D. and Hamilton, A. J. S. (2005). Information content of the non-linear matter power spectrum. *Monthly Notices of the Royal Astronomical Society: Letters*, 360(1):L82–L86.

[Scoccimarro, 1997] Scoccimarro, R. (1997). Cosmological perturbations: Entering the nonlinear regime. *The Astrophysical Journal*, 487(1):1–17.

[Scoccimarro et al., 1999] Scoccimarro, R., Couchman, H. M. P., and Frieman, J. A. (1999). The bispectrum as a signature of gravitational instability in redshift space. *The Astrophysical Journal*, 517(2):531–540.

[Scoccimarro et al., 2004] Scoccimarro, R., Sefusatti, E., and Zaldarriaga, M. (2004). Probing primordial non-gaussianity with large-scale structure. *Physical Review D*, 69(10).

[Sefusatti, 2009] Sefusatti, E. (2009). One-loop perturbative corrections to the matter and galaxy bispectrum with non-gaussian initial conditions. *Physical Review D*, 80(12).

[Sefusatti and Komatsu, 2007] Sefusatti, E. and Komatsu, E. (2007). Bispectrum of galaxies from high-redshift galaxy surveys: Primordial non-gaussianity and nonlinear galaxy bias. *Physical Review D*, 76(8).

[Sheth and Tormen, 1999] Sheth, R. K. and Tormen, G. (1999). Large-scale bias and the peak background split. *Monthly Notices of the Royal Astronomical Society*, 308(1):119–126.

[Sirko, 2005] Sirko, E. (2005). Initial conditions to cosmological N-body simulations, or how to run an ensemble of simulations. *Astrophys. J.*, 634:728–743.

[Skibba et al., 2006] Skibba, R., Sheth, R. K., Connolly, A. J., and Scranton, R. (2006). The luminosity-weighted or 'marked' correlation function. , 369(1):68–76.

[Smith, 2017] Smith, L. N. (2017). Cyclical learning rates for training neural networks.

[Springel et al., 2018] Springel, V., Pakmor, R., Pillepich, A., Weinberger, R., Nelson, D., Hernquist, L., Vogelsberger, M., Genel, S., Torrey, P., Marinacci, F., and Naiman, J. (2018). First results from the IllustrisTNG simulations: matter and galaxy clustering. , 475(1):676–698.

[Springel et al., 2005] Springel, V., White, S. D. M., Jenkins, A., Frenk, C. S., Yoshida, N., Gao, L., Navarro, J., Thacker, R., Croton, D., Helly, J., Peacock, J. A., Cole, S., Thomas, P., Couchman, H., Evrard, A., Colberg, J., and Pearce, F. (2005). Simulations of the formation, evolution and clustering of galaxies and quasars. , 435(7042):629–636.

[Srivastava et al., 2014] Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., and Salakhutdinov, R. (2014). Dropout: A simple way to prevent neural networks from overfitting. *Journal of Machine Learning Research*, 15(56):1929–1958.

[Trenti and Hut, 2008] Trenti, M. and Hut, P. (2008). N-body simulations (gravitational). *Scholarpedia*, 3(5):3930. revision #91544.

[Verde, 2008] Verde, L. (2008). A practical guide to basic statistical techniques for data analysis in cosmology.

[Villaescusa-Navarro et al., 2022] Villaescusa-Navarro, F., Genel, S., Anglés-Alcázar, D., Thiele, L., Dave, R., Narayanan, D., Nicola, A., Li, Y., Villanueva-Domingo, P., Wandelt, B., Spergel, D. N., Somerville, R. S., Zorrilla Matilla, J. M., Mohammad, F. G., Hassan, S., Shao, H., Wadekar, D., Eickenberg, M., Wong, K. W. K., Contardo, G., Jo, Y., Moser, E., Lau, E. T., Machado Poletti Valle, L. F., Perez, L. A., Nagai, D., Battaglia, N., and Vogelsberger, M. (2022). The camels multifield data set: Learning the universe's fundamental parameters with artificial intelligence. *The Astrophysical Journal Supplement Series*, 259(2):61.

[Villaescusa-Navarro et al., 2020a] Villaescusa-Navarro, F., Hahn, C., Massara, E., Banerjee, A., Delgado, A. M., Ramanah, D. K., Charnock, T., Giusarma, E., Li, Y., Allys, E., Brochard, A., Uhlemann, C., Chiang, C.-T., He, S., Pisani, A., Obuljen, A., Feng, Y., Castorina, E., Contardo, G., Kreisch, C. D., Nicola, A., Alsing, J., Scoccimarro, R., Verde, L., Viel, M., Ho, S., Mallat, S., Wandelt, B., and Spergel, D. N. (2020a). The quijote simulations. *The Astrophysical Journal Supplement Series*, 250(1):2.

[Villaescusa-Navarro et al., 2020b] Villaescusa-Navarro, F., Hahn, C., Massara, E., Banerjee, A., Delgado, A. M., Ramanah, D. K., Charnock, T., Giusarma, E., Li, Y., Allys, E., Brochard, A., Uhlemann, C., Chiang, C.-T., He, S., Pisani, A., Obuljen, A., Feng, Y., Castorina, E., Contardo, G., Kreisch, C. D., Nicola, A., Alsing, J., Scoccimarro, R., Verde, L., Viel, M., Ho, S., Mallat,

S., Wandelt, B., and Spergel, D. N. (2020b). The quijote simulations. *The Astrophysical Journal Supplement Series*, 250(1):2.

[Wang et al., 2015] Wang, L., Spurzem, R., Aarseth, S., Nitadori, K., Berczik, P., Kouwenhoven, M. B. N., and Naab, T. (2015). nbody6++gpu: ready for the gravitational million-body problem. *Monthly Notices of the Royal Astronomical Society*, 450(4):4070–4080.

[Weinberg, 2008] Weinberg, S. (2008). *Cosmology*.

[Workman et al., 2022] Workman, R. L. et al. (2022). Review of Particle Physics. *PTEP*, 2022:083C01.

[Zel'dovich, 1970] Zel'dovich, Y. B. (1970). Gravitational instability: An approximate theory for large density perturbations. , 5:84–89.