Cosmic Frontiers: From the Origins to Gravitational Wave Astronomy

by

Amirhossein Dehghanizadeh

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Applied Mathematics

Waterloo, Ontario, Canada, 2025

© Amirhossein Dehghanizadeh 2025

Examining Committee Membership

The following served on the Examining Committee for this thesis. The decision of the Examining Committee is by majority vote.

External Examiner:	Maria Celeste Artale Assistant Professor, Dept. of Physics Universidad Andres Bello
Supervisor:	Ghazal Geshnizjani Associate Adjunct Professor, Dept. of Applied Mathematics University of Waterloo Faculty, Perimeter Institute for Theoretical Physics
Internal Member:	Florian Girelli Associate Professor, Dept. of Applied Mathematics University of Waterloo
Internal Member:	Michael Hudson Professor, Dept. of Physics and Astronomy University of Waterloo
Internal-External Member:	Will Percival Professor, Dept. of Physics and Astronomy University of Waterloo

Author's Declaration

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

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

Statement of Contributions

I am the sole author of chapters 1, 2, 4, 5, and 8.

Chapter 3 is based on a paper in the publication process [1] coauthored by Jerome Quintin and Ghazal Geshnizjani. The paper's general concept was initially designed by Ghazal Geshnizjani and Jerome Quintin, and I contributed to its development. The analysis was conducted in parallel by me and Jerome Quintin to ensure validity through nearly independent analyses. I performed the numerical calculations and generated all the plots, while Jerome Quintin separately reproduced them and conducted analytical estimates for verification. The initial draft was written by me and was then refined by Jerome Quintin. Ghazal Geshnizjani supervised the project, led discussions, verified the calculations, and edited the text.

Chapter 6 is also based on a paper in the publication process [2] written in collaboration with Leo Kim, Dorsa Sadat Hosseini, Alex Krolewski, Suvodip Mukherjee, and Ghazal Geshnizjani. The paper's general idea was initially proposed by Ghazal Geshnizjani and Suvodip Mukherjee, with all of us contributing to its development. I took the lead in conducting the analysis for this project. Leo Kim, Dorsa Sadat Hosseini, and I developed the initial codebase, which I subsequently expanded with numerous modifications and improvements. This resulted in a modular code capable of taking any galaxy catalog as input and generating the desired results. Then I carried out the numerical analyses and produced most of the resulting plots, with a few generated by Leo Kim and Dorsa Sadat Hosseini. They also carried out numerous validation analyses that were crucial in the early stages of the project. Additionally, Leo Kim cleaned and provided the galaxy catalogs. I was also the lead author in charge of writing the manuscript, with other collaborators contributing to the revisions. Suvodip Mukherjee, Alex Krolewski, and Ghazal Geshnizjani each guided and supervised different scientific aspects of the project, ranging from setting up the framework to construct the gravitational wave siren catalogs, handling galaxy surveys, and connecting them to the theoretical framework while verifying the calculations and results.

Chapter 7 will soon be submitted for publication as a separate paper [3] and serves as a follow-up phase to Chapter 6. Hence, the same team conducted the research, with Dorsa Sadat Hosseini leading the project, generating the results, and writing the manuscript. My contributions included proceeding the basis for initial codes and verifying the validity of some of the subsequent codes, checking numerical calculations, collaboration on cross-examining the results between the two projects and finally providing feedback to the text, ensuring the coherence of the manuscript's text.

Abstract

In this thesis, we present different aspects of cosmology, focusing on the early universe, large scale structure formation, and gravitational wave astrophysics. We begin by introducing cosmological perturbation theory as a framework to describe the origin of the universe, emphasizing its observational implications. We discuss early universe models and, motivated by the avoidance of singularities, we propose bouncing cosmologies. In particular, we discuss the Cuscuton model; which modifies gravity to generate a bounce without introducing new degrees of freedom, provides a smooth transition between contraction and expansion, and generates scale-invariant entropy perturbations. We analyze scalar perturbations and demonstrate that the model remains stable and weakly coupled. Additionally, we compute three-point functions, showing negligible non-Gaussianities on observable scales, with potential signals arising from the conversion of isocurvature to curvature perturbations.

Building on this foundation, we investigate the formation of large scale structures and study bias parameters that describe the statistical distribution of dark matter halos and galaxies. We extend this analysis to gravitational wave (GW) sources, introducing the GW bias parameter. We propose a methodology to link binary black hole (BBH) mergers to their host galaxies using phenomenological GW host-galaxy probability functions. These functions are constructed based on observed astrophysical properties of galaxies, including stellar mass, star formation rate, and metallicity. We calculate the GW bias using the angular power spectrum for photometric surveys and the 3D power spectrum for spectroscopic surveys. We find that the GW bias depends on the parameters of the host-galaxy probability functions, which are linked to the formation channels of BBHs. Future measurements of the GW bias could provide insights into the astrophysical processes governing GW sources and their relation to large-scale structure.

Lastly, we emphasize that cosmology may be entering a golden era, driven by a wealth of upcoming experiments, including the next generation of Cosmic Microwave Background (CMB) telescopes, galaxy surveys, and GW detectors. These experiments have the potential to provide hints for modifications to ΛCDM cosmology, the nature of dark matter and dark energy, black hole physics, and possible deviations from general relativity, paving the way for new theoretical and observational breakthroughs.

Acknowledgements

First and foremost, I would like to express my deepest gratitude to my supervisor, Ghazal Geshnizjani, for her invaluable guidance, support, and encouragement throughout this research. Her patience in allowing me to learn new topics and skills created a safe environment where I could step back and think deeply about physics. Beyond being a research supervisor, she was a caring mentor who valued more than just research results — always ready to offer support, even with personal challenges. Thank you for your unwavering support; I have grown immensely as both a researcher and a person during my PhD, and I have learned so much along the way.

Next, I would like to thank Jerome Quintin, who was not just a research collaborator but also a true friend. I learned so much from him, including his professionalism and dedication to research. Our many constructive discussions, with him patiently explaining various aspects of cosmology, were invaluable. I wish you all the best in your career, Jerome, and I hope we can continue our pool games in the future. I also thank Niayesh Afshordi — not only for his warm and welcoming presence in the cosmology research community here but also, together with Ghazal, for generously hosting us in their home several times, making us feel truly at home.

Many thanks to my research collaborators, Alex Krolewski and Suvodip Mukherjee, for their valuable guidance and insights throughout my research. I am also deeply grateful to my research teammates, Dorsa Sadat Hosseini and Leo Kim, whose contributions were instrumental in my work. I thank Bryden for being such a great officemate and groupmate, always bringing positivity to our workspace, and Elly for the thoughtful touches she brought to our group meetings.

I am grateful to all the members of my wonderful family, especially my mother, who may not have been physically present, but whose moral support I always felt. Additionally, life would not be the same without the presence of my dearest friends here, including — but not limited to — Atiyeh, who was not just a friend but more like a caring sister; Samira, Mojtaba, Sina, the members of the Tafsir group, Hasan, Keyvan, Greg, Parisa, Ali, Somaye, Zahra, Milad, and Negar, whose friendship and support have meant so much to me.

I appreciate all the people who created a positive and welcoming atmosphere during my time in the Applied Mathematics department at the University of Waterloo, the Waterloo Center for Astrophysics (WCA), and the Perimeter Institute (PI). In particular, I would like to thank the bistro staff at PI for the many wonderful memories and amazing food. I also appreciate the facility, administrative, and IT staff, as well as those who kept *Symmetry* running — a key computational tool for my research — all of whom made working at PI

such a pleasure. I also thank the Fields Institute for a stimulating atmosphere in the early stages of my work.

This research in this thesis was supported by a Discovery Grant from the Natural Science and Engineering Research Council of Canada (NSERC). It was also supported in part by the Perimeter Institute for Theoretical Physics. Research at the Perimeter Institute is supported by the Government of Canada through the Department of Innovation, Science and Economic Development and by the Province of Ontario through the Ministry of Colleges and Universities. This research made use of `MathGR` [4]. I thank ChatGPT, which I occasionally have used as an editing tool to refine my own writing and identify grammatical errors. I thank the Wide Field Astronomy Unit (WFAU) for publicly providing the 2MPZ and WISC data. This work used Astropy:[1] a community-developed core Python package and an ecosystem of tools and resources for astronomy [5–7]. The computational aspects of this work were performed at the computing clusters at the Perimeter Institute (PI) for Theoretical Physics and the Center for Advanced Computing (CAC) at Queen's University, supported in part by the Canada Foundation for Innovation.

[1] http://www.astropy.org

Dedication

To my mother, whose boundless sacrifices and kindness transcend words. She always dreamed of seeing me become a scientist — I hope I have made her proud.

To Mahsa Amini, Sarina Esmailzadeh, and all those who have given their lives for *Woman, Life, Freedom*. May freedom one day shine upon my beloved Iran and all those striving for it around the world.

Table of Contents

Examining Committee ii

Author's Declaration iii

Statement of Contributions iv

Abstract v

Acknowledgements vi

Dedication viii

List of Figures xiv

List of Tables xvii

1 Introduction 1

1.1 Cosmology in a nutshell 1

1.2 Walk-through guide 3

1.3 Notations and conventions 5

2 Cosmological solutions for the very early universe 6

2.1 Homogeneous and isotropic universe 6

2.2 Cosmological perturbation theory 9

2.2.1 Perturbing the FLRW universe 9

2.2.2 Two-point function 14

2.2.3 Non-Gaussianities . . . 17
2.3 Early universe models . . . 20
2.3.1 Observational constraints . . . 21
2.3.2 Horizon and flatness problems . . . 23
2.3.3 Inflationary models . . . 25
2.3.4 Alternatives; bouncing cosmology . . . 29
2.4 Singularity problem . . . 32
2.4.1 Geodesic Incompleteness . . . 32
2.4.2 Non-singular cosmologies . . . 34

3 Cuscuton Bounce Beyond the Linear Regime: Bispectrum and Strong Coupling Constraints 36

3.1 Introduction . . . 36
3.2 The Cuscuton bounce in the linear regime . . . 40
3.2.1 General setup, background equations, and linear perturbations . . . 40
3.2.2 Explicit models and analytical approximations . . . 46
3.2.3 Numerical examples . . . 60
3.3 Beyond the linear regime . . . 65
3.3.1 Theoretical insights and expectations . . . 65
3.3.2 Third-order action . . . 68
3.3.3 Toward estimating the strong coupling scale . . . 71
3.3.4 Bispectrum calculations . . . 80
3.4 Discussion and conclusions . . . 92

4 Large scale structure 95

4.1 Linear evolution; Transfer function . . . 95
4.2 Non-linear evolution; Halo and galaxy biases . . . 100

5 Gravitational waves **104**

5.1 Theory . . . 104

5.2 Sources . . . 108

5.3 Observation . . . 111

5.3.1 Interferometers . . . 112

5.3.2 Summary of observations . . . 112

6 The Gravitational Wave Bias Parameter from Angular Power Spectra: Bridging Between Galaxies and Binary Black Holes **116**

6.1 Introduction . . . 116

6.2 Modeling Gravitational Wave bias parameter from BBHs based on observed galaxies . . . 119

6.2.1 GW bias parameter . . . 119

6.2.2 Phenomenological approach: assigning mock BBHs to observed galaxies 120

6.3 Summary of the galaxy catalogs used in the analysis . . . 123

6.3.1 2MPZ and WISC catalogs . . . 123

6.3.2 Galaxy redshifts . . . 126

6.3.3 Determining the stellar masses of galaxies . . . 126

6.4 Populating galaxies with BBHs . . . 127

6.4.1 The merger rate and $N(z)$ of BBHs . . . 127

6.4.2 Host-galaxy probability for BBHs . . . 130

6.4.3 BH mass distribution . . . 132

6.5 Calculating the angular power spectra for sirens and galaxies . . . 133

6.5.1 Power spectrum measurement . . . 133

6.5.2 Model angular power spectrum . . . 136

6.6 Results for gravitational wave (GW) bias and discussion . . . 140

6.6.1 Variation in pivot mass $M_{\mathcal{K}}$ . . . 141

6.6.2 Variation in slopes δ_l and δ_h . . . 144

6.6.3 Variation with redshift . . . 146

6.6.4 Variation in $\ell_{\max}$ . . . 149
6.6.5 Comparison with previous studies . . . 149
6.7 Conclusion . . . 150

7 Gravitational Wave Bias from 3D Power Spectrum: Tracing Binary Black Hole Mergers Through Galaxy Surveys 153
7.1 Introduction . . . 153
7.2 Modeling gravitational wave bias parameter . . . 154
7.2.1 Gravitational wave bias parameter . . . 154
7.2.2 A Phenomenological framework for measuring GW bias parameters using mock GW source catalogs generated from galaxy surveys . . . 155
7.3 Galaxy catalog . . . 156
7.3.1 Sampling from SDSS DR7 . . . 157
7.3.2 Physical properties of galaxies in SDSS DR7 . . . 158
7.4 Generating mock GW source catalogs from galaxy catalog . . . 162
7.4.1 Binary merger rates and merger counts for binary black holes (BBHs) 162
7.4.2 Selecting host galaxies for BBH mergers using astrophysical host-galaxy probability functions . . . 164
7.5 3D power spectrum estimation . . . 168
7.5.1 Measuring the anisotropic power spectrum and covariance . . . 168
7.5.2 Measuring the GW bias from the power spectrum multipoles . . . 170
7.6 Discussion on the GW bias parameter results . . . 171
7.6.1 GW bias parameter exclusively derived from M_* host-galaxy probability (M_* model) . . . 172
7.6.2 Results for the GW bias parameter using the entire M_*-Z-SFR model 173
7.7 Conclusion . . . 179

8 Conclusion 180

References 183

APPENDICES 220

A Three-point functions for two locally non-Gaussian scalar fields **221**

A.1 One field 221

A.2 Two fields 223

B Redshift conversion **225**

C Testing uniform host-galaxy probability **226**

D Determining the best choice for $\ell_{\rm min}$ **228**

E Bias with a different photometric redshift convolution **230**

F 3D plots of the combined effect of δ_l, δ_h and $M_{\mathcal{K}}$ on GW bias **234**

G 1D distribution of galaxies as a function of Mass, SFR, and Metallicity **237**

H 3D power spectrum of the SDSS DR7 galaxy catalog-validation **239**

I Testing uniform M_*-Z-SFR host-galaxy probability **241**

J GW bias vs BH masses **243**

J.1 Modeling BH masses in a binary system 243

J.2 Results for GW bias vs BH masses 246

List of Figures

2.1 Tiny anisotropies on the CMB temperature map. 7

2.2 Reconstructed primordial power spectrum for curvature perturbations. . . . 22

2.3 Potential shape for inflation. 28

2.4 The joint n_s-r constraint form Planck collaboration. 29

2.5 Schematic behavior of the scale factor in three different cases. 31

3.1 Numerical background solutions for the Hubble parameter H and its time derivative $\dot{H}$ as functions of the physical time t. 61

3.2 Numerical solutions for the absolute value of the curvature perturbations $|\zeta_k(t)|$ on superhorizon scales. 63

3.3 Behavior of $k^{3/2}|\chi_k(t)|$ from the numerical solutions for the entropy perturbations, $\chi_k(t)$. 64

3.4 Results for the nonlinear corrections from $\zeta\chi^2$ bispectrum. 89

3.5 Numerical and analytical estimates of $|C^{\chi}_{\rm NL}\zeta_{\rm G}|$ and $|f^{\zeta}_{\rm NL}\zeta_{\rm G}|$. 90

3.6 The level of non-Gaussianities and shape function in Cuscuton model. 91

4.1 The transfer function for different cosmological models. 98

5.1 A schematic depiction of how test particles move under the passing of GWs. 106

5.2 Possible GW sources along with their observing tools. 109

5.3 The reconstructed waveform for the first direct GW observation. 110

5.4 Schematic illustration of the Advanced LIGO detector design. 113

5.5 Summary of the most recent LIGO-Virgo-KAGRA results. 114

5.6 The prospects of full LIGO-Virgo-KAGRA network. 115

6.1 Flowchart showing our procedure for modeling the GW bias by populating galaxy catalogs with mock BBHs. 121
6.2 2D sky maps of galaxies after masking, displayed in equatorial coordinates. . 123
6.3 Stellar mass distribution of the 2MPZ and WISC galaxy catalogs. 124
6.4 Distribution of heliocentric redshifts and Stellar mass limit for 2MPZ and WISC. 125
6.5 BBH merger rate and number of BBH mergers. 129
6.6 Stellar mass-dependent BBH host-galaxy probability. 132
6.7 The distribution of photometric and spectroscopic redshifts for the subset of the 2MPZ catalog that includes both photometric and spectroscopic redshifts. 138
6.8 The offset between the spectroscopic and photometric redshift distribution. . 139
6.9 Comparison between the photometric redshift distribution and the convolved redshift distribution. 140
6.10 Bias for varying the pivot mass $M_{\mathcal{K}}$ in the GW host-galaxy probability for 2MPZ and WISC catalogs. 143
6.11 Mean stellar mass of host galaxies for different $M_{\mathcal{K}}$ values. 143
6.12 Bias for varying δ_l and δ_h in the GW host-galaxy probability function setting. 145
6.13 Comparisons of GW bias relative to $M_{\mathcal{K}}$ for 2MPZ and WISC under different δ_l and δ_h choices. 147
6.14 The redshift dependence of GW bias. 148
6.15 Same as Fig. 6.10 but for different values of $k_{\max}$. 150

7.1 The flowchart that shows the framework for the GW bias parameter calculation. 157
7.2 Counts map of the galaxy and sirens catalogs. 158
7.3 The distribution of stellar mass, SFR, and metallicity of the galaxies in the redshift range of $0.07 \leq z < 0.2$. 160
7.4 The 2D projected distribution of the SFR, stellar mass and metallicity of all galaxies in the SDSS DR7 catalogs. 161
7.5 M_*, SFR and Z host-galaxy probability functions. 166
7.6 Comparison between our results and those of Chapter 6. 172
7.7 GW bias parameter as a function of stellar mass break $M_{\mathcal{K}}$. 175
7.8 GW bias parameter as a function of star-formation rate turnover, $\mathrm{SFR}_{\mathcal{K}}$. . . 177

7.9 GW bias parameter as a function of $Z_{\mathcal{K}}$. 178

C.1 Comparison of galaxy and the mock siren catalog power spectra. 227

C.2 Bias comparison for galaxy catalog and siren catalog based on uniform host-galaxy probability. 227

D.1 The distribution of χ^2 for 50 host-galaxy probabilities with different values of $M_{\mathcal{K}}$ for 2MPZ. 229

D.2 Like Fig. D.1, but for WISC. 229

E.1 The spread of spectroscopic redshift distribution around a particular photometric redshift. 231

E.2 Bias for 2MPZ catalog with six different photometric redshift convolution schemes. 232

E.3 Fraction of GW hosts with measured spectroscopic redshift, $f_{\rm spec}$, versus $M_{\mathcal{K}}$, for mock siren catalogs in 2MPZ. 232

F.1 Bias for varying $M_{\mathcal{K}}$, δ_l, and δ_h in Eq. (6.14) for 2MPZ catalog. 235

F.2 Bias for varying $M_{\mathcal{K}}$, δ_l, and δ_h in Eq. (6.14) for WISC catalog. 236

G.1 The distribution of galaxies as a function of stellar mass, SFR and metallicity. 238

H.1 Comparison of our galaxy power spectrum from the Ross et al measurement. 240

I.1 Number density of galaxies and sirens. 242

I.2 Ratio of sirens and galaxies power spectra. 242

J.1 GW bias parameter for varying $M_{\mathcal{K}}$ using M_*-Z-SFR model and considering different redshift and BH mass ranges. 247

List of Tables

6.1 Parameters of Eq. (6.22) fitted to the data points in Figure 6.10. 142

I.1 Number of galaxies and sirens in each redshift bin. 241

I.2 Calculated averaged ratio of sirens and galaxies biases. 242

Chapter 1

Introduction

1.1 Cosmology in a nutshell

Since ancient times, people have been mesmerized by the night sky, imagining the face of their god in the patterns of the stars. Millennia later, we are still captivated by its mysteries — only now, we seek answers not from gods, but from science. Cosmology is the science of studying the universe as a whole, its origin, evolution, structure, composition, and fate. To achieve this goal, we have come a long way, from Galileo Galilei's first use of telescopes in astronomy to the Euclid mission, designed to observe billions of galaxies, with many breakthroughs along the way.

One of the greatest revolutions in cosmology was led by Edwin Hubble, who revealed two crucial insights. First, the Milky Way is not the entire universe — it is just one of the countless galaxies. This dramatically expanded our understanding of the vastness of the universe, showing that it is orders of magnitude larger than previously believed. Second, he discovered that these galaxies are moving away from us[1], proving that the universe is not static but expanding. This discovery, together with the fact that looking at distant objects in the sky means looking into the past, transformed cosmology into a science akin to archaeology. We must search for clues and evidence using every tool available to unravel the history of the universe and, perhaps, its origins.

The discovery that the universe is expanding had a profound implication: as we move backward in time, the universe must have been denser and hotter. At some point, it was

[1] In fact, Hubble used redshift measurements made by V. M. Slipher, who had already pointed out that the positivity of redshifts indicates galaxies are moving away from us [8]. Hubble's main contribution was establishing a linear relationship between redshift and distance. This relation was also suggested by G. Lemaître, albeit using less accurate data [8].

so hot that even neutral atoms could not form. This led to the prediction and eventual observation of the cosmic microwave background (CMB), the relic light emitted when the universe cooled enough for neutral atoms to form. Observations provide us with evidence of even earlier, hotter, and denser stages of the universe, where matter was broken down into its fundamental constituents, quarks, leptons, and gauge bosons. This extreme phase is known as the *hot Big Bang*. The observation of the CMB was indeed a major breakthrough. Along with modern galaxy surveys, it marked the beginning of the era of *Precision Cosmology* — an effort to describe the entire universe with a mathematical model whose parameters can be tightly constrained using highly accurate cosmological observations. This led to the development of the standard concordance model of cosmology (ΛCDM), which describes the universe as homogeneous and isotropic, predominantly filled with two mysterious components: dark energy (Λ) and cold dark matter (CDM).

Since then, this archaeological effort has continued, refining and expanding the cosmological model. The current picture introduces an early Big Bang phase, which explains the homogeneity and isotropy of the universe while also generating small curvature fluctuations that later seeded the formation of cosmic structures. Various models have been proposed for this phase, with two major classes: one involving a period of quasi-exponentially accelerated expansion and the other proposing a contracting universe followed by a bounce into expansion. However, the physics governing these early epochs remains beyond direct experimental reach. So far, we only have indirect observational constraints from the CMB. Moreover, the standard model of particle physics lacks a candidate for the type of matter that could have driven this early Big Bang phase. As a result, much remains to be explored and discovered in this area of cosmology.

The universe then evolved from the hot Big Bang stage to the time of last scattering surface (LSS), where the universe became transparent, allowing us to observe CMB photons. Since the physics governing this evolution is well understood through the standard model of particle physics, we can predict the features of the CMB radiation. In addition, at some point between the hot Big Bang and the LSS, the universe's energy density transitioned from a radiation-dominated phase to a matter-dominated phase. After CMB photons started to stream freely, the universe entered a long period of cosmic darkness, during which CMB photons traveled freely, and the matter in the universe gradually began to form structures. This phase is also modeled by general relativity, the standard theory of gravity. Around 13 billion years ago, known as the Cosmic Dawn, the first stars formed. About 10 billion years ago, referred to as Cosmic Noon (3 billion years after Cosmic Dawn), the universe experienced the peak of star formation. Finally, around 5 billion years ago, the energy density of the universe underwent another transition, becoming dominated by dark energy, which accelerated the expansion of the universe. Throughout the history of the universe,

many fascinating physical phenomena have occurred. Here we are on Earth, an insignificant corner of the cosmos, observing their signals, from radio waves to gamma rays.

We have been fortunate enough to witness yet another major observational breakthrough in astrophysics. In 2015, the first direct detection of gravitational waves — ripples in space-time itself — initiated a new era in multi-messenger astronomy. Unlike almost all previous observations[2], which relied on light, gravitational waves opened an entirely new window into the universe. This breakthrough provided us with a powerful tool for studying dark objects, most notably the mysterious black holes. Gravitational waves offer valuable insights into the extreme gravitational fields surrounding black holes and may even reveal hints of modified gravity or quantum gravity.

Despite many successes in cosmology over the years, numerous open questions remain. Here, we list some of the most important ones without further explanation: cosmological singularity, the true geometry and topology of the universe, the role of quantum gravity and trans-Planckian physics in cosmology, primordial cosmology, primordial tensor perturbations, decoherence in the early universe, the reheating phase, CPT violation, baryon asymmetry, neutrino physics, primordial black holes, CMB anomalies, CMB dipole and its relation to isotropy, the nature of dark energy, the nature of dark matter, modifications to general relativity, very high-redshift galaxies, puzzles in dark matter halo and galaxy formation, the formation mechanism of supermassive black holes, black hole physics and the information paradox, sources of gravitational waves, and tensions in measuring cosmological parameters such as the H_0 and S_8 tensions.

In this thesis, we study a few of these open questions in cosmology. We present possible models for the early Big Bang phase and identify the current observables that can constrain the parameters of these models. This study is primarily motivated by the goal of resolving the singularity problem in cosmology. Additionally, we investigate the potential insights that gravitational wave sources can offer, focusing in particular on the cosmological information encoded in their spatial clustering. With this brief overview, the next subsection provides a more detailed guide for the reader on how to navigate this thesis.

1.2 Walk-through guide

This thesis mainly covers two topics. The first, discussed in Chapters 2 and 3, focuses on early universe cosmology. Chapter 2 introduces the basics of the standard model of cosmology. We then review cosmological perturbation theory up to third order in perturbations,

[2]We have some additional astronomical information coming from neutrinos and cosmic rays.

exploring the possible observables and reporting the current status of observational constraints. Following that, we present examples of models compatible with the observations, and in the final section, we briefly review the singularity problem and explain why bouncing cosmology could provide a resolution.

While Chapter 3 builds on the background laid out in Chapter 2, it is based on a paper in the publication process [1] that can be read independently, assuming the reader is familiar with the context. In this chapter, we discuss a specific bouncing model, the Cuscuton bounce, and provide the calculations for the bispectrum and the non-Gaussianity signal. We demonstrate that this scenario is a robust, stable, weakly coupled, and nonsingular cosmological model, consistent with observations.

We then change the topic to the late universe and, as an intermediate step, briefly discuss the formation of large-scale structure in Chapter 4 to illustrate how the universe evolves from the early stages to its current state, filled with galaxies and stars. We introduce a parameter called galaxy bias, which summarizes the relationship between galaxies and dark matter halos. This quantity is used in the following chapter.

Next we review the basics of gravitational waves (GWs) in Chapter 5. We first present the underlying theory for the formation and propagation of GWs. We then explore the possible sources that generate GWs and discuss how these sources may have formed throughout cosmic history. We finish this chapter by reporting the status of current and future GW observations.

In Chapter 6, we introduce a new cosmological parameter called GW bias, which is similar to the previously discussed galaxy bias but applies to a different type of tracers —astrophysical GW sources. We present a novel method for theoretically modeling GW bias using observed galaxy catalogs and demonstrate how these two parameters can differ in various scenarios for the formation of binary black holes. This chapter is also based on a paper accepted for publication [2], which can be read independently, assuming the reader is familiar with the context.

Next, Chapter 7 extends the analysis from Chapter 6 by incorporating spectroscopic surveys, allowing access to galaxies' star formation rate and metallicity. This chapter will soon be submitted as a separate paper [3], so readers in the future may wish to refer to the most recent published version for the latest version.

Finally, in Chapter 8, we conclude the thesis and discuss potential future directions.

1.3 Notations and conventions

Throughout this thesis, we use the plus metric signature $(-,+,+,+)$ and units where the speed of light and Planck's reduced constant are set to unity, $c = \hbar = 1$. The reduced Planck mass is defined in terms of Newton's gravitational constant as $M_{\rm Pl} \equiv (8\pi G_{\rm N})^{-1/2}$, and we use $t_{\rm Pl} \equiv 1/M_{\rm Pl}$ as the Planck time. Finally, Greek tensorial indices run over spacetime coordinates and are contracted with the spacetime metric $g_{\mu\nu}$, while latin tensorial indices run over spatial coordinates only. We use t and τ for cosmic time and conformal time respectively. Accordingly, the following defines derivatives with respect to them; $\dot{} \equiv d/dt$, $' \equiv d/d\tau$, $\ddot{} \equiv d^2/dt^2$, and $'' \equiv d^2/d\tau^2$.

Chapter 2

Cosmological solutions for the very early universe

2.1 Homogeneous and isotropic universe

Studying the vast and complex universe is a difficult task, however, as observations guide us, we can make simplifying assumptions. The main assumption, known as the *Cosmological Principle*, states that, on average, no point in the cosmos is special (homogeneity) and there are no preferred directions (isotropy). In addition to philosophical arguments, observational evidence supports this assumption. The most prominent is the Cosmic Microwave Background (CMB) radiation. Around 10^5 years after the Big Bang (approximately 13 billion years ago), the photon gas in the universe decoupled from the matter and began an almost interaction-free journey, propagating in all directions. We observe this radiation today and find that it has the same temperature in all directions. The measured temperature is 2.7260 ± 0.0013 K [9]. The latest CMB observations from the *Planck* satellite have confirmed this uniformity with remarkable precision, at the level of μK [10], see Figure 2.1. The variance of temperature fluctuations is around tens of μK, and the largest fluctuations are about 100μK.

As Einstein theory of gravity suggests, the geometry of the universe is connected to the matter and energy distribution in the universe. Therefore, homogeneity and isotropy induce translational and rotational symmetries in geometry. Mathematically, this translates to a metric that is only a function of time, known as Friedmann-Lemaître-Robertson-Walker (FLRW) metric. With the flat FLRW metric, the line element is

$$ds^2 = g_{\mu\nu}dx^\mu dx^\nu = -dt^2 + a(t)^2\delta_{ij}dx^i dx^j = a(\tau)(-d\tau^2 + dx^i dx^j)\,, \tag{2.1}$$

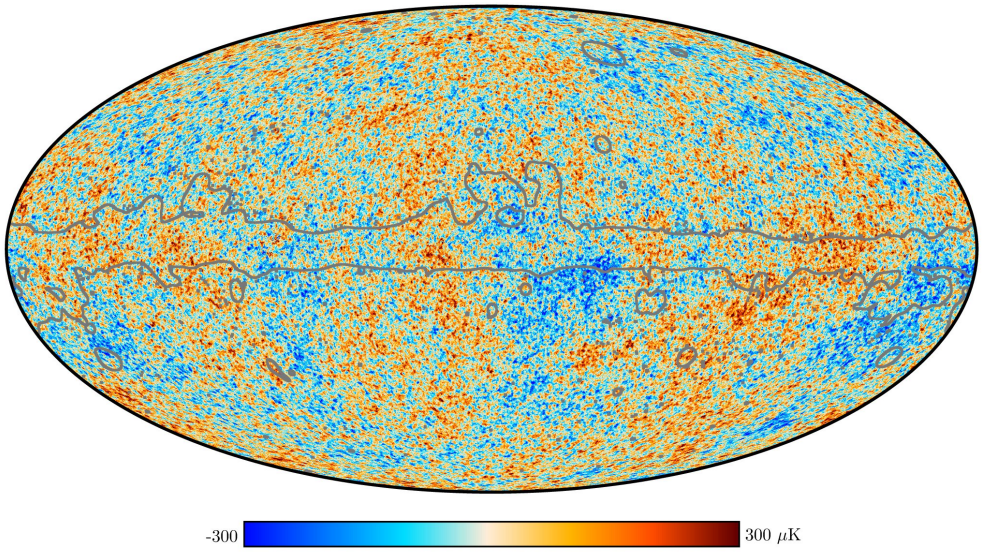


Figure 2.1: Tiny anisotropies on the CMB temperature map [10]. The variance of temperature fluctuations is around tens of μK, and the largest fluctuations are about 100μK. Note that in this figure, the dipole due to the Earth's motion with respect to the CMB frame is subtracted. The gray outline shows the extent of the confidence mask. See here for the full list of *Planck* publications.

where t, and τ are known as cosmic time, and conformal time respectively where $dt^2 = a^2 d\tau^2$. First, let us introduce the following parameters that will appear frequently in the equation.

$$H \equiv \frac{\dot{a}}{a}\,, \qquad \mathcal{H} = \frac{a'}{a}\,, \quad \Rightarrow \mathcal{H} = aH\,,$$
$$\epsilon \equiv -\frac{\dot{H}}{H^2} = 1 - \frac{\mathcal{H}'}{\mathcal{H}^2}\,, \tag{2.2}$$

where H(or $\mathcal{H}$) is the Hubble (conformal Hubble) parameter, $\dot{} \equiv d/dt$ is the derivative with respect to cosmit time (t), $' \equiv d/d\tau$ is the derivative with respect to conformal time (τ), and ϵ is known as the equation of state parameter for the reason that will be explained shortly.

Similar to the metric, the mathematical description of matter in the universe should also respect the cosmological principle. Therefore, the matter energy-momentum tensor could be fairly described by a homogeneous and isotropic perfect fluid as

$$T^{\nu}_{\mu} = \mathrm{diag}(-\bar{\rho}, \bar{p}\ \delta^j_i)\,, \tag{2.3}$$

where $\bar{\rho}$, and $\bar{p}$ are the total energy density and the total pressure which are only functions of time. Substituting both the metric and the energy-momentum tensor into the Einstein equation, we obtain the following (known as Friedmann equations within the flat universe),

$$3M_{\rm Pl}^2 H^2 = \bar{\rho}\,, \qquad 2M_{\rm Pl}^2 \dot{H} = -(\bar{\rho} + \bar{p})\,. \tag{2.4}$$

From the conservation of energy-momentum tensor ($\nabla_\mu T^\mu_\nu = 0$), we can derive the following conservation equation. However, this is not an independent equation and can be derived from the two Friedmann equations.

$$\dot{\bar{\rho}} + 3H(\bar{\rho} + \bar{p}) = 0\,. \tag{2.5}$$

The first Friedmann equation is just a constraint equation, and upon identifying the type of the perfect fluid (a relation between $\bar{\rho}$, and $\bar{p}$), we can express the second equation only in terms of $a(t)$, and its time derivative. Therefore, there is only one degree of freedom (how the universe is expanding), manifested through $a(t)$.

The relation between $\bar{\rho}$, and $\bar{p}$ is known as the equation of state. We consider a barotropic fluid, which assumes that the two are proportional, i.e. $\bar{p}{=}w\bar{\rho}$, where w is called the equation of state parameter. In principle, the universe can be filled with different components with their own equation of state, however, here we assume one component is dominant over the others. By plugging the former equation into the conservation equation and integrating, we have

$$\bar{\rho} \propto a^{-3(1+w)}\,, \tag{2.6}$$

and lastly plugging this into the first Friedmann equation and integrating, we obtain

$$a \propto |t|^{\frac{2}{3(1+w)}}\,, \tag{2.7}$$

except for the case of $w = -1$[1] in which $a \propto \exp(Ht)$. It is worth noting that the Friedmann equations can be rewritten by dividing the second equation by the first equation, yielding the following form,

$$\epsilon = \frac{3}{2}(1 + w)\,. \tag{2.8}$$

It is evident that there is a one-to-one relation between ϵ, and w, hence the former is also called the equation of state parameter.

Instead of a perfect fluid, we can model the matter content by introducing some fields at the level of action. In particular, scalar fields are suggested by most early universe models (we discuss this in more details in Section 2.3.3). Here, we show the simple case of one scalar field as

$$S = \int d^4\boldsymbol{x}\sqrt{-g}\left(\frac{M_{\rm Pl}^2}{2}R - \frac{1}{2}\partial_\mu\bar{\phi}\partial^\mu\bar{\phi} - V(\bar{\phi})\right)\,, \tag{2.9}$$

[1]note for $w < -1$, this equation is still valid but the range for time is $[t_{\rm initial}, 0]$. The scale factor then diverges at $t = 0$, which means that the universe reaches a big-rip end in finite time. This is a consequence of an expansion that is even faster than the exponential expansion.

where $V(\phi)$ is the potential. One can derive the equation of motion by varying the action or, equivalently, show that the following holds for a purely time-dependent field.

$$\bar{\rho} = \frac{1}{2}\dot{\bar{\phi}}^2 + V(\bar{\phi})\,, \qquad \bar{p} = \frac{1}{2}\dot{\bar{\phi}}^2 - V(\phi)\,. \tag{2.10}$$

By substituting these into the Friedmann equations, one can obtain solutions for a scalar field. Two interesting cases arise: (i) when kinetic energy dominates over the potential, corresponding to $w \simeq 1$; and (ii) when the potential dominates, resembling $w \simeq -1$. In addition, by substituting these into the continuity equation, one can derive the following for the evolution of $\bar{\phi}$

$$\ddot{\bar{\phi}} + 3H\dot{\bar{\phi}} + V'(\bar{\phi}) = 0\,. \tag{2.11}$$

This would be useful later.

Let us briefly introduce a modification to the flat FLRW metric in Eq. (2.1). We can introduce spatial curvature, meaning that three-dimensional spatial hypersurfaces do not need to be flat while still preserving homogeneity and isotropy. Therefore, the metric can be rewritten as

$$g_{\mu\nu}dx^\mu dx^\nu = -dt^2 + a(t)^2\left(\frac{dr^2}{1-kr^2} + r^2\left(d\theta^2 + \sin(\theta)^2 d\phi^2\right)\right)\,, \tag{2.12}$$

where (r, θ, ϕ) are the spherical coordinates, and k can take the values of 0, 1, or -1 corresponding to flat, positively curved, and negatively curved spatial geometries, respectively. However, observations constrain the contribution of spatial curvature to be very small [11], so it is often neglected in the literature.

2.2 Cosmological perturbation theory

2.2.1 Perturbing the FLRW universe

The homogeneous and isotropic universe is an approximation that is only valid on large scales up to leading order. Therefore, we should account for small perturbations on top of the homogeneous and isotropic background. We follow the standard cosmological perturbation theory in [12–16]. We redefine the metric as

$$g_{\mu\nu}(\tau, \boldsymbol{x}) = \bar{g}_{\mu\nu}(\tau) + \delta g_{\mu\nu}(\tau, \boldsymbol{x})\,, \tag{2.13}$$

where $\bar{g}_{\mu\nu}(\tau)$ is the FLRW metric previously introduced in in Eq (2.1), and $\delta g_{\mu\nu}(\tau, \boldsymbol{x})$ represent the perturbative corrections.

Scalar-vector-tensor decomposition Similar to the background metric, metric perturbations can be decomposed into temporal and spatial parts. On top of that, the perturbations can be decomposed into scalar, vector, and tensor parts as follows.

$$\begin{aligned}\delta g_{\mu\nu}dx^\mu dx^\nu =& a(\tau)^2\big[-2Ad\tau^2+2(\partial_i B+S_i)dx^i d\tau \\ &+(2C\delta_{ij}+2\partial_i\partial_j E+2\partial_{(i}F_{j)}+h_{ij})dx^i dx^j\big]\,.\end{aligned} \tag{2.14}$$

We list all the 10 DoFs (degrees of freedom) of this metric,

- 4 scalar DoFs: $A(\boldsymbol{x},\tau)$, $B(\boldsymbol{x},\tau)$, $C(\boldsymbol{x},\tau)$, and $E(\boldsymbol{x},\tau)$.
- 4 vector DoFs: since both $F_i(\boldsymbol{x},\tau)$ and $S_i(\boldsymbol{x},\tau)$ are transverse vectors i.e. $\partial_i F^i = \partial_i S^i = 0$, each of them possesses two DoFs.
- 2 tensor DoFs: since $h_{ij}(\boldsymbol{x},\tau)$ is a transverse traceless tensor i.e. $\partial_i h^i_j = 0$, and $h^i_i = 0$, it possesses two DoFs. In fact, these represent the two intrinsic DoFs in general relativity as a spin-2 field theory. As we discuss later in this chapter, and in more detail in Chapter 5, they can also be interpreted as the two polarization states of propagating waves.

This is particularly useful, since the scalar, vector, and tensor perturbations evolve independently of each other at the linear level. Although, here we showed the decomposition around the FLRW metric, the perturbations can be introduced around any background metric. In addition, this can be translated to the well-known Arnowitt–Deser–Misner (ADM) decomposition. For more information, see [17, 18]. Similarly to metric perturbations, we need to perturb the matter content around its background value. For a canonical scalar field introduced in Eq (2.9) (with $\bar{\phi}$ replaced by ϕ), we simply have

$$\phi(\tau,\boldsymbol{x}) = \bar{\phi}(\tau) + \delta\phi(\tau,\boldsymbol{x})\,. \tag{2.15}$$

Gauge degrees of freedom The 10 DoFs introduced in Eq. (2.14) are not all physical; four of them, known as gauge artifacts, can be eliminated due to diffeomorphism invariance. In other words, when transforming from one coordinate system to another, a parameter (metric or fields perturbations) might show artificial fluctuations. In general, there are two approaches (I) one can work with gauge invariant quantities, (II) or one can fix the gauge. We briefly introduce both.

The gauge transformation, or small transformation of the coordinate system, is in the following form

$$x^\mu \to \tilde{x}^\mu = x^\mu + \xi^\mu\,. \tag{2.16}$$

ξ^μ can be decomposed into two scalars; ξ^0 and ξ, and two vector DoFs; $\xi^i_{\rm T}$ as follows

$$\xi = (\xi^0, \xi^i) \qquad \xi^i = \xi^i_{\rm T} + \partial^i \xi \,, \tag{2.17}$$

where $\xi^i_{\rm T}$ is the transverse part ($\partial_i \xi^i_{\rm T} = 0$). There is no tensorial DoF in ξ^μ, therefore, the tensor perturbations of Eq (2.14) are gauge invariant.

In this thesis, we only show the equations for the scalar and tensor perturbations, however, those of vector perturbations are straightforward to derive in similar steps. Under the gauge transformation of Eq (2.16), the scalar perturbations of the metric in Eq (2.14) are transformed as [2]

$$\begin{aligned} A &\to A - (\xi^0)' - \mathcal{H}\xi^0 \,, \\ B &\to B + \xi^0 - \xi' \,, \\ C &\to C - \mathcal{H}\xi^0 \,, \\ E &\to E - \xi \,. \end{aligned} \tag{2.18}$$

Similarly, for a canonical scalar field with a scalar potential, the perturbations transform as

$$\delta\phi \to \delta\phi - (\bar{\phi})'\xi^0 \,, \tag{2.19}$$

As mentioned previously, 2 of the 4 DoFs in the Eq (2.14) can be gauged away, and there are two ways to achieve this. One can define gauge invariant quantities, i.e. a quantity that preserves its functional form under any gauge transformation. These quantities are not unique and one can define them in different ways, depending on which one is suitable for the problem. Here we just introduce two of the most famous gauge-invariant quantities known as the Bardeen potentials,

$$\begin{aligned} \Phi &\equiv A + \frac{1}{a}\left[a(B - E')\right]' \,, \\ \Psi &\equiv -C - \mathcal{H}(B - E') \,. \end{aligned} \tag{2.20}$$

It is evident that 4 scalars of the metric, i.e. A, B, C, and E are compressed into 2 new scalars Φ, and Ψ, reducing the Dofs.

[2] Note for a general rank (0,2) tensor, the gauge transformation is

$$\widetilde{\delta O}_{\mu\nu} = \delta O_{\mu\nu} - \xi^\rho{}_{,\mu}\bar{O}_{\rho\nu} - \xi^\sigma{}_{,\nu}\bar{O}_{\mu\sigma} - \bar{O}_{\mu\nu,\alpha}\xi^\alpha \,,$$

where $O_{\mu\nu} = \bar{O}_{\mu\nu} + \delta O_{\mu\nu}$. The second, and third terms on the right-hand side come from the general coordinate transformation, and the last term comes from the small displacement of points. See [16] and this for more information.

On the other hand, the more common method for gauging away the unphysical DoFs is to simply choose a gauge by fixing some of these parameters. As long as by putting any constraint on the parameter space, two DoFs disappear, it provides a viable gauge; however, some gauges are more convenient in some cases. Here we introduce the two most common gauges. First, the *conformal Newtonian gauge* is built by setting $B = E = 0$. This translates to $\Phi = E$, and $\Psi = -C$, and the scalar part of Eq (2.14) reads

$$\delta g_{\mu\nu} dx^\mu dx^\nu = a(\tau)^2[-2\Phi d\tau^2 - 2\Psi \delta_{ij} dx^i dx^j] \,. \tag{2.21}$$

This is in particular useful for late-time cosmology as in the Newtonian limit Φ translates to the classical Newtonian potential. In addition, in most cases where there is no anisotropic stress, it can be shown that $\Psi = \Phi$.

Since the scalar perturbations of the metric are related to the matter perturbations through the Einstein equation, one can equivalently choose the gauge by fixing the matter perturbations. In particular, setting $E = 0$, and $\delta\phi = 0$[3] gives the so called *comoving gauge*. Since the value of C in this gauge is greatly useful (for most of the models, it is conserved on super-Hubble scales), it is given a new name as $\zeta \equiv -C$. Therefore, the scalar part of Eq (2.14) reads

$$\delta g_{\mu\nu} dx^\mu dx^\nu = a(\tau)^2[-2A d\tau^2 + 2\partial_i B dx^i d\tau - 2\zeta \delta_{ij} dx^i dx^j] \,. \tag{2.22}$$

ζ then, in our definition, represents the spatial part of the perturbed metric in the comoving gauge. It is shown that on large scales, ζ is equivalent to the curvature perturbation on uniform density hypersurfaces [19], hence, it is often called "comoving curvature perturbations", or shorter "curvature perturbations". This gauge is broadly used in the context of early-universe cosmology.

Equations of motion Once the gauge degrees of freedom have been eliminated, the next step is to determine the dynamics of the remaining ones via the equations of motion (EoM). Generally, there are two ways to obtain EoM, one can simultaneously perturb the Einstein tensor, and the energy-momentum tensor to solve the Einstein equation order by order. Equivalently, one can perform the perturbative expansion at the level of action. In other words, we schematically have $S = S^{(0)} + S^{(1)} + S^{(2)} + ...$, where each order is derived by considering the same order of perturbations appearing in the action. Next, by applying the

[3]To be more precise, the original definition of this gauge is to set $\delta u = 0$, where $\delta u_i = \partial_i \delta u$, and δu_i is the perturbations of 4-velocity in a perfect fluid. This is why it is called comoving. On large scales, this is equivalent to set $\delta\rho = 0$, where $\delta\rho$ is the perturbations to the energy density of the perfect fluid.

action principle on the action up to $N+1$ order, the N^{th} order of EoM is achieved[4]. The latter method has the advantage of allowing for quantizing the action. The full calculation for EoM for different DoFs (metric or field perturbations) is relatively long and can be derived by incorporating the ADM formalism (see [17, 20, 21] for more details[5]). In particular, the form of the EoMs can vary depending on the type of matter present in the universe. Therefore, we only report the final result for the action introduced in Eq (2.9) (with $\bar{\phi}$ replaced by ϕ) in the comoving gauge.

For the metric represented in Eq (2.22), we can vary the action with respect to three parameters; A, B, and ζ. Variation with respect to the first two gives the following

$$\begin{aligned} A =& \frac{\zeta'}{\mathcal{H}} \\ \frac{B}{a} =& -\frac{\zeta}{\mathcal{H}} - \epsilon\nabla^{-2}\zeta' + \frac{3\left((\bar{\phi})' + 2a^2V(\bar{\phi}) - 6M_{\rm Pl}^2\mathcal{H}^2\right)}{4M_{\rm Pl}^2\mathcal{H}}\nabla^{-2}\zeta' = -\frac{\zeta}{\mathcal{H}} - \epsilon\nabla^{-2}\zeta' \,, \end{aligned} \tag{2.23}$$

where ϵ is given in Eq (2.2), and in the second line, the background EoM in Eq (2.4) is used to simplify the expression. ∇^{-2} is the inverse Laplacian operator i.e. $\nabla^{-2}\nabla^2$ is the identity operator (note $\nabla^2 = \partial_i\partial^i$). An important feature of the above equations is that A, and B are non-dynamical (non-propagating). Even though, here we showed this for a particular action, this could be quite generic as in ADM formalism, two EoM turn out to be Hamiltonian and momentum constraints.

To summarize, we started with a metric with 4 scalar DoFs. Introducing a scalar field to the action, provided the 5th scalar DoF. Then, we eliminated two unphysical DoFs by fixing a gauge, and two more were found to be just non-dynamical i.e. their EoM turned out to be constraint equations. Therefore, this leaves only one dynamical scalar DoF which here manifests itself through ζ. By substituting the solution to A, and B back into the action, using background EoMs, and simplifications through integration by parts, one can rewrite the action only in terms of ζ,

$$S^{(2)}_{\text{scalar}} = M_{\rm Pl}^2 \int \mathrm{d}\tau \mathrm{d}^3\boldsymbol{x}\, a^2\epsilon \left((\zeta')^2 - \partial_i\zeta\partial^i\zeta\right) \,, \tag{2.24}$$

which give the following EoM

$$\zeta'' + 2\frac{z'}{z}\zeta' - \nabla^2\zeta = 0 \,, \tag{2.25}$$

[4]Equivalently one can apply the Euler-Lagrange equation for $S^{(N)}$. However, these two statements are essentially the same, since the Euler-Lagrange equation is derived by adding small perturbations to DoFs in the Lagrangian which acts like increasing the order of perturbations by 1.

[5]You can also use *"MathGR"* which is a *Mathematica* package developed for the same purpose for arbitrary actions [4].

where z is defined as follows to make a canonical available as will be explained shortly.

$$z^2 \equiv 2M_{\rm Pl}^2 a^2 \epsilon \,. \tag{2.26}$$

Similarly, the second-order action for tensor perturbations reads

$$S^{(2)}_{\rm tensor} = \frac{M_{\rm Pl}^2}{8} \int {\rm d}\tau {\rm d}^3\boldsymbol{x}\, a^2 \left(h'_{ij}(h^{ij})' - \partial_l h_{ij} \partial^l h^{ij} \right) \,. \tag{2.27}$$

2.2.2 Two-point function

Quantization In most recent models of primordial cosmology, small perturbations in the universe's curvature, denoted by ζ (which can be transformed into energy density perturbations via a gauge transformation), are considered to arise from quantum fluctuations at small scales. These perturbations are subsequently stretched to scales larger than the Hubble radius (the typical curvature radius of the universe) as the background geometry evolves, resulting in the current density fluctuations observed in the universe today. Here, we explain the mathematical framework underlying this process.

Quantizing the fields in the early universe can naturally provide fluctuations (see [22] or the last chapter of [23] for a review of quantum fields in curved spacetime). Since ζ is the only DoF here, we start by quantizing the action in Eq (2.24). Following the standard methodology of quantum field theory, we transfer this action to the Fourier space which provides a system of independent harmonic oscillators. In addition, to create canonical harmonic oscillators, we perform a field redefinition as follows

$$\nu \equiv z\zeta \,, \tag{2.28}$$

where z is defined in Eq (2.26), and ν is known as *Sasaki-Mukhanov* variable. Then the action reads

$$S^{(2)}_{\rm scalar} = \frac{1}{2} \int {\rm d}\tau {\rm d}^3\boldsymbol{k} \left((\nu'_{\boldsymbol{k}})^2 - (k^2 - \frac{z''}{z})\nu^2_{\boldsymbol{k}} \right) \,, \tag{2.29}$$

where

$$\nu(\tau, \boldsymbol{x}) = \int \frac{{\rm d}^3\boldsymbol{k}}{(2\pi)^3} \nu_{\boldsymbol{k}}(\tau) e^{i\boldsymbol{k}.\boldsymbol{x}} \,, \tag{2.30}$$

and $k^2 = \boldsymbol{k}.\boldsymbol{k}$. Lastly, we arrived at a set of canonical harmonic oscillators[6] which we know how to quantize. The only difference with quantization of a free field in Minkowski spacetime is that here the frequency of the oscillators is time-dependent. Therefore we propose

[6]Note when the frequency of oscillation becomes imaginary, i.e. when $k^2 < z''/z$, they stop oscillating. However, we still stick to the standard terminology of quantum mechanics and call them harmonic oscillators.

the following ansatz for $\nu_{\boldsymbol{k}}$ to upgrade it to a quantum operator,

$$\hat{\nu}_{\boldsymbol{k}}(\tau) \equiv \nu_k(\tau)\hat{a}^-_{\boldsymbol{k}} + \nu^*_k(\tau)\hat{a}^+_{-\boldsymbol{k}}\,, \tag{2.31}$$

where $\nu_k(\tau)$ (known as *mode functions*) carries the information about the time evolution, and $\hat{a}^-_{\boldsymbol{k}}$ and $\hat{a}^+_{\boldsymbol{k}}$ carries the information about non-commutativity of quantum operators. Note, as we show later the EoM for ν_k will be a function of k (not $\boldsymbol{k}$), therefore we use ν_k instead of $\nu_{\boldsymbol{k}}$. Inserting this into Eq (2.30) gives the real-space ansatz for $\hat{\nu}(\tau, \boldsymbol{x})$,

$$\hat{\nu}(\tau, \boldsymbol{x}) = \int \frac{\mathrm{d}^3\boldsymbol{k}}{(2\pi)^3}[\nu_k(\tau)\hat{a}^-_{\boldsymbol{k}}e^{i\boldsymbol{k}.\boldsymbol{x}} + \nu^*_k(\tau)\hat{a}^+_{\boldsymbol{k}}e^{-i\boldsymbol{k}.\boldsymbol{x}}]\,. \tag{2.32}$$

We then impose the following commutation relations,

$$[\hat{\nu}(\tau, \boldsymbol{x}), \hat{\pi}(\tau, \boldsymbol{y})] = i\delta^3(\boldsymbol{x} - \boldsymbol{y})\,, \qquad [\hat{\pi}(\tau, \boldsymbol{x}), \hat{\pi}(\tau, \boldsymbol{y})] = [\hat{\nu}(\tau, \boldsymbol{x}), \hat{\nu}(\tau, \boldsymbol{y})] = 0\,, \tag{2.33}$$

where $\hat{\pi}(\tau, \boldsymbol{x}) = \hat{\nu}'(\tau, \boldsymbol{x})$. This translates to

$$[\hat{a}^-_{\boldsymbol{k}_1}, \hat{a}^+_{\boldsymbol{k}_2}] = (2\pi)^3\delta^3(\boldsymbol{k}_1 - \boldsymbol{k}_2)\,, \qquad [\hat{a}^-_{\boldsymbol{k}_1}, \hat{a}^-_{\boldsymbol{k}_2}] = [\hat{a}^+_{\boldsymbol{k}_1}, \hat{a}^+_{\boldsymbol{k}_2}] = 0\,, \qquad \nu'_k\nu^*_k - \nu_k\nu^{*'}_k = -i\,, \tag{2.34}$$

where the last equality is known as the Wronskian condition. The EoM for mode functions reads

$$\nu''_k + (k^2 - \frac{z''}{z})\nu_k = 0\,. \tag{2.35}$$

Upon identifying the functional form of $z(\tau)$, this equation can be solved, thereby solving the system of quantum harmonic oscillators as well. Lastly, by considering $\hat{a}^-_{\boldsymbol{k}}|0\rangle = 0$, we can evaluate the *two-point correlation function* as follows

$$\begin{aligned}\langle\hat{\nu}_{\boldsymbol{k}_1}\hat{\nu}_{\boldsymbol{k}_2}\rangle =& \langle 0|\hat{\nu}_{\boldsymbol{k}_1}\hat{\nu}_{\boldsymbol{k}_2}|0\rangle = \nu_{k_1}\nu^*_{k_2}\langle 0|[\hat{a}^-_{\boldsymbol{k}_1}, \hat{a}^+_{\boldsymbol{k}_2}]|0\rangle = (2\pi)^3\nu_{k_1}\nu^*_{k_2}\delta^3(\boldsymbol{k}_1 - \boldsymbol{k}_2)\\ \equiv& (2\pi)^3 P_\nu(k_1)\delta^3(\boldsymbol{k}_1 - \boldsymbol{k}_2)\,,\end{aligned} \tag{2.36}$$

where the second line is the definition of the *power spectrum*. It is conventional to also define the dimensionless power spectrum as

$$\mathcal{P}_\nu(k, \tau) \equiv \frac{k^3}{2\pi^2}P_\nu(k, \tau) = \frac{k^3}{2\pi^2}|\nu_k(\tau)|^2\,, \tag{2.37}$$

which for ζ translates to

$$\mathcal{P}_\zeta(k, \tau) = \frac{k^3}{2\pi^2}\frac{|\nu_k(\tau)|^2}{z(\tau)^2}\,. \tag{2.38}$$

Similarly, we perform the following Fourier transform for tensor perturbations,

$$h_{ij}(\tau, \boldsymbol{x}) = \int \frac{\mathrm{d}^3\boldsymbol{k}}{(2\pi)^3}\sum_{\lambda=+,\times}\epsilon^{(\lambda)}_{ij}h_{\boldsymbol{k},\lambda}(\tau)e^{i\boldsymbol{k}.\boldsymbol{x}}\,, \tag{2.39}$$

where $h_{\boldsymbol{k},+}(\tau)$, and $h_{\boldsymbol{k},\times}(\tau)$ are the two polarization states and $\epsilon_{ij}^{(\lambda)}$ satisfy

$$k^i\epsilon_{ij}^{(\lambda)} = 0\,, \qquad g^{ij}\epsilon_{ij}^{(\lambda)} = 0\,, \qquad g^{im}g^{jn}\epsilon_{ij}^{(\lambda)}\epsilon_{mn}^{(\lambda')} = 2\delta^{\lambda\lambda'}\,, \tag{2.40}$$

so that h_{ij} remains transverse and traceless. Similar steps to those for ζ can be followed; however, the action for tensor perturbations differs, leading to a modified definition of the Sasaki-Mukhanov variable as

$$u_k \equiv z_{\rm t} h_k\,, \qquad z_{\rm t} \equiv \frac{M_{\rm Pl}}{2}a\,, \tag{2.41}$$

where we dropped the polarization index for simplicity. Eq (2.29), and Eq (2.35) still hold (ν_k, and z replaced by u_k, and $z_{\rm t}$ respectively), and the power spectrum evaluates to

$$\mathcal{P}_{\rm t}(k,\tau) = \frac{k^3}{\pi^2}\frac{|u_k(\tau)|^2}{z_{\rm t}(\tau)^2}\,, \tag{2.42}$$

where an extra factor of 2 is multiplied because of two states of polarization. $u_k(\tau)$ is also governed by a similar EoM as follows

$$u_k'' + (k^2 - \frac{z_{\rm t}''}{z_{\rm t}})u_k = 0\,. \tag{2.43}$$

Then, the *tensor-to-scalar ratio* is also defined as

$$r \equiv \frac{\mathcal{P}_{\rm t}}{\mathcal{P}_\zeta} = 2\frac{|u_k|^2}{|\nu_k|^2}\left(\frac{z_{\rm t}}{z}\right)^{-2}\,. \tag{2.44}$$

So far we remained agnostic about the background evolution which enters the picture through the parameter $z(\tau)$, and $z_{\rm t}(\tau)$ (which are essentially functions of $a(\tau)$). In Section 2.3, we explain how different background cosmologies can create a power spectrum that matches the observations.

Initial conditions The last piece of the puzzle to solve the EoM in Eq (2.35) is the initial condition. Let us first examine Eq (2.35) in two different regimes,

$$\begin{aligned} k^2 \gg |z''/z| \quad &\Rightarrow \quad \nu_k'' + k^2\nu_k \simeq 0\,, \\ k^2 \ll |z''/z| \quad &\Rightarrow \quad \nu_k'' - \frac{z''}{z}\nu_k \simeq 0\,, \end{aligned} \tag{2.45}$$

where they are known as sub-Hubble, and super-Hubble regimes respectively. The reason for this terminology is that for vanilla inflationary models (canonical, single field, slow roll), z''/z approximately corresponds to $\mathcal{H}^2$. However, in more exotic models, the situation may be more complicated. Therefore, in vanilla models $k^2 \gg |z''/z|$ translates to $\lambda_{\rm comoving} \propto k^{-1} \ll$

$|\mathcal{H}|^{-1}$ which means modes whose wavelength is smaller than the Hubble horizon and are referred to as sub-Hubble. In inflation, another way to characterize the sub-Hubble regime is by asking $k|\tau| \gg 1$ (as opposed to supper-Hubble where $k|\tau| \ll 1$). This simultaneously corresponds to sufficiently small scales, and a far enough past (in most early universe models τ starts from $-\infty$) so the comoving Hubble Horizon is large enough; together ensuring the mode is sub-Hubble.

We then set the initial conditions in the sub-Hubble regime, where the spacetime resembles Minkowski spacetime ($k^2 \gg |z''/z|$ means the modes do not feel the expansion) and the behavior of quantum fields is well understood. The solution to the equation of motion in this regime is the standard harmonic oscillator solution, i.e.

$$\nu_k(\tau) \overset{k|\tau|\gg 1}{=} C_1 e^{-ik\tau} + C_2 e^{ik\tau} \,. \tag{2.46}$$

Requiring zero particle production in far past[7] for the quantum field sets $C_2 = 0$ and applying the Wronskian condition gives $C_1 = 1/\sqrt{2k}$. Therefore, the proper initial condition is formulated as

$$\nu_k(\tau) \overset{k|\tau|\gg 1}{=} \frac{e^{-ik\tau}}{\sqrt{2k}} \,. \tag{2.47}$$

This is known as the *adiabatic vacuum*[8] or *Bunch-Davies vacuum* [25]. Note that we are working in the Heisenberg picture, where states remain fixed while operators evolve. Thus, once the initial state is set, it remains valid until the end of inflation.

2.2.3 Non-Gaussianities

Similar to the power spectrum quantifying the two-point function of curvature perturbations, $\langle \hat{\zeta}_{\boldsymbol{k}_1} \hat{\zeta}_{\boldsymbol{k}_2} \rangle$, we can define the *bispectrum*, $B_\zeta(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3)$, as a measure of the three-point function, given by

$$(2\pi)^3 \delta^3(\boldsymbol{k}_1 + \boldsymbol{k}_2 + \boldsymbol{k}_3) B_\zeta(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3) \equiv \langle \hat{\zeta}_{\boldsymbol{k}_1} \hat{\zeta}_{\boldsymbol{k}_2} \hat{\zeta}_{\boldsymbol{k}_3} \rangle \,. \tag{2.48}$$

[7] Note a time-dependent frequency for a quantum harmonic oscillator naturally leads to the creation of particles. One can construct the number count operator as $\hat{N}_{\boldsymbol{k}} = \hat{a}^+_{\boldsymbol{k}_2} \hat{a}^-_{\boldsymbol{k}_1}$ and minimize the particle production. See [22] for more details.

[8] To be more precise, the choice of vacuum is not uniquely defined. In principle, any set of mode functions corresponds to quantum operators that have their own ground state, which can be considered a vacuum state. For example, instead of using the adiabatic vacuum to minimize particle production, one could choose to minimize the Hamiltonian (the minimum energy level), which could lead to a different vacuum state. Check [24] for more discussions.

δ^3 appears due to translational invariance and it indicates that three wave vectors should form a triangle. Each $\zeta_{\boldsymbol{k}}$, and δ^3 has dimensions of k^{-3}, therefore, the bispectrum has dimensions of k^{-6}. Hence the dimensionless bispectrum is defined as

$$\tilde{B}_\zeta(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3) \equiv \frac{(k_1 k_2 k_3)^3}{k_1^3 + k_2^3 + k_3^3} B_\zeta(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3)\,, \tag{2.49}$$

and the $f_{\rm NL}$ parameter as

$$f_{\rm NL}(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3) \equiv \frac{B_\zeta(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3)}{2\left(P_\zeta(k_1)P_\zeta(k_2) + P_\zeta(k_2)P_\zeta(k_3) + P_\zeta(k_1)P_\zeta(k_3)\right)}\,, \tag{2.50}$$

which is a dimensionless quantity, and NL stands for Non-Linear. Note $\tilde{B}_\zeta$, and $f_{\rm NL}$ are both dimensionless quantities, however, the latter also quantifies the ratio of the three-point function to the two-point function. As we will explain shortly, the three-point function is generated by the interaction terms in the Lagrangian. Therefore, the three-point function also quantifies the level of deviations from Gaussianity (or Non-Gaussianities) in the statistics of any perturbative dynamical DoF, here specifically in ζ. In this sense, assuming that non-Gaussianities are generated locally in $\zeta(\boldsymbol{x})$, one can expand the real-space value of ζ as follows[9]

$$\zeta(\boldsymbol{x}) = \zeta_{\rm G}(\boldsymbol{x}) + f_{\rm NL}^{\rm local}\left(\zeta_{\rm G}(\boldsymbol{x})^2 - \langle\zeta_{\rm G}(\boldsymbol{x})^2\rangle\right)\,, \tag{2.51}$$

where $\zeta_{\rm G}(\boldsymbol{x})$ represents the Gaussian part, and $f_{\rm NL}^{\rm local}$ quantifies deviations from Gaussianity. As shown in Appendix A (where $f_{\rm NL}^{\rm local}$ is replaced by $f_{\rm NL}^{\zeta}$), the above expansion for $\zeta(\boldsymbol{x})$ gives the following for the bispectrum,

$$B_\zeta(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3) = 2 f_{\rm NL}^{\rm local}\left(P_\zeta(k_1)P_\zeta(k_2) + P_\zeta(k_1)P_\zeta(k_3) + P_\zeta(k_2)P_\zeta(k_3)\right)\,, \tag{2.52}$$

which is the same as Eq (2.50). In other words, the original definition for B_ζ, and $f_{\rm NL}$ is motivated by the local expression for B_ζ here.

The local description is not the only form of non-Gaussianities that could potentially appear in ζ. Hence, it is common to define the following which is known as the dimensionless shape function[10]

$$\mathcal{F}(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3) \equiv C k_1^2 k_2^2 k_3^2 B_\zeta(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3)\,, \tag{2.53}$$

[9] If one defines the $f_{\rm NL}^{\rm local}$ in terms of a three-point function for $\delta\phi$, a coefficient of $\frac{3}{5}$ appears in front of $f_{\rm NL}^{\rm local}$ after a gauge transformation to ζ. However, for simplicity, here we ignore this coefficient.

[10] Not to be confused with the dimensionful quantity, also known as the shape function and defined as

$$\mathcal{A}(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3) \equiv C k_1^3 k_2^3 k_3^3 B_\zeta(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3)\,.$$

where C is a normalization which has to be specified and in principle could be k-dependent. Comparing this with Eq (2.50), we notice $\mathcal{F}$ has the same nature as the $f_{\rm NL}$ function, therefore, it equivalently represents the level of non-Gaussianities.

As mentioned before, $\boldsymbol{k}_1$, $\boldsymbol{k}_2$, and $\boldsymbol{k}_3$ form a triangle. Rotational invariance indicates that the bispectrum should not depend on the orientation of this triangle, hence only its overall shape matters. The overall shape of the triangle can be parametrized in terms of two of its sides (for example k_2/k_1, and k_3/k_1) and its overall size (k_1). In fact, the goal of introducing the normalization factor, C, is to split the bispectrum into two parts, (I) a shape function that quantifies how the bispectrum depends on the shape of the triangle, (II) an overall amplitude quantifying how the bispectrum depends on the overall size of the triangle. However, in general, this splitting is ambiguous, and one must remain consistent when comparing observations with models. We discuss this in more detail in Section 3.3.4. If the bispectrum is scale-invariant, i.e. $B_\zeta(\lambda\boldsymbol{k}_1, \lambda\boldsymbol{k}_2, \lambda\boldsymbol{k}_3) = \lambda^{-6} B_\zeta(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3)$, it is generally more straightforward to perform the normalization [26].

Any model for the early universe can have its own shape function. Therefore, one can plot the full shape function to report the level of non-Gaussianities in a model. An alternative way is to define a few standard template shapes as follows

$$\mathcal{F}(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3) \equiv f_{\rm NL}^{\triangle} \mathcal{F}^{\triangle}_{\rm template}(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3)\,, \tag{2.54}$$

where $\triangle$ is just a schematic symbol representing different types of shapes, i.e. $\triangle \in$ {local, squeezed, equilateral, orthogonal, folded}. Most of the early universe models generate non-Gaussianities in the form of one (or more) of these template shapes. Then, the amplitude of the non-Gaussianities of a particular shape is encoded in the corresponding $f_{\rm NL}$ parameter. Here, we do not bring the expressions for the template shapes since they are relatively lengthy expressions. The reader can refer to [26–29] for the exact definition, motivation, and implications of each shape function. The template shape functions are constructed so that each one is maximized in the following limits

$$\begin{aligned}
k_1 \ll k_2 \simeq k_3 \qquad & \text{(squeezed shape)} \\
k_1 = k_2 = k_3 \qquad & \text{(equilateral shape)} \\
k_1 = \sqrt{2}k_2 = \sqrt{2}k_3 \qquad & \text{(orthogonal shape)} \\
k_1 = 2k_2 = 2k_3 \qquad & \text{(folded shape)}
\end{aligned} \tag{2.55}$$

justifying their names as these limits create a triangle with the same shape.

So far, we have shown how the three-point function can be represented in some parameters. Now, the question is what generates a three-point function, or non-Gaussianities. The answer can be derived using the in-in formalism for quantum fields, which calculates the

expectation value of an operator in a time-evolving background[11]. The expectation value to the leading order is given by (see [30–32] for a detailed derivation)

$$\langle \hat{O} \rangle = 2\mathrm{Im}\langle 0|\hat{O} \int_{\tau_\mathrm{i}}^{\tau} d\tilde{\tau}\ \hat{\mathrm{H}}_\mathrm{int}|0\rangle\,, \tag{2.56}$$

where τ_i is when the vacuum state is set ($-\infty$ for the Bunch-Davis vacuum), and $\hat{\mathrm{H}}_\mathrm{int}$ is the interaction Hamiltonian given by

$$\hat{\mathrm{H}}_\mathrm{int} \simeq -L^{(3)}\,, \tag{2.57}$$

where

$$\int d\tau L^{(3)} \equiv S^{(3)}\,, \tag{2.58}$$

and $S^{(3)}$ is the third-order action in perturbations which can be calculated by performing similar steps in calculating $S^{(2)}$ (see the discussion around Eq (2.24)). In Chapter 3, we study a particular model and show the relevant calculations. Lastly, note that, in principle, non-Gaussianities can arise not only from $\langle\zeta\zeta\zeta\rangle$ but also from $\langle hhh\rangle$, $\langle h\zeta\zeta\rangle$, and $\langle hh\zeta\rangle$, where h schematically represents tensor perturbations. However, as we show in Section 2.3.1, even the signal for the two-point function of tensor perturbations has not been detected yet. Therefore, the signal for three-point functions including tensor perturbations is not observable with current observations.

Throughout this subsection, we showed the cosmological perturbation theory and derived some observable quantities. In the next subsection, we show observational constraints on these observables and build some early universe models which respect these constraints.

2.3 Early universe models

Inflation, historically speaking, has been suggested as an early universe phase characterized by accelerated expansion, intended to address the horizon and flatness issues (which will be elaborated on soon). Subsequently, it was discovered that the combination of inflationary theory and quantum mechanics provides an explanation for the origin of cosmological perturbation which later seeded the formation of stars and galaxies. After the first inflation model was proposed, many more models were suggested. The paper [33] (also see [34]), titled *Inflationary Encyclopedia*, lists over 100 different models[12] for single-field inflation with minimal

[11] This is in contrast to the in-out formalism, which gives the scattering amplitudes between vacuum states at past and future infinity.

[12] The first version of the paper was published in 2014, before the release of the Planck satellite results, and thus included models that had not yet been ruled out. The authors updated the article in 2024, incorporating the most recent data.

kinetic terms and smooth potential. One can relax some of these assumptions and propose even more complicated models. For instance, see [35] for a review of multi-field inflation, and [36] for inflationary models in the context of modified gravity. As we will show shortly, accelerated expansion is not the only possibility; a contracting universe is also viable which further enlarges the realm of early universe models. Hence, in what follows, we explain the general ideas which guide the model construction for the early universe. We summarize the current observational constraints, and briefly introduce the horizon and flatness problems. We present the mathematical structure of inflationary models that satisfy these constraints and address the horizon and flatness problems. Finally, we explore alternative early universe scenarios, with a particular focus on bouncing cosmology.

2.3.1 Observational constraints

The two-point statistics of the temperature fluctuations shown in Figure 2.1 can be represented in an angular power spectrum. Since the physics from the onset of the radiation-dominated universe to the CMB phase is fairly known, one can reconstruct the primordial power spectrum for curvature perturbations from the CMB angular power spectrum. This has been done by the Planck collaboration [37], and their result for the dimensionless power spectrum is shown in Figure 2.2. The key feature of this power spectrum is that, in the well-constrained region, it varies only slightly with the scale k, in other words, the dimensionless power spectrum is scale-invariant. Hence, the following parametrization is suggested

$$\mathcal{P}_\zeta(k) = A_s \left(\frac{k}{k_{\text{pivot}}}\right)^{n_s - 1} , \tag{2.59}$$

where k_{pivot} is an arbitrary pivot scale, A_s is the amplitude of the power spectrum, and n_s is known as the *spectral index*. The former is also known as the red tilt since the power spectrum is tilted toward lower wavenumbers. One can assume n_s itself is also k dependent and define the following parameter known as *running of the spectral index*,

$$\alpha_s \equiv \frac{dn_s}{d\ln(k)} . \tag{2.60}$$

However, the current observational constraints are consistent with $\alpha_s = 0$. The current best constraints on A_s, and n_s are [11]

$$\ln\left(10^{10} A_s\right) = 3.043 \pm 0.014 \qquad n_s = 0.9652 \pm 0.0042 \qquad (68\ \%\ \text{CL}) , \tag{2.61}$$

where they set $k_{\text{pivot}} = 0.002\ \text{Mpc}^{-1}$.

In addition to CMB temperature fluctuations, its polarization map can also be observed. Several independent telescopes, apart from Planck, are specifically designed to observe the

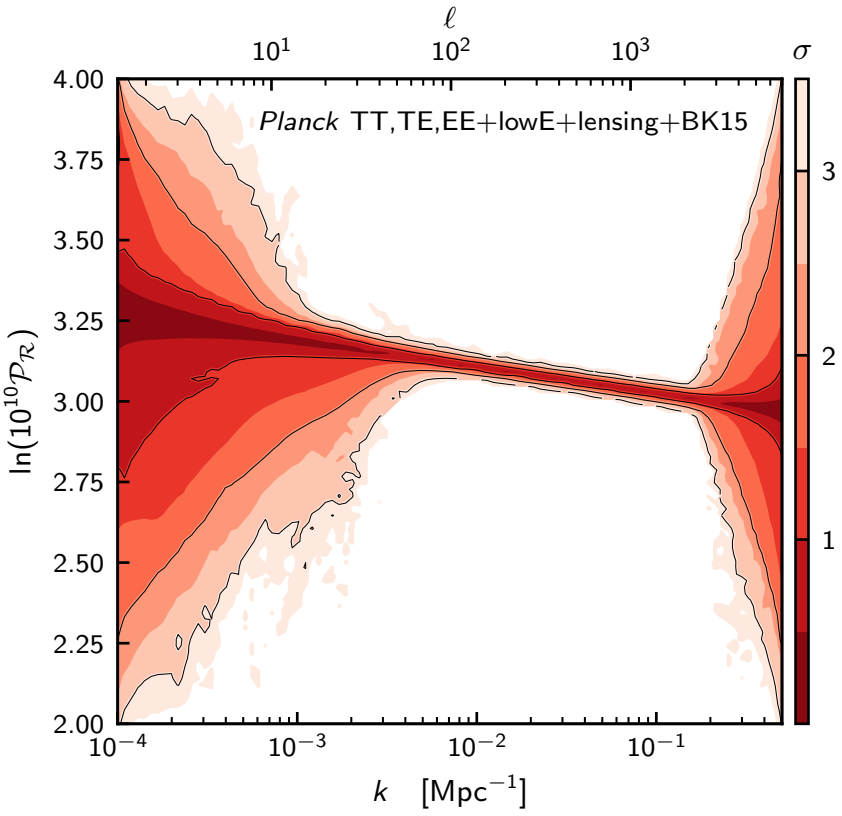


Figure 2.2: Reconstructed primordial power spectrum for curvature perturbations. The Coloured bounds show different confidence levels. The figure is adapted from [37], note they use a different letter, $\mathcal{R}$ (instead of ζ), for curvature perturbations.

CMB polarization map. The most significant among them is the BICEP/Keck observation (BK) [38]. Scalar perturbations can generate both E-mode and B-mode polarization, while tensor perturbations can only generate B-modes. Hence, detecting the B-mode polarization is direct evidence for primordial tensor perturbations. However, B-mode polarization is also produced by astrophysical sources, creating foreground contamination to the signal, so the detection has been unsuccessful so far. Therefore, primordial tensor perturbations are only constrained by upper bounds, represented in terms of r as [37]

$$r_{0.002} < 0.056 \qquad (95\ \%\ \mathrm{CL})\,, \tag{2.62}$$

where the number 0.002 represents the value of the pivot scale, i.e. $k_{\mathrm{pivot}} = 0.002\ \mathrm{Mpc}^{-1}$. There are a variety of future CMB experiments, some of them are already under construction or collecting data and some of them are planned to launch in the near future. In general, there are three categories of these experiments: (I) ground-based experiments including AdvACT, Polarbear, CLASS, Keck/BICEP3, Simons Array, and SPT3G (II) balloons including EBEX 10k, ExploreR (Piper), and Spider (III) satellites including CMBPol, COrE, PIXIE, PRISM, and LiteBIRD. Ground-based and balloon experiments mostly have fewer frequency channels and their goal is to improve the sensitivity over some specific range of multipoles. On the other hand the goal of space-borne experiments is mostly spanning a larger multipoles range. Also the frequency channels are more than the ground-based ones in order to better control systematic errors. [39–43].

Lastly, the three-point statistics of the CMB temperature map can be calculated and connected to the non-Gaussianity parameters introduced in Section 2.2.3. The first measurements of the non-Gaussianity signal was done using the WMAP satellite telescope [29, 44, 45], and the current best constraints from Planck collaboration are [27]

$$f_{\mathrm{NL}}^{\mathrm{local}} = -0.9 \pm 5.1 \qquad f_{\mathrm{NL}}^{\mathrm{equil}} = -26 \pm 47 \qquad f_{\mathrm{NL}}^{\mathrm{ortho}} = -38 \pm 24\,, \tag{2.63}$$

where all the error bars are 68% confidence level.

2.3.2 Horizon and flatness problems

As we mentioned in Section 2.1 (see Figure 2.1), the universe possessed nearly identical conditions at all the points on the surface where the CMB radiation was emitted. How is this possible? This implies that all the points on the CMB patch must have been in causal contact at some earlier phase before the emission of the CMB. If the universe has a beginning, t_{initial}, one can calculate the maximum comoving distance between points with causal contact. This, known as the *comoving horizon*; d_{comoving}, is given by the distance that

a light particle can travel. Setting $ds^2 = 0$ in Eq (2.1) gives

$$d_{\text{comoving}}(t) = \int_{t_{\text{initial}}}^{t} \frac{d\tilde{t}}{a(\tilde{t})} = \int_{a_{\text{initial}}}^{a} \frac{d\ln(a)}{aH(a)} = \int_{a_{\text{initial}}}^{a} \frac{d\ln(a)}{\mathcal{H}(a)}\,, \tag{2.64}$$

where $a_{\text{initial}} = a(t = t_{\text{initial}})$, and $\mathcal{H}^{-1}$ is the known as the *comoving Hubble horizon*. Eq (2.7) gives $\mathcal{H}(a)^{-1} \propto a^{(1+3w)/2}$. If we calculate d_{comoving} for the regular matter ($w = 0$ for dust, or $w = 1/3$ for radiation) from t_{initial} to the time of CMB, it would be much smaller than the size of the whole observable CMB patch. This is known as the Horizon problem.

The obvious solution to this problem is to increase d_{comoving}. For an always expanding universe, $\mathcal{H}^{-1}$ is increasing if $w > -1/3$, therefore, its contribution in the result of the integral becomes smaller as we move closer to t_{initial}. Hence, in order to increase d_{comoving}, we need a phase in which $\mathcal{H}^{-1}$ is decreasing i.e. the comoving Hubble horizon is shrinking. Generally, there are two ways to achieve this; (I) We simply ask $w < -1/3$ which corresponds to an accelerated phase of expansion (inflation) (II) We keep $w > -1/3$, but we ask $\mathcal{H}$ itself to be negative which correspond to a contracting phase. Both options will be discussed later.

Next, in order to discuss the flatness problem, let us first modify the Friedmann equation based on the metric given in Eq (2.12) as follows,

$$3M_{\text{Pl}}^2(H^2 + \frac{k}{a^2}) = \bar{\rho}\,. \tag{2.65}$$

We then define the critical energy density, ρ_{crit}, as the energy density that gives the exact flat space, i.e. $k = 0$,

$$\rho_{\text{crit}} \equiv 3M_{\text{Pl}}^2 H^2\,. \tag{2.66}$$

By defining $\Omega \equiv \bar{\rho}/\rho_{\text{crit}}$, and $\Omega_k \equiv -k/(aH)^2 = -\mathcal{H}^{-2}k$ (both known as density parameters), we can once more rewrite the Friedmann equation as the following simple form,

$$\Omega_k + \Omega = 1\,. \tag{2.67}$$

We mentioned below Eq (2.12) that observations indicate the contribution of spatial curvature, k, to be negligible, however, they do not constrain k directly, instead the present-day value of Ω is constrained to be close to 1 [11], indicating the present-day value Ω_k is close to 0, $\Omega_K = 0.0007 \pm 0.0037$ (95 % CL) [37].

As the definition of Ω_k indicates, its evolution is governed by the behavior of the comoving Hubble horizon, $\mathcal{H}^{-1}$. Consequently, if $\mathcal{H}^{-1}$ increases over time, the universe must have started with an extremely fine-tuned Ω_k close to 0, so that after being amplified through the evolution of $\mathcal{H}^{-1}$, it remains consistent with its current value. This is known as the flatness problem, and similar to the horizon problem, it can be resolved by proposing a phase in which the comoving Hubble horizon shrinks.

2.3.3 Inflationary models

From a phenomenological point of view, a successful theory of inflation as argued before must meet the following list of observational constraints

- As a solution to the horizon and flatness problems, there should be a phase in which the comoving Hubble horizon shrinks. Inflation, by construction, meets this requirement. In addition, the inflationary phase must last long enough to generate a sufficiently large causal patch, at least as large as the CMB surface.
- A scale-invariant power spectrum for scalar perturbations should be generated. The observational constraints on the shape of the power spectrum is characterized by A_s (the amplitude), n_s (spectral index), and α_s (running). In principle, it is possible to consider higher derivatives of the power spectrum or special features on small scales, however, currently, there are no observational constraints on them.
- No significant tensor perturbations should be generated meeting the current upper bound on r.
- Generation of non-Gaussianities should be under control and meeting current bounds on $f_{\rm NL}$.

Any model that satisfies all of these requirements could be a viable inflationary model. Here, we focus only on models driven by a single scalar field, which, in the context of inflation, is referred to as the *inflaton*.

Let us first introduce the basic toy model known as *de Sitter inflation*. For the model with a canonical scalar field, if we set $(\bar{\phi})' = 0$, then $\bar{\phi}$ is a constant and consequently, its potential is just a constant i.e. $V(\bar{\phi}) \equiv \Lambda_{\rm I}$[13]. As we mentioned below Eq (2.10), this translates to $w = -1$ ($\epsilon = 0$), and $a = \exp(Ht)$, where $H = \sqrt{\Lambda_{\rm I}/(3M_{\rm Pl}^2)}$ is just a constant. Therefore, for the comoving Hubble horizon, we have

$$\mathcal{H}^{-1} = H^{-1}e^{-Ht}\,, \tag{2.68}$$

which is exponentially decaying. This, as it was expected from a phase with $w < -1/3$, resolves the Horizon and flatness problems[14]. However, the main issue with the exact de

[13]The value of this constant is different from the late time dark energy parameter, hence we use $\Lambda_{\rm I}$ instead of just Λ.

[14]It is also possible to discuss this model at the level of perturbation. However, note that in this case $\epsilon = 0$, and the action for ζ in Eq (2.24) evaluates to zero. This is because setting both $\delta\phi$ and $(\bar{\phi})'$ to zero makes the gauge transformation in Eq (2.19) ill-defined. Therefore, one has to calculate everything in another gauge.

Sitter solution is that it never ends, leading to eternal exponential expansion. Thus, it serves only as a toy model.

The simplest modification to the de Sitter model is to require $(\bar{\phi})'$ to be close to zero but not exactly zero. This allows for the possibility of designing a potential that allows small values of $(\bar{\phi})'$ to sustain a sufficiently long period of inflation, followed by a transition to a phase with larger $(\bar{\phi})'$ which ends inflation. In other words, w does not have to be exactly -1, and the possible range becomes $-1 < w < -1/3$ which using Eq (2.8), translates to $0 < \epsilon < 1$. Therefore, instead of an exponentially accelerated expansion, we parameterize the expansion using a power-law as follows,

$$a(t) \propto |t|^p \,, \tag{2.69}$$

where p is a real positive constant. This can be rewritten in terms of conformal time,

$$a(\tau) \propto |\tau|^{p/(1-p)} \,. \tag{2.70}$$

The equation of state parameter then evaluates to $\epsilon = 1/p$, therefore $0 < \epsilon < 1$ is equivalent to $\infty > p > 1$. Note, this parametrization of the scale factor does not allow for a time-varying equation of state. Recalling Eq (2.26), and Eq (2.41), and considering the fact that ϵ is constant, we have

$$\frac{z''}{z} = \frac{z''_{\rm t}}{z_{\rm t}} = \frac{a''}{a} = \frac{v^2 - 1/4}{\tau^2} \,, \tag{2.71}$$

where v is defined as

$$v^2 \equiv \frac{(1-3p)^2}{4(1-p)^2} \,. \tag{2.72}$$

Next, the general solution to the EoM for $\nu_k(\tau)$ (or $u_k(\tau)$) is given by

$$\nu_k(\tau) = \sqrt{\tau}\left[C_1 H^{(1)}_{|v|}(k|\tau|) + C_2 H^{(2)}_{|v|}(k|\tau|)\right] \tag{2.73}$$

where $H^{(1)}_{|v|}$, and $H^{(2)}_{|v|}$ are respectively Hankel functions of type 1, and type 2. Upon imposing the initial condition given in Eq (2.47), we have $C_2 = 0$, and $C_1 = \sqrt{\pi}/2$ up to an irrelevant phase.

The sub-Hubble solution has a strongly blue spectrum. In contrast, as we show below, the super-Hubble solution could have a nearly scale-invariant power spectrum. Therefore, in cosmology, we are interested in the modes that start from the sub-Hubble regime, then cross the horizon, and become super-Hubble. The energy scale of the background then needs to

be adjusted such that the modes corresponding to CMB scales cross the horizon. We then should take $k\tau \to 0$ limit of the previous solution, which including Eq (2.26) gives[15]

$$\mathcal{P}_\zeta(k,\tau) = \frac{4^{|v|-2}\Gamma(|v|)^2}{\pi^3 \epsilon M_{\rm Pl}^2} \frac{k^{3-2|v|}}{a(\tau)^2 \tau^{2|v|-1}}\,, \tag{2.74}$$

and we can read off the spectral index as

$$n_s - 1 = 3 - 2|v| = 3 - \left|\frac{1-3p}{1-p}\right| \overset{p\gg 1}{\simeq} 0\,, \tag{2.75}$$

where the last approximation ($p \gg 1$ translates to $\epsilon \to 0$) indicates that in order to generate a nearly scale-invariant power spectrum, the background evolution should be close to that of de Sitter space. In this limit, we have $|\tau| \simeq 1/(aH)$, and we can read off the amplitude of the power spectrum as

$$A_s = \frac{1}{8\pi^2}\frac{H^2}{M_{\rm Pl}^2 \epsilon}\,, \tag{2.76}$$

and by a similar process, we have

$$r = 16\epsilon\,. \tag{2.77}$$

For this model, it can also be shown that $f_{\rm NL}^{\rm local} \propto \epsilon$ (see [46] for a detailed derivation). Hence, the smallness of ϵ can generate observables that match current constraints. In addition, as mentioned earlier, inflation must last long enough, meaning that ϵ should also vary slowly. This kinematic condition can be expressed as follows

$$\eta \equiv \frac{\dot\epsilon}{H\epsilon} \ll 1\,. \tag{2.78}$$

The conditions of $\epsilon \ll 1$, and $\eta \ll 1$ can be translated into conditions on $\bar\phi$, and its potential through the background EoMs. First, as we already pointed out, $\epsilon \ll 1$ requires $\dot{\bar\phi}/2 \ll V(\phi)$. In order to maintain this condition for a long time, there should be a balance between the friction force, $3H\dot{\bar\phi}$, and the deriving force, $V'(\bar\phi)$, in Eq (2.11) i.e. $3H\dot{\bar\phi}+V'(\bar\phi) \simeq 0$. In other words, the acceleration term, $\ddot{\bar\phi}$, should be small compared to the other two terms. We can then summarize these conditions as follows

$$\frac{1}{2}\dot{\bar\phi} \ll V(\bar\phi)\,, \qquad |\ddot{\bar\phi}| \ll 3H|\dot{\bar\phi}|\,. \tag{2.79}$$

Using the approximated background equations, these can be shown to be equivalent to

$$\epsilon_V \equiv \frac{M_{\rm Pl}^2}{2}\left(\frac{V'(\bar\phi)}{V(\bar\phi)}\right)^2 \ll 1\,, \qquad |\eta_V| = M_{\rm Pl}^2\left|\frac{V''(\bar\phi)}{V(\bar\phi)}\right| \ll 1\,, \tag{2.80}$$

[15] We are ignoring the decaying mode. For more details on this derivation, see [18].

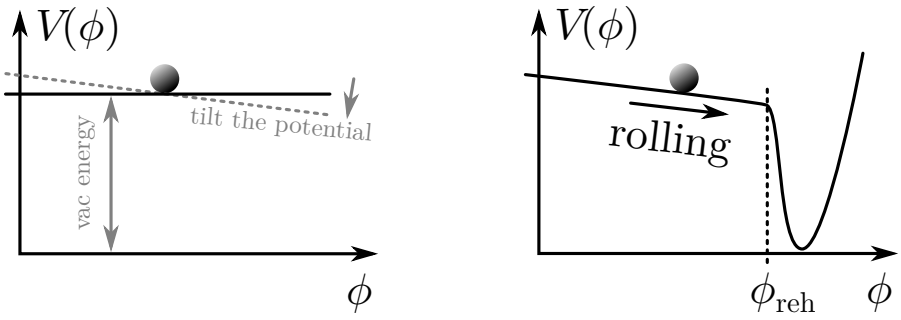


Figure 2.3: The left panel shows the potential shape in the case of the de Sitter universe which is essentially a constant, and the value of ϕ is fixed. By tilting the potential and attaching it to a reheating phase, the potential in the right panel is obtained which can generate a slow-roll inflation followed by a reheating phase. The figure is adapted from [32].

where ϵ_v, and η_V are known as *slow-roll* parameters, these conditions are known as slow-roll conditions, and a model satisfying these conditions is called slow-roll inflation. If the slow-roll conditions are satisfied, one can show the following holds

$$\epsilon \simeq \epsilon_V \,, \qquad \eta = 4\epsilon_V - 2\eta_V \,. \tag{2.81}$$

Hence, ϵ and η are also commonly referred to as slow-roll parameters. The reason for this terminology is that, as shown in Figure 2.3, the slow-roll conditions generally require an almost flat region in the potential, causing the field to roll down slowly. The inflation then ends (indicated in the plot by $\phi_{\rm reh}$) where the slow-roll conditions do not hold anymore. The subsequent phase is called the *reheating phase*, as the universe transfers from inflation to the radiation-dominated epoch.

An example of such a potential could be the simple quadratic potential,

$$V(\phi) = \frac{1}{2} m^2 \phi^2 \,. \tag{2.82}$$

Then the slow-roll parameters are

$$\epsilon_V = \eta_V = 2\left(\frac{M_{\rm Pl}}{\phi}\right)^2 \,. \tag{2.83}$$

It is obvious that to satisfy the slow-roll conditions, the field has to acquire super-Planckian values i.e. $\phi > M_{\rm Pl}$.

The joint n_s-r constraint is shown in Figure 2.4, alongside the prediction of some simple inflationary models. Hence the n_s-r constraint is a powerful tool in model selection for the early universe. One can work out the potentials suggested in Figure 2.4 and derive the corresponding conditions on the parameters of the potential. As mentioned earlier, there are

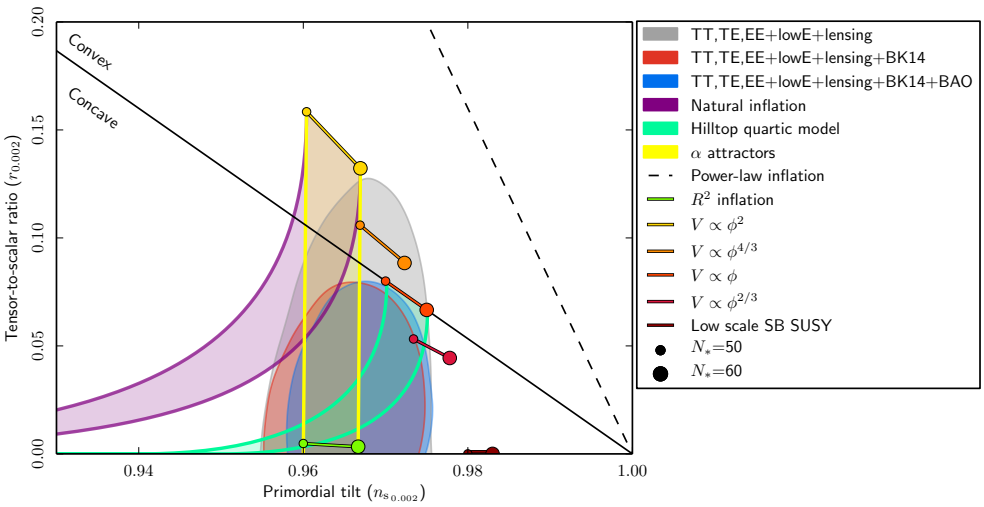


Figure 2.4: The joint n_s-r constraint form Planck collaboration [37], alongside the prediction of some inflationary models. The number 0.002 next to n_s, and r represents the value of the pivot scale, i.e. $k_{\text{pivot}} = 0.002$ Mpc^{-1}.

many viable inflationary potentials. In addition, one can propose early universe models by introducing more than one field, or by modifying gravity. Beyond observational constraints, theoretical considerations can also guide model selection, such as avoiding trans-Planckian scales and singularities. Motivated by this, we introduce a whole new class of early universe models; bouncing cosmologies.

2.3.4 Alternatives; bouncing cosmology

Contracting phase As previously mentioned, an alternative resolution to the horizon and flatness problems is to introduce a contracting phase (negative $\dot{a}$ and hence negative H) with $w > -1/3$. We still can use the parametrization given in Eq (2.69) for the scale factor, where in the contracting case, t is negative and its range is $[-\infty, 0]$. Therefore, Eq (2.75) is still valid. However, the $p \gg 1$ limit does not work anymore since it yields $w < -1/3$. Instead, the particular value of $p = 2/3$ gives the exact scale invariance for the power spectrum. This model is known as *matter bounce* since $p = 2/3$ means $w = 0$, and that resembles the equation of state of a dust-like matter. In terms of a scalar field deriving the background evolution, $w = 0$ is obtained by assuming $\dot{\bar{\phi}}^2/2 \sim V(\phi)$, recall Eq (2.10). If we consider the example potential given in Eq (2.82), we then have $\dot{\bar{\phi}}^2 \sim m^2\bar{\phi}^2$. Based on Eq (2.11), if $m \gg |H|$, such a background evolution can occur.

Although the matter bounce model can generate a scale-invariant power spectrum, it comes with several shortcomings. First of all, as Figure 2.4 shows, the exact case of $n_s - 1 = 0$

is ruled out by the observations, hence w should be slightly negative. In addition, as Eq (2.77) shows and noting $\epsilon = 3/2$ in this case, r would be larger than the current observational upper bound. Furthermore, this model has a serious theoretical problem. It can be shown that anisotropies in the universe evolve with a^{-6}, equivalent to $\epsilon = 3$, or $w = 1$. In an expanding universe, as $a(t)$ grows, anisotropies dilute faster than the inflaton energy density ($w \to 0$), ensuring that the homogeneous and isotropic background solution remains valid. In contrast, in a contracting universe, where $a(t)$ is decreasing, both anisotropies and $\rho(\bar{\phi})$ grow. If the growth rate of $\rho(\bar{\phi})$ is slower than that of anisotropies, the homogeneous and isotropic background is spoiled. This will happen inevitably for any $w < 1$ unless the initial value of anisotropies is fine-tuned to be small enough.

To get away from the enhancement of anisotropies, the equations of state parameter should obey $w \geq 1$. The case of $w = 1$ (known as stiff matter) marginally address the former issue as both the background filed and anisotropies grow with a^{-6}. This can be achieved by a massless field i.e. setting $V(\bar{\phi}) = 0$ in Eq (2.10). To be safer, it is possible to set $w > 1$ (known as ultra-stiff matter), and this is achievable by the following potential

$$V(\bar{\phi}) = -V_0 e^{-\sqrt{\frac{2}{p}}\bar{\phi}}\,, \tag{2.84}$$

where $V_0 > 0$, and $0 < p < 1/3$. Solving the background EoMs gives the same scale factor evolution of (2.69), hence $w > 1$. This model is known as *Ekpyrotic Cosmology*, and the form of the potential is motivated by string theory [47]. The problem with Ekpyrotic Cosmology is that it does not generate a scale-invariant power spectrum. To resolve this issue, it is suggested to add a second scalar field, known as *entropy field*, which is independently responsible for generating scale-invariant perturbations. This, in principle, could be a viable early universe scenario, meeting all the observational constraints. We study a similar model in extensive detail in Chapter 3.

Bounce phase So far, we have discussed how a contracting universe can be generated and examined which scenarios align with observational constraints. A non-trivial missing piece is a mechanism that smoothly transitions the contracting universe into the subsequent expanding phase. This piece is known as the *bounce* phase. To examine what is the main requirement for a bouncing cosmology, let us take a step back and move backward in time from the current universe. The late universe, ignoring the very recent accelerated epoch, is decelerating ($\ddot{a} < 0$). Considering the rate of change of the Hubble parameter $H = \dot{a}/a$, this simply implies $\dot{H} < 0$, where

$$\dot{H} = \frac{\ddot{a}}{a} - \frac{\dot{a}^2}{a^2}\,. \tag{2.85}$$

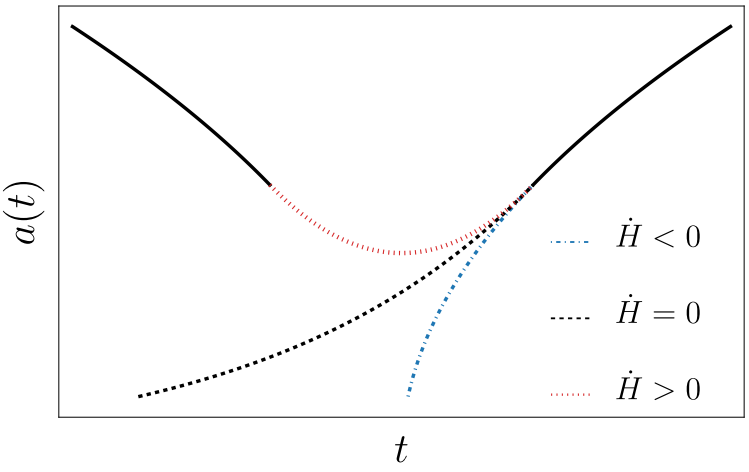


Figure 2.5: Schematic behavior of the scale factor in three different cases. The solid black curve on the right shows the decelerating late-time universe. Its mirror on the left represents a decelerating contracting phase. The other curves show different options for the continuation of the late time universe toward the past.

Roughly speaking and only considering the kinematic behavior of the scale factor, we can imagine three different cases for the continuation of the late-time universe toward the past. Schematically, this is shown in Figure 2.5. The first option (The blue dotdashed line in Figure 2.5) is to continue the $\ddot{a} < 0$ (which implies $\dot{H} < 0$) behavior and we eventually hit the $a = 0$ line. This suffers from both a Big Bang singularity and horizon-flatness problems. The second option is to glue that universe to an inflationary phase ($\ddot{a} > 0$), but making sure $\dot{H}$ stays still negative, i.e $\ddot{a}/a < \dot{a}^2/a^2$. The extreme case would be $\dot{H} = 0$, corresponding to de Sitter universe (The black dashed line in Figure 2.5). Finally, the last option is to have a bounce phase in which $\dot{a}$ goes from positive to negative, hence there is a point (bounce point) where $\dot{a} = 0$. Therefore, looking at Eq (2.85), there is a neighborhood around the bounce point in which $\dot{H} > 0$.

Attaining a phase with positive $\dot{H}$ is challenging. Form the definition of ϵ, a positive $\dot{H}$ means $\epsilon < 0$, and hence $w < -1$. Using the Friedmann equations, if the Einstein gravity holds, this translates to the condition $\bar{\rho} + \bar{p} < 0$. For a perfect fluid, this condition violates the null energy condition. This normally leads to instabilities or a superluminal speed of sound [48, 49]. To avoid violation of the NEC, one may break the relation between $\dot{H} > 0$ and $\bar{\rho} + \bar{p} < 0$ through some modified gravity theories [50, 51]. In Chapter 3, we propose a model in which a smooth bounce phase is allowed while the null energy condition is satisfied for the matter sector.

2.4 Singularity problem

The mathematical framework of general relativity is fundamentally built on the concept of smooth differential manifolds. The requirement of smoothness is essential for the axioms and definitions of differential geometry to remain valid within general relativity. However, the presence of a singularity can violate the smoothness. The eternal-like behavior of inflationary models (quasi-de Sitter) might suggest the idea that the singularity problem can be solved by inflation, however, it has been shown that inflationary models generically remain geodesically past incomplete [52, 53]. Defining a singularity is inherently challenging. A common intuition is that a singularity corresponds to an ill-defined point in a coordinate system, such as a time coordinate being truncated at a specific value. In this sense, inflation might appear to provide a resolution; however, since coordinates in general relativity have no invariant meaning, attempting to resolve the past singularity with an eternal time coordinate may be misleading.

2.4.1 Geodesic Incompleteness

One might think of the definition of singularity as a problematic point in space-time. There are several caveats to this definition. Tracking points in a coordinate-independent way is usually difficult. Also, if we want to talk about individual points, we need the complete solution for the geometry of the universe. We can have these solutions for very symmetric cases, like the FLRW universe in cosmology or some black hole solutions. For those, there are indeed some problematic points where a representation of curvature or energy density blows up. However, in reality, the universe is deviated from the symmetric state. Another problem is that the word '*point*' makes sense only when we have a well-behaved space-time solution at that point. In other words, when we already know that a '*singularity*' is where something is ill-defined, the notion of '*place*' will not work anymore. Therefore, we must abandon the notion of a singularity as a place or point in space-time [54]. Attempting to define a singularity as a curvature divergence also fails, because it may artificially arise due to actually approaching places at infinity. In addition, one can come up with examples of having a singularity without any of its direct illustration in the curvature tensor or curvature scalars [54].

Giving up defining singularity as a place, we may search for pathological behavior through what a hypothetical observer may experience. For this, we can describe the world-lines of observers as some geodesics, either timelike or null. Then the inability to boundlessly travel through these geodesics (either toward the future or past) is called *geodesic incompleteness*. It means the existence of the hypothetical observer comes to an end, which is interpreted

as an indication of a singularity. To mathematically describe this concept, we can start from a geodesic congruence which roughly is a set of geodesics. Here we show it for timelike geodesics and similar calculations exist for null geodesics. We parameterize a one-parameter sub-family of the congruence by $\gamma_s(r)$ where $s \in \mathbb{R}$. We then can define two vector fields: (I) the tangent vector field, $U^\mu = (\partial/\partial r)^\mu$; (II) the deviation vector field $S^\mu = (\partial/\partial s)^\mu$ which shows the infinitesimal spatial displacement from one geodesic to its neighbor. Then we can define a new tensor field as

$$B_{\mu\nu} = \nabla_\nu U_\mu \,, \tag{2.86}$$

and using $\mathcal{L}_U S^\mu = 0$, where $\mathcal{L}$ is the Lie derivative, we can show the following equation holds:

$$U^\nu \nabla_\nu S^\mu = B_\nu{}^\mu S^\nu \,. \tag{2.87}$$

This shows the geometrical interpretation of $B_\mu{}^\nu$ is the failure of S^μ to be parallelly transported. In other words, $B_{\mu\nu}$ quantifies the evolution of the geodesic congruence. We are more interested in the trace of $B_{\mu\nu}$ which is denoted by θ and is called the *expansion* scalar. θ measures the average expansion of geodesics that are infinitesimally close to each other. In FLRW spacetimes, θ relates to the Hubble parameter and measures the expansion (or contraction) of geodesics of comoving observers. If we extract the spatial-trace part of the geodesic deviation equation (which is essentially the extended congruence version of the geodesic equation), we get the following relation known as Raychaudhuri's equation:

$$\frac{d\theta}{d\eta} = -\frac{1}{3}\theta^2 - \sigma_{\mu\nu}\sigma^{\mu\nu} + \omega_{\mu\nu}\omega^{\mu\nu} - R_{\mu\nu}U^\mu U^\nu \,, \tag{2.88}$$

where η is an affine parameter along the geodesic, $\sigma_{\mu\nu}$ is known as the shear tensor, $\omega_{\mu\nu}$ is the vorticity tensor, and $R_{\mu\nu}$ is the Ricci tensor. The second term on the right-hand side is non-positive, and the third term is zero if the congruence is hypersurface orthogonal. The last term by virtue of the Einstein equation can be translated to quantities related to the energy-momentum tensor, i.e., with Newton's gravitational constant and the speed of light set to unity,

$$R_{\mu\nu}U^\mu U^\nu = 8\pi(T_{\mu\nu}U^\mu U^\nu + \frac{1}{2}T) \,, \tag{2.89}$$

where T denotes the trace of the energy-momentum tensor, $T_{\mu\nu}$. We ask for the right-hand side to be positive; this condition is known as the strong energy condition. Therefore, if the strong energy condition holds for the matter and Einstein's gravity is the gravity theory here, we can confidently state

$$\frac{d\theta}{d\eta} + \frac{1}{3}\theta^2 \leq 0 \,, \tag{2.90}$$

which means

$$\frac{d}{d\eta}(\theta^{-1}) \geq \frac{1}{3}\,, \tag{2.91}$$

and finally,

$$\theta^{-1}(\eta) \geq \theta_0^{-1} + \frac{1}{3}\eta\,, \tag{2.92}$$

where θ_0 is the initial expansion(or contraction) at $\eta = 0$. Eq (2.92) is at the heart of most geodesic incompleteness theorems. It simply shows that if at a particular moment (basically the time that refers to θ_0), the expansion becomes negative, i.e., if θ_0 is negative, then as the time elapses toward the future, the left-hand side should pass through zero, which means eventually θ diverges and goes to $-\infty$. Therefore, θ diverges within a finite time, $\eta \leq \frac{3}{|\theta_0|}$. This is the inability to travel through the geodesic forever.

The derivation was brief here and we will skip the next steps as well (for more details, the reader can refer to [54] or other standard GR textbooks), but essentially, the pathological behavior of θ here, alongside some other propositions and theorems, leads to geodesic incompleteness theorems. There are different versions of geodesic incompleteness theorems (with different assumptions and consequently different conclusions), however here for reference, we quote one of them [54] (in fact, this is a vintage one, and there are more recent versions):

> "Let $(M, g_{\mu\nu})$ be a globally hyperbolic spacetime with $R_{\mu\nu}U^\mu U^\nu \geq 0$ for all timelike U^μ, which will be the case if Einstein's equation is satisfied with the strong energy condition holding for matter. Suppose there exists a smooth (or, at least C^2) spacelike Cauchy surface Σ for which the trace of the extrinsic curvature (for the past directed normal geodesic congruence) satisfies $K \leq C < 0$ everywhere, where C is a constant. Then no past directed timelike curve from Σ can have length greater than $3/|C|$. In particular, all past directed timelike geodesics are incomplete."

This one in particular refers to timelike geodesic incompleteness in the context of cosmology. If U^μ is assumed to be null instead of timelike, T does not appear in Eq (2.89). Hence $R_{\mu\nu}U^\mu U^\nu \geq 0$ is translated to $T_{\mu\nu}U^\mu U^\nu \geq 0$ for a null U^μ, known as the null energy condition (NEC). For a perfect fluid, the null energy condition translates to $\bar{\rho} + \bar{p} \geq -1$ ($w \geq -1$). Therefore, if the NEC is respected, as it is the case in inflationary models, singularities would be inevitable.

2.4.2 Non-singular cosmologies

Here, we briefly list the possible resolutions to the singularity problem.

- *Keeping the Lorentzian manifold structure*: In this case, the Lorentzian geometry is still at play, however, to avoid singularities, at least one of the assumptions of the singularity theorem must be violated. One option is to introduce a matter sector that violates the NEC. Additionally, it is possible to modify gravity, hence not respecting the Einstein equation part of the singularity theorems. One exciting option is to introduce spacetime extensions which means gluing the problematic spacetime to a larger one in which geodesics are complete. Note that doing this smoothly is non-trivial and not always possible.

- *Quantum gravity approaches*: It is believed that the classical description of space and time will break down once the universe approaches scales that are relevant to quantum phenomena. Examples are string theory, loop quantum gravity, causal set theory, etc. Since the singularity theorems are based on a classical description of spacetime, a successful theory of quantum gravity could potentially resolve these singularities.

- *Altering the Lorentzian manifold structure*: By relaxing the Lorentzian geometry assumption, the manifold is not necessarily required to be smooth in the first place. Mathematically, these are well developed; examples are piecewise smooth manifolds, curvature bounded spaces, pre-length spaces, Riemannian (Euclidean) manifolds, etc. However, their ability to describe nature is uncertain.

In bouncing cosmology, singularities could be generally avoided as there is a minimum scale for the size of the universe. As mentioned in Section 2.3.4, a bouncing cosmology could violate the NEC, hinting at a solution for the singularity problem as it does not respect the assumption of the singularity theorems. However, note that these theorems are not if-and-only-if statements: one can only investigate if the assumptions are satisfied, then the solution is geodesically incomplete. In the contrapositive version, if the geodesics are complete, it indicates one of the assumptions of the singularity theorem does not hold anymore, for example, an energy condition is violated, or the constraint on the extrinsic curvature is not respected. Proving geodesic completeness requires an alternative approach.

Geodesic completeness means their affine parameter should span $[-\infty, \infty]$. Here we examine a simple scenario of a flat FLRW universe which possess a bounce and was contracting before the bounce. In flat FLRW, the affine length of a null geodesic is given by

$$\lambda = \int_{t_\mathrm{i}} a(t) dt \,. \tag{2.93}$$

In a bouncing universe $a \to \infty$ as $t \to -\infty$, and a never reaches 0, so clearly in this case, the length diverges as $t_\mathrm{i} \to -\infty$ providing the necessary range for the affine parameter.

Chapter 3

Cuscuton Bounce Beyond the Linear Regime: Bispectrum and Strong Coupling Constraints

3.1 Introduction

Bouncing cosmology offers a compelling alternative to inflationary cosmology as the theory of the very early universe. While the inflationary paradigm provides a seemingly simple early universe scenario (resolving the puzzles of standard big bang cosmology and generating the seeds of our universe's large-scale structures; see, e.g., [55–58]), it is not free of any conceptual or theoretical issues (see, e.g., [59–64]). For instance, at least within its simplest (semi)classical realizations, it generally predicts an initial singularity in the past [65] (see [53, 66] and references therein for a detailed view of the issue). In contrast, a key aspect of bouncing cosmology is to provide a mechanism that avoids any cosmological singularity. However, this is precisely the source of one of the main challenges that bouncing cosmology faces (see, e.g., [67, 68] for reviews): how can the bounce — the transition from contraction to expansion — consistently and successfully occur?

Even considering a simple cosmological background (i.e., flat, homogeneous, and isotropic), it is already challenging to find bouncing solutions. This cannot be achieved with Einstein gravity and 'normal matter' satisfying the null energy condition (NEC). Therefore, we must seek a modification or an extension to the theory, which could account for high-curvature or nonlocal effects from a complete theory of quantum gravity (e.g., string theory [69–77], loop quantum gravity [78–84], etc.). Propositions for resolving big bang singularities into classically nonsingular cosmological bounces are abundant in the literature. However,

while many succeed at describing the background evolution, they often suffer from instabilities or other kinds of pathologies once inhomogeneities are included. In this work, we shall focus on one particular theory, which shows great promise since it has not exhibited any of these problems thus far: the Cuscuton [85, 86]. This theory proposes to add a new scalar field to general relativity (GR) which breaks the time diffeomorphism invariance precisely such that the theory does not propagate any new local degree of freedom (DoF) compared to GR (see, e.g., [85–97]).

The Cuscuton generically allows for nonsingular cosmologies that are free of instabilities at the linear level, i.e., there is no Ostrogradsky, ghost, gradient or tachyonic instability; see [97–100]. This is already a remarkable feature on its own since this is often very difficult to get from other types of modified gravity such as models of Horndeski theory that admit nonsingular bounces (for models that are claimed to work and others that simply do not work, see, e.g., [101–125]). Within Horndeski theory, one typically needs very special choices of functions (see the previous references) or introduce beyond-Horndeski terms (e.g., within degenerate higher-order scalar-tensor theories; see, e.g., [120–123, 126–135]). Even in the cases that are claimed to work in (beyond-)Horndeski theory, it is often difficult to remain under perturbative control in the high-curvature regime of the bounce [136], as corrections from higher-order terms (expanding the Lagrangian to third order and beyond) can become of the order of (or even larger than) second-order perturbations of the Lagrangian. This can be manifest, for instance, from the behavior of the sound speed, which may become very small in the bounce phase. One often finds the ratio of terms in the third-order action (computed on shell) to that of terms in the second-order action to scale as $1/c_{\rm s}^2$ (see, e.g., [137]), and control of the perturbation theory may become problematic if this ratio becomes larger than unity. This is usually a sign of strong coupling or even nonunitarity [112, 138–140]; at the very least, it may imply that large non-Gaussianities are generated through the bounce. (For related discussions in other alternative scenarios, see, e.g., [123, 139–147].) For the Cuscuton, it appears that the sound speed may remain far from very small values [99]. In fact, it might even allow for slight superluminality in the vicinity of the bounce and for high wavenumbers, without violating causality (this is not uncommon in this context; see, e.g., [148–150]).[1] Therefore, one might be able to show that the Cuscuton theory is further well behaved, in that it remains under perturbative control through a nonsingular bounce, free of strong coupling issues, with reasonable non-Gaussianities. This shall be the main goal of this chapter.

[1] It is worth noting the so-called sound speed does not always correspond the waves' propagation speed. However, a proper understanding and interpretation of $c_{\rm s} > 1$ in this context would have to be the subject of a future study on its own, as it comes with several caveats. It certainly does not appear to imply any kind of instability or pathology.

A second challenge for bouncing cosmologies, if seen as alternatives to inflation, is the generation of cosmological perturbations in the prebounce contracting phase, which can lead to predictions in agreement with cosmic microwave background (CMB) observations. Different models have been proposed such as matter domination (vanishing equation of state [EoS]) [151–153], the ekpyrotic scenario (ultra-stiff EoS) [154–156], or pre-big bang cosmology (stiff EoS) [157–159]. Purely adiabatic scenarios (such as matter domination) are known to face several issues, e.g., the overproduction of primordial gravitational waves or non-Gaussianities [160, 161], background instability with respect to anisotropies [97, 162–164], or strong coupling issues [137] (see [97, 165–168] for workarounds). Therefore, we will focus on 'curvaton-like' scenarios, in which isocurvature modes (a.k.a. entropy perturbations) are generated with a nearly scale-invariant power spectrum before being converted to curvature perturbations. The most recent ekpyrotic models make use of this mechanism [169–171], but it can also be implemented in a background evolving with a simple massless scalar field driving contraction with a stiff EoS [158, 172]. (Ekpyrosis has the advantage of being an attractor washing out anisotropies [173–180]; a massless scalar field marginally addresses this issue.) These models can be implemented such that the entropy modes remain Gaussian in the contracting phase prior to their conversion into curvature perturbations [169, 170, 181, 182]. This will allow us to focus on and isolate the impact of the nonsingular bounce phase, driven by the Cuscuton field, on the predicted bispectrum.

Outline The chapter is organized as follows: in Sec. 3.2, we introduce and review the Cuscuton bounce model, the nonsingular background evolution, and the evolution of linear perturbations. Specifically, we derive the general equations for the background and for the cosmological perturbations in Sec. 3.2.1. In Sec. 3.2.2, we present different scenarios for the generation of large-scale fluctuations in the contracting phase and derive approximate analytical solutions for the evolution of the background and of the perturbations, including through the bounce phase. Some numerical examples are shown in Sec. 3.2.3. We move on to the issues of strong coupling and non-Gaussianities in Sec. 3.3. After outlining our expectations in Sec. 3.3.1, the full third-order action is derived in Sec. 3.3.2. The cubic terms are then estimated and compared to those in the second-order action in Sec. 3.3.3 to estimate the strong coupling scale. We calculate the bispectra in various limits in Sec. 3.3.4, both analytically and numerically. This both reinforces our results regarding the validity of the perturbative regime and provides an estimate for the magnitude of the non-Gaussianities. We end with a discussion in Sec. 3.4.

Summary of the main results Since the chapter is relatively long, we summarize the main results here for the reader's convenience.

- We present two models for the contracting phase, which generate scalar perturbations that can be consistent with observations: a scalar field with an ekpyrotic potential and a massless field. For both models, we analytically and numerically show that perturbations of observational interest receive insignificant growth in the infrared (IR) as they pass through the bounce. This holds both for entropy and adiabatic perturbations, as well as if one is converted into the other right before the bounce. Adiabatic modes entering the bounce from a phase dominated by a massless scalar are the most prone to enhancement; still, the growth is at most a small fraction of the perturbations' pre-bounce amplitude. All other modes effectively remain constant to leading order in the IR.

- By comparing the terms in third- and second-order actions, we show that perturbativity is well under control in the IR, and in the UV, it appears to break down when the fluctuations' physical momenta approach the scale $\sqrt{|H_{\mathrm{b}-}|M_{\mathrm{Pl}}}$, where $|H_{\mathrm{b}-}|$ is the highest Hubble rate reached in the universe (in absolute value), right by the start of the bounce phase. This scale is controlled by the Cuscuton model parameters, and it is usually well below the Planck scale. Higher interaction terms due to the Cuscuton in the bounce phase may bring the strong coupling scale down to $(H_{\mathrm{b}-}^2 M_{\mathrm{Pl}})^{1/3}$, but in all situations the background energy scale is shown to remain below the scale of strong coupling.

- We compute all nonzero 3-point correlation functions in Fourier space and find them to be highly suppressed in the IR (the bispectra are blue). Both entropy and adiabatic perturbations thus remain Gaussian and within the observational bounds to a very good approximation, at least assuming no additional enhancement during the conversion process. We show that there exists a reasonable range of Cuscuton parameters over which this remains true as the perturbations pass through the bouncing phase. Ultimately, the details of the process for converting isocurvature perturbations into curvature perturbations is the main remaining possibility responsible for producing potentially observable non-Gaussianities.

3.2 The Cuscuton bounce in the linear regime

3.2.1 General setup, background equations, and linear perturbations

The setup shall consist of three scalar fields coupled to gravity: the Cuscuton φ with Lagrangian

$$\mathcal{L}_\varphi = -M_{\rm L}^2\sqrt{-\partial_\mu\varphi\partial^\mu\varphi} - U(\varphi)\,, \tag{3.1}$$

which is responsible for modifying the gravitational dynamics, in particular allowing for nonsingular bouncing spacetimes, though without introducing any new dynamical DoFs [85–96]; the adiabatic scalar field ψ with Lagrangian

$$\mathcal{L}_\psi = X - V(\psi)\,, \qquad X \equiv -\frac{1}{2}\partial_\mu\psi\partial^\mu\psi\,, \tag{3.2}$$

which corresponds to a dominant matter field in the theory and which is responsible for driving the background cosmological evolution; and last an entropy scalar field χ with Lagrangian

$$\mathcal{L}_\chi = -\frac{1}{2}M_{\rm Pl}^2F(\psi,X)\partial_\mu\chi\partial^\mu\chi\,, \tag{3.3}$$

which is a spectator field (i.e., unaffecting the background), responsible for the generation of late-time scalar perturbations. The total action is thus

$$S = \int \mathrm{d}^4x\,\sqrt{-g}\left(\frac{M_{\rm Pl}^2}{2}R + \mathcal{L}_\varphi + \mathcal{L}_\psi + \mathcal{L}_\chi\right)\,, \tag{3.4}$$

where g is the determinant of the metric and R is the Ricci scalar. A few comments are in order: we choose a negative sign for the Cuscuton's 'kinetic term' since this is the one that allows for nonsingular bouncing cosmology [98–100] — both signs are equally valid [85, 86] and the Cuscuton is nondynamical (there is no ghost). The Cuscuton field has dimensions of mass, and we use $M_{\rm L}$ to denote the new mass scale associated with the Cuscuton 'kinetic term' — the subscript 'L' is meant to indicate that this is a Limiting curvature scale, in the spirit of a limiting curvature theory (e.g., [100, 183–185] and references therein). It is called μ in [85, 86, 89, 98], but we shall reserve this variable for the dimensionless ratio of the Cuscuton's mass scale to the Planck scale, i.e., $\mu \equiv M_{\rm L}/M_{\rm Pl}$. We note that the scalar field ψ is simply a proxy for the dominant matter DoF, in this case taken to be a canonical scalar field with some arbitrary potential $V(\psi)$ at this point. Finally, we note that the scalar field χ is kinetically coupled to the matter field via the dimensionless function $F(\psi,X)$ (in the spirit of, e.g., [170, 172, 186]) — the possible form of this function will be specified shortly, but the idea is that it should be a function of either ψ or its derivative. It was previously

mentioned that χ acts as a spectator field; it shall soon become clear that this results in χ being relevant only perturbatively. For this reason, we define χ as a dimensionless scalar by explicit insertion of $M_{\rm Pl}^2$ in (3.3).

Let us write the metric in the Arnowitt-Deser-Misner formalism as $g_{\mu\nu}\mathrm{d}x^\mu\mathrm{d}x^\nu = -N^2\mathrm{d}t^2+h_{ij}(N^i\mathrm{d}t+\mathrm{d}x^i)(N^j\mathrm{d}t+\mathrm{d}x^j)$, where the lapse, shift, and spatial metric are perturbed about a Friedmann-Lemaître-Robertson-Walker (FLRW) background as

$$N(t,\boldsymbol{x}) = 1+\alpha(t,\boldsymbol{x})\,, \qquad N_i(t,\boldsymbol{x}) = \partial_i\beta(t,\boldsymbol{x})\,, \qquad h_{ij}(t,\boldsymbol{x}) = a(t)^2e^{2\zeta(t,\boldsymbol{x})}\delta_{ij}\,. \tag{3.5}$$

These are only the scalar perturbations, which are going to be the perturbations of interest throughout this work — we will briefly comment on tensor perturbations at the end. We further perturb the scalar fields as

$$\varphi(t,\boldsymbol{x}) = \bar\varphi(t)+\delta\varphi(t,\boldsymbol{x})\,, \qquad \psi(t,\boldsymbol{x}) = \bar\psi(t)\,, \qquad \chi(t,\boldsymbol{x}) = \bar\chi(t)+\delta\chi(t,\boldsymbol{x})\,. \tag{3.6}$$

Here we are explicitly choosing the comoving gauge with respect to the matter field ψ by setting $\delta\psi(t,\boldsymbol{x}) \equiv \psi(t,\boldsymbol{x}) - \bar\psi(t) = 0$. In what follows, we will abuse notation and simply express the matter field as $\psi(t)$, and its kinetic energy will be $X = \dot\psi^2/2$. Note that we are free to perform this choice of gauge (under some reasonable assumptions, e.g., such that close to FLRW, constant-ψ surfaces form spacelike hypersurfaces), and any other gauge would result in the same second-order action (see, however, [99] for a slight subtlety when using the spatially flat gauge). In the above, $a(t)$ is the background scale factor and $\zeta(t,\boldsymbol{x})$ is the curvature perturbation.

The background equation of motion (EoM) for χ is given by $\ddot{\bar\chi} + \big[3H + \dot F(\psi,X)\big]\dot{\bar\chi} = 0$, where $H \equiv \dot a/a$ is the Hubble parameter. We note that $\dot{\bar\chi} = 0$ is always a solution, in which case $\bar\chi$ never contributes to the background evolution — in this regime χ is only a spectator field. Therefore, going forward, we shall assume $\bar\chi = 0$ — the value of the constant is irrelevant, and this solution has been shown to be stable [170, 172, 186] for the models of interest that will be explained later. The field χ thus only enters at the perturbation level as $\delta\chi$, so we will simply write $\delta\chi(t,\boldsymbol{x}) = \chi(t,\boldsymbol{x})$ and directly use the latter variable to describe the perturbation.

Given the spectator field χ does not contribute to the backgound, we may write the remaining background equations as

$$3M_{\rm Pl}^2H^2 = \frac{1}{2}\dot\psi^2 + V(\psi) + U(\bar\varphi)\,, \tag{3.7a}$$

$$-2\dot H = M_{\rm Pl}^{-2}\dot\psi^2 - \mu^2|\dot{\bar\varphi}|\,, \tag{3.7b}$$

$$\ddot\psi + 3H\dot\psi + V_{,\psi}(\psi) = 0\,, \tag{3.7c}$$

$$3\mu^2H = M_{\rm Pl}^{-2}U_{,\varphi}(\bar\varphi)\,, \tag{3.7d}$$

where a comma denotes a partial derivative. Those are the Friedmann constraint equation, the background evolution equation, the matter EoM, and the Cuscuton constraint equation, respectively. Note that from here on, without loss of generality, we will work under the assumption that $\dot{\bar{\varphi}} > 0$ in order to avoid the use of absolute values. We are free to make this choice since the Cuscuton on a cosmological background is defined for timelike $\partial_\mu\varphi$, so the sign of $\dot{\bar{\varphi}}$ will never change.[2] Defining

$$\epsilon \equiv -\frac{\dot{H}}{H^2}\,, \qquad \epsilon_\psi \equiv \frac{\dot{\psi}^2}{2M_{\rm Pl}^2H^2}\,, \qquad \sigma \equiv \epsilon_\psi - \epsilon\,, \tag{3.8}$$

the Friedmann constraint and evolution equations above can also be rewritten as $U(\bar{\varphi}) = M_{\rm Pl}^2(3-\epsilon_\psi)H^2 - V(\psi)$ and $\mu^2\dot{\bar{\varphi}} = 2\sigma H^2$. Note that since $\dot{\bar{\varphi}} > 0$, it follows that $\sigma > 0$, and in the GR limit $\mu^2 \to 0$ implies $\sigma \to 0$ for $H \neq 0$. We also notice from the above that the Cuscuton does not follow any dynamical equation but rather only a constraint equation — it is the nature of the Cuscuton not to introduce any new DoF but to modify the gravitational dynamics. Taking a derivative of the Cuscuton constraint (3.7d), we note that

$$U_{,\varphi\varphi}(\bar{\varphi}) = \frac{3\mu^2M_{\rm Pl}^2\dot{H}}{\dot{\bar{\varphi}}}\,. \tag{3.9}$$

We can thus see from (3.7d) and (3.9) how an appropriate choice of Cuscuton potential allows for nonsingular solutions. In particular, if $U_{,\varphi}$ is bounded, so will be H; additionally, where it crosses zero corresponds to a bouncing point. If this occurs while $U_{,\varphi\varphi}$ admits an interval where it is positive, then according to (3.7d) this will also allow for a bouncing phase (a transition from contraction [$H \propto U_{,\varphi} < 0$] to expansion [$H \propto U_{,\varphi} > 0$]) where $\dot{H} > 0$, implying that the effective NEC is violated. Conversely, away from the bounce phase, the effective NEC is restored as long as $U_{,\varphi\varphi}$ becomes negative. Examples of such solutions are given in [98, 99], and we will show this explicitly in the next subsection.

We will soon consider the case where ψ is massless as an example. For such a case, it is actually straightforward to see in what sense the Cuscuton is nondynamical. Furthermore, if $V(\psi)$ is set to zero, the constraint equations allow for a simple change of variable from t to $\bar{\varphi}$ and significant simplifications of the differential equations. So let us explore this example. First, we note that once the potential $U(\varphi)$ is prescribed, the Hubble parameter is set by the constraint equation (3.7d) as a function of the $\bar{\varphi}$ profile. Then, if ψ has vanishing potential, its energy density is $\rho_\psi = \dot{\psi}^2/2$, and we can write [using (3.7a) and (3.7d)]

$$\rho_\psi = 3M_{\rm Pl}^2H^2 - U(\bar{\varphi}) = \frac{U_{,\varphi}(\bar{\varphi})^2}{3\mu^4M_{\rm Pl}^2} - U(\bar{\varphi})\,, \tag{3.10}$$

[2] In fact, it is known that the theories with $\partial_\mu\varphi$ timelike vs spacelike are fundamentally different, in particular with regard to the number of propagating DoFs [85, 88, 91]. In this work, we always consider the theory that has $\partial_\mu\varphi$ timelike.

hence once the Cuscuton potential is chosen, one immediately obtains $\rho_\psi(\bar\varphi)$. The second time derivative of the scalar field as a function of $\bar\varphi$, if ever needed, is then obtained from the ψ EoM (3.7c), $\ddot\psi = -3H\dot\psi = \mp(\mu M_{\rm Pl})^{-2}U_{,\varphi}(\bar\varphi)\sqrt{2\rho_\psi(\bar\varphi)}$, where the sign depends on whether $\dot\psi$ is positive or negative, which depends on the chosen initial conditions. Finally, one can isolate $\dot H$ (and then find ϵ) by combining (3.7b), (3.9), and (3.10):

$$\dot H(\bar\varphi) = \frac{2\rho_\psi(\bar\varphi)U_{,\varphi\varphi}(\bar\varphi)}{M_{\rm Pl}^2\big(3\mu^4M_{\rm Pl}^2 - 2U_{,\varphi\varphi}(\bar\varphi)\big)} \quad\Rightarrow\quad \epsilon(\bar\varphi) = 3\left(1-\frac{3\mu^4M_{\rm Pl}^2U(\bar\varphi)}{U_{,\varphi}(\bar\varphi)^2}\right)\left(1-\frac{3\mu^4M_{\rm Pl}^2}{2U_{,\varphi\varphi}(\bar\varphi)}\right)^{-1}. \tag{3.11}$$

From those equations, one can see that every background quantity can be expressed as a function of $\bar\varphi$. To get temporal evolution, a strategy is to find the solution for $\bar\varphi(t)$, after which every function of $\bar\varphi$ found above becomes a function of time, e.g., $\dot H(t) = \dot H(\bar\varphi(t))$. The solution for $\bar\varphi(t)$ can be found from (3.9) and (3.11):

$$\dot{\bar\varphi} = \frac{3\mu^2M_{\rm Pl}^2\dot H(\bar\varphi)}{U_{,\varphi\varphi}(\bar\varphi)} = \frac{2}{\mu^2M_{\rm Pl}^2}\frac{U_{,\varphi}(\bar\varphi)^2 - 3\mu^4M_{\rm Pl}^2U(\bar\varphi)}{3\mu^4M_{\rm Pl}^2 - 2U_{,\varphi\varphi}(\bar\varphi)}\,. \tag{3.12}$$

Given a potential $U(\varphi)$, this is a first-order ordinary differential equation that can be solved given some initial condition. The initial condition is just an arbitrary choice of reference time; often we shall set $\bar\varphi(t=0)=0$ to be the bounce point as in the numerical examples of Sec. 3.2.3.

Let us move on to the general description of the linear perturbations. (What follows reproduces calculations performed in [89, 97–99, 172]; the reader familiar with this work may skip ahead to Sec. 3.2.2.) Upon perturbing the action to second order in α, β, $\delta\varphi$, ζ, and χ [defined in (3.5)–(3.6)], one finds

$$\begin{aligned} S^{(2)} = M_{\rm Pl}^2\int {\rm d}^3\boldsymbol{x}\,{\rm d}t\,a^3\Bigg(&H^2(\epsilon_\psi-3)\alpha^2 - 3\mu^2H\alpha\,\delta\varphi + 6H\alpha\dot\zeta + 3\mu^2\dot\zeta\,\delta\varphi - 3\dot\zeta^2 + \frac{(\partial_i\zeta)^2}{a^2} \\ &- 2\alpha\frac{\partial^2\zeta}{a^2} - \Big(2H\alpha - 2\dot\zeta + \mu^2\delta\varphi\Big)\frac{\partial^2\beta}{a^2} + \frac{\mu^4}{4\sigma}\Big(3\epsilon\delta\varphi^2 + \frac{(\partial_i\delta\varphi)^2}{a^2H^2}\Big) \\ &+ \frac{1}{2}F(\psi,X)\Big(\dot\chi^2 - \frac{(\partial_i\chi)^2}{a^2}\Big)\Bigg), \end{aligned} \tag{3.13}$$

where we used integration by parts and the background EoMs to simplify the expression. Note that $\partial^2 \equiv \delta^{ij}\partial_i\partial_j$ denotes the spatial Laplacian. We notice that the perturbations of the lapse and shift (α and $\partial_i\beta$) are nondynamical, hence they just represent constraints (the Hamiltonian and momentum constraints), and they can be removed from the action. Variation with respect to α and $\partial_i\beta$, respectively, yields the constraints

$$2H^2(\epsilon_\psi-3)\alpha + 6H\dot\zeta - \frac{2}{a^2}\left(\partial^2\zeta + H\partial^2\beta\right) = 3\mu^2H\,\delta\varphi\,, \qquad 2(H\alpha-\dot\zeta) = -\mu^2\,\delta\varphi\,, \tag{3.14}$$

which can be solved for α and β as

$$\alpha = \frac{1}{H}\left(\dot{\zeta} - \frac{\mu^2}{2}\delta\varphi\right), \qquad \beta = -\frac{\zeta}{H} + a^2\epsilon_\psi Q\,, \tag{3.15}$$

where Q is defined according to

$$\partial^2 Q \equiv \dot{\zeta} - \frac{\mu^2}{2}\delta\varphi\,. \tag{3.16}$$

It is understood that $\partial^{-2}(\partial^2 Q) = Q$, so ∂^{-2} is the inverse spatial Laplace operator. Similar to the lapse and shift, $\delta\varphi$ also appears nondynamically in the action (as expected since the Cuscuton is nondynamical). Upon substituting the solutions for α and β in (3.13), $\delta\varphi$ is found to satisfy a constraint equation, whose solution is

$$\delta\varphi = -\frac{2\sigma H}{\mu^2}\mathcal{D}^{-2}\left(\epsilon_\psi H\dot{\zeta} - \frac{\partial^2\zeta}{a^2}\right), \qquad \mathcal{D}^2 \equiv \frac{\partial^2}{a^2} - (\sigma+3)\epsilon_\psi H^2\,; \tag{3.17}$$

here $\mathcal{D}^{-2}$ is a transformed inverse Laplace operator.

Note that the solutions (3.15) may appear invalid at a bounce point when $H = 0$, but it turns out that no divergences actually arise [99]. Indeed, one can appreciate this by expanding first the solution to the Cuscuton constraint (3.17) about $H = 0$, which yields

$$\frac{\mu^2}{2}\delta\varphi \overset{H\approx 0}{\simeq} \dot{\zeta} - \frac{1}{\epsilon_\psi H}\frac{\partial^2\zeta}{a^2} - \frac{3}{\sigma}\dot{\zeta} + \frac{1}{\sigma\epsilon_\psi H^2}\frac{\partial^2\dot{\zeta}}{a^2} + \mathcal{O}(H^3)\,, \tag{3.18}$$

recalling ϵ, ϵ_ψ, and σ — defined in (3.8) — are $\mathcal{O}(H^{-2})$ about $H = 0$. This means from (3.15),

$$\alpha \overset{H\approx 0}{\simeq} \frac{1}{\epsilon_\psi H^2}\frac{\partial^2\zeta}{a^2} + \frac{3}{\sigma H}\dot{\zeta} - \frac{1}{\sigma\epsilon_\psi H^3}\frac{\partial^2\dot{\zeta}}{a^2} + \mathcal{O}(H^2)\,, \tag{3.19a}$$

$$\beta \overset{H\approx 0}{\simeq} -\frac{1}{\sigma H^2}\dot{\zeta} + \frac{3\epsilon_\psi}{\sigma}a^2\partial^{-2}\dot{\zeta} + \mathcal{O}(H)\,. \tag{3.19b}$$

This will be useful later.

Next, inserting the solution (3.17) back into the second-order action, performing more integration by parts, and using the background equations, one finally finds an action that separates as $S^{(2)} = S^{(2)}_\zeta + S^{(2)}_\chi$, with

$$S^{(2)}_\zeta = M_{\rm Pl}^2\int \mathrm{d}^3\boldsymbol{x}\,\mathrm{d}t\,a^3\left(\mathcal{G}_1\dot{\zeta}^2 + \mathcal{G}_2\frac{(\partial_i\dot{\zeta})^2}{a^2} - \mathcal{F}_1\frac{(\partial_i\zeta)^2}{a^2} - \mathcal{F}_2\frac{(\partial_i\partial_j\zeta)^2}{a^4} - \mathcal{F}_3\frac{(\partial_i\partial_j\partial_k\zeta)^2}{a^6}\right) \tag{3.20}$$

and

$$S^{(2)}_\chi = \frac{M_{\rm Pl}^2}{2}\int \mathrm{d}^3\boldsymbol{x}\,\mathrm{d}t\,a^3F(\psi,X)\left(\dot{\chi}^2 - \frac{1}{a^2}(\partial_i\chi)^2\right). \tag{3.21}$$

In the above, the coefficients are $\mathcal{G}_1 = \epsilon_\psi \mathcal{D}^{-4}\left(a^{-4}\partial^4 - 3\epsilon_\psi H^2\left(a^{-2}\partial^2 + \mathcal{D}^2\right)\right)$, $\mathcal{G}_2 = \sigma\epsilon_\psi^2 H^2 \mathcal{D}^{-4}$,

$$\begin{aligned}
\mathcal{F}_1 = \epsilon\epsilon_\psi H^2\mathcal{D}^{-6}\Bigg(&\frac{1}{\epsilon_\psi H^2}\frac{\partial^6}{a^6} + \left[5\epsilon + 2\frac{\epsilon_\psi}{\epsilon}(3+2\eta_\psi) - (15 + 5\epsilon_\psi + 2\eta + 2\eta_\psi)\right]\frac{\partial^4}{a^4} \\
&- \left[4\epsilon - 5\epsilon_\psi - 3(4+\eta) + \frac{\epsilon_\psi}{\epsilon}(3 + \epsilon_\psi + 3\eta_\psi)\right]\left[3\epsilon_\psi H^2\frac{\partial^2}{a^2} - (\sigma+3)\epsilon_\psi^2 H^4\right]\Bigg)\,, \\
\mathcal{F}_2 = \mathcal{D}^{-6}\Bigg(&-2\sigma\frac{\partial^4}{a^4} + \epsilon_\psi H^2\Big[4\epsilon^2 - \epsilon(12 + 7\epsilon_\psi + \eta + \eta_\psi) + \epsilon_\psi\big(3\epsilon_\psi + 2(6+\eta_\psi)\big)\Big]\frac{\partial^2}{a^2} \\
&- \epsilon_\psi^2 H^4\Big[2\epsilon^3 + \epsilon_\psi\big(6 + \epsilon_\psi^2 + \epsilon_\psi(5 - 2\eta_\psi)\big) - \epsilon^2(4 + 3\epsilon_\psi + \eta + \eta_\psi) \\
&+ \epsilon\big(3(\eta_\psi - \eta - 2) + \epsilon_\psi(-1 + \eta + 3\eta_\psi)\big)\Big]\Bigg)\,,
\end{aligned} \tag{3.22}$$

and $\mathcal{F}_3 = -\sigma\mathcal{D}^{-4}$, where we further defined

$$\eta \equiv \frac{\dot\epsilon}{H\epsilon}\,, \qquad \eta_\psi \equiv \frac{\dot\epsilon_\psi}{H\epsilon_\psi}\,. \tag{3.23}$$

The above actions are somewhat easier to read when expressed in Fourier space[3] and in conformal time (defined as $\mathrm{d}\tau \equiv a^{-1}\mathrm{d}t$):

$$S_\zeta^{(2)} = \frac{M_{\rm Pl}^2}{2}\int_{\boldsymbol{k}}\int \mathrm{d}\tau\, z^2\left(|\zeta_{\boldsymbol{k}}'|^2 - c_{\rm s}^2 k^2|\zeta_{\boldsymbol{k}}|^2\right)\,; \qquad S_\chi^{(2)} = \frac{M_{\rm Pl}^2}{2}\int_{\boldsymbol{k}}\int \mathrm{d}\tau\, z_\chi^2\left(|\chi_{\boldsymbol{k}}'|^2 - k^2|\chi_{\boldsymbol{k}}|^2\right)\,. \tag{3.24}$$

Note that a prime denotes a derivative with respect to τ, and defining $\mathcal{H} \equiv a'/a = aH$, the coefficients in the actions are now

$$z^2 = 2a^2\epsilon_\psi\frac{k^2 + 3\epsilon_\psi\mathcal{H}^2}{k^2 + (\sigma+3)\epsilon_\psi\mathcal{H}^2}\,, \qquad c_{\rm s}^2 = \frac{k^4 + \mathcal{H}^2\mathcal{B}_1 k^2 + \mathcal{H}^4\mathcal{B}_2}{k^4 + \mathcal{H}^2\mathcal{A}_1 k^2 + \mathcal{H}^4\mathcal{A}_2}\,, \qquad z_\chi^2 = a^2 F(\psi, X)\,, \tag{3.25}$$

with

$$\begin{aligned}
\mathcal{A}_1 &= \epsilon_\psi(\sigma+6)\,, & \mathcal{B}_1 &= \epsilon_\psi(6 + 3\sigma + \eta - \eta_\psi) - \sigma(6 + 2\sigma + \eta + \eta_\psi)\,, \\
\mathcal{A}_2 &= 3\epsilon_\psi^2(\sigma+3)\,, & \mathcal{B}_2 &= \epsilon_\psi\big(3\epsilon_\psi(3 + \sigma + \eta - \eta_\psi) - 4\sigma(3+\sigma) - 3\sigma\eta\big)\,.
\end{aligned} \tag{3.26}$$

The resulting EoMs for the mode functions ζ_k and χ_k are

$$\zeta_k'' + 2\frac{z'}{z}\zeta_k' + c_{\rm s}^2 k^2\zeta_k = 0\,, \qquad \chi_k'' + 2\frac{z_\chi'}{z_\chi}\chi_k' + k^2\chi_k = 0\,, \tag{3.27}$$

[3] Our convention is

$$\zeta(t,\boldsymbol{x}) = \int_{\boldsymbol{k}}\zeta_{\boldsymbol{k}}(t)\exp(i\boldsymbol{k}\cdot\boldsymbol{x})\,, \qquad \int_{\boldsymbol{k}} \equiv \int\frac{\mathrm{d}^3\boldsymbol{k}}{(2\pi)^3}\,,$$

and we denote $k \equiv \|\boldsymbol{k}\| \equiv \sqrt{\boldsymbol{k}\cdot\boldsymbol{k}} \equiv \sqrt{\delta_{j\ell}k^j k^\ell}$ (the standard Euclidean norm and dot product).

whose specific solutions depend on the models of interest. Note that z^2 and $c_{\rm s}^2$ are both time and wavenumber dependent. It is straightforward to see that z^2 is always positive [89], so there is no ghost, and $c_{\rm s}^2$ can also remain positive (and close to unity) given a reasonable Cuscuton potential (see [98, 99]; this will be discussed later as well).

For future reference, it is useful to look at the action for ζ in different regimes of interest: in the IR to leading order in the gradient expansion, we find

$$S_\zeta^{(2,\rm IR)} \simeq M_{\rm Pl}^2 \int \mathrm{d}^3\boldsymbol{x}\,\mathrm{d}t\,a^3 \frac{\epsilon_\psi}{1+\sigma/3}\left(\dot\zeta^2 - \frac{3\epsilon_\psi(3+\sigma+\eta-\eta_\psi)-4\sigma(3+\sigma)-3\sigma\eta}{3\epsilon_\psi(\sigma+3)}\frac{(\partial_i\zeta)^2}{a^2}\right); \tag{3.28}$$

oppositely, in the UV where spatial gradients dominate, we have

$$S_\zeta^{(2,\rm UV)} \simeq M_{\rm Pl}^2 \int \mathrm{d}^3\boldsymbol{x}\,\mathrm{d}t\,a^3\epsilon_\psi\left(\dot\zeta^2 - \frac{(\partial_i\zeta)^2}{a^2}\right). \tag{3.29}$$

This makes it explicit in what sense the Cuscuton affects IR perturbations, while leaving the UV sector of the theory untouched. The story in the UV is a little more complicated at the bounce point where $H=0$, though (see [99]). Indeed, since ϵ_ψ is $\mathcal{O}(H^{-2})$ about $H=0$, it would appear that (3.29) diverges at the point where $H=0$. However, this is solely an issue about how to take the UV limit since first expanding about $H\approx 0$ and then taking a UV limit yields a finite expression,

$$S_\zeta^{(2,\rm UV)} \overset{H\approx 0}{\simeq} M_{\rm Pl}^2 \int \mathrm{d}^3\boldsymbol{x}\,\mathrm{d}t\,\frac{a^3}{\sigma H^2}\left(\frac{(\partial_i\dot\zeta)^2}{a^2} - \frac{\sigma}{\epsilon_\psi}\frac{(\partial^2\zeta)^2}{a^4}\right), \tag{3.30}$$

since σH^2 and σ/ϵ_ψ are $\mathcal{O}(H^0)$. Note that everywhere else ($H\neq 0$), (3.29) is a perfectly good approximation, and this subtlety does not affect the evolution of the perturbations in the UV [99].

3.2.2 Explicit models and analytical approximations

Prebounce contraction

We shall consider two main models in what follows for the prebounce contracting phase: the case where ψ is an ekpyrotic field; and the case where ψ is a massless field. In the former situation, this means the background scalar field has a steep and negative potential of the form

$$V(\psi) = -V_0 e^{-\sqrt{2\bar\epsilon}\psi/M_{\rm Pl}} \tag{3.31}$$

for some real constants $V_0>0$ and $\bar\epsilon>3$, while in the latter situation this simply means $V(\psi)$ vanishes identically. The motivation for these particular models is as follows: in

[98], a massless scalar field was considered in the matter sector as the simplest proxy for matter — the focus was on the physics of the bouncing phase and on the stability of the solution. Then, in order to explore the evolution of cosmological perturbations through the bounce, especially scale-invariant perturbations that could account for the observations, an appropriate coupling to the spectator field χ was considered in [172]. In the spirit of testing the Cuscuton bounce, we will revisit this model here again. The other model where one has an early ekpyrotic phase has the advantage of being a proper attractor [173, 175], thus properly smoothing the universe (just like inflation) before the bounce [176, 180, 187–194]. Even when ψ is massless, anisotropies are under control as long as the evolution starts close enough to FLRW. Let us note that in concrete realizations of ekpyrotic cosmology, it is often the case that the ekpyrotic phase of contraction is followed by a phase of kination, i.e., the potential goes to 0, so the scalar field becomes massless and kinetic driven.[4] Therefore, the two models we will consider could actually be considered as being two separate phases of the whole prebounce contraction.

Ekpyrotic potential Let us review these models in more detail. In the far past, well before the bounce phase, we will ensure the Cuscuton field has a negligible contribution, hence we neglect it in the discussion for now. If the scalar field ψ has the ekpyrotic potential (3.31), then this will lead, early in the contracting phase, to the well-known scaling solution (see, e.g., [156, 195])

$$\begin{aligned} a(\tau) &= \left(\frac{\tau}{\tau_{\mathrm{b}-}}\right)^{\frac{1}{\bar{\epsilon}-1}}, \\ \psi(\tau) &= M_{\mathrm{Pl}}\left(\frac{\sqrt{2\bar{\epsilon}}}{\bar{\epsilon}-1}\ln\left(-\tau M_{\mathrm{Pl}}\right)+\frac{1}{\sqrt{2\bar{\epsilon}}}\ln\left(\frac{(\bar{\epsilon}-1)^2}{\bar{\epsilon}-3}\frac{V_0}{M_{\mathrm{Pl}}^4}\frac{1}{(-\tau_{\mathrm{b}-}M_{\mathrm{Pl}})^{\frac{2}{\bar{\epsilon}-1}}}\right)\right), \end{aligned} \tag{3.32}$$

for $\tau < \tau_{\mathrm{b}-} < 0$. Note that, for practical reasons, we choose to normalize the scale factor to unity at the time $\tau_{\mathrm{b}-}$, where the subscript 'b−' indicates the end of the standard contracting phase (equivalently the onset of the bounce phase where the Hubble parameter is maximal in absolute value), but physical quantities do not depend on this choice of scale factor normalization. From the above, we note

$$\mathcal{H}(\tau) = \frac{1}{(\bar{\epsilon}-1)\tau}, \qquad \epsilon = \epsilon_\psi = \bar{\epsilon}, \tag{3.33}$$

and the energy density in the ekpyrotic field scales as $\rho_\psi \propto a^{-2\bar{\epsilon}}$, hence ψ dominates over other fields as $a \to 0^+$, including massless fields and shear anisotropies whose energy densities

[4] This could be modeled with a potential of the form $V(\psi) = -V_0/(e^{\sqrt{2\bar{\epsilon}}\psi/M_{\mathrm{Pl}}} + e^{-b_V\sqrt{2\bar{\epsilon}}\psi/M_{\mathrm{Pl}}})$, for some constant $b_V > 0$, as in [104, 105]. At the level of the discussion for this chapter, we will treat a single-exponential phase and a massless one distinctly.

scale as a^{-6} (recall $\bar{\epsilon} > 3$). In fact, an ekpyrotic field often has $\bar{\epsilon} \gg 3$, hence its isotropizing power.

Let us then discuss the evolution of adiabatic perturbations in a phase of ekpyrotic contraction. Defining the Mukhanov-Sasaki variable $v_k \equiv M_{\rm Pl} z \zeta_k$, the EoM for ζ_k in (3.27) may be rewritten as

$$v_k'' + \left(c_{\rm s}^2 k^2 - \frac{z''}{z} \right) v_k = 0 \,. \tag{3.34}$$

It is straightforward to see from (3.25) that when $\epsilon = \epsilon_\psi$, we have $z^2 = 2a^2\epsilon_\psi$, and here $\epsilon_\psi = \bar{\epsilon}$ is a constant, so

$$\frac{z''}{z} = \frac{a''}{a} = \frac{\nu^2 - 1/4}{\tau^2} \,, \qquad \nu = \frac{\bar{\epsilon} - 3}{2(\bar{\epsilon} - 1)} \,, \tag{3.35}$$

and furthermore $c_{\rm s}^2 = 1$. The general solution is thus

$$v_k(\tau) = \sqrt{-\tau} \left(C_1 {\rm H}_\nu^{(1)}(-k\tau) + C_2 {\rm H}_\nu^{(2)}(-k\tau) \right) \,, \tag{3.36}$$

where $C_{1,2}$ are integration constants and ${\rm H}_\nu^{(1,2)}$ are the Hankel functions of the first and second kind, respectively. Imposing an adiabatic vacuum in the early time subhorizon[5] limit,

$$v_k(\tau) \overset{-k\tau \gg 1}{\simeq} \frac{1}{\sqrt{2k}} e^{-ik\tau} \,, \tag{3.37}$$

determines the integration constants, and up to an irrelevant phase, the solution is

$$v_k(\tau) = -\frac{\sqrt{-\pi\tau}}{2} {\rm H}_\nu^{(1)}(-k\tau) \,. \tag{3.38}$$

In the late-time superhorizon limit, the solution is well approximated by

$$v_k(\tau) \overset{-k\tau \ll 1}{\simeq} i \frac{\Gamma(\nu)}{2} \sqrt{\frac{-\tau}{\pi}} \left(\frac{2}{-k\tau} \right)^\nu \quad \Rightarrow \quad \zeta_k(\tau) \simeq -i \frac{\Gamma(\nu)}{\sqrt{\pi\bar{\epsilon}} M_{\rm Pl} \tau_{\rm b-} k^{3/2}} \left(\frac{-k\tau_{\rm b-}}{2} \right)^{\frac{\bar{\epsilon}}{\bar{\epsilon}-1}} \,, \tag{3.39}$$

[5] In this work, the subhorizon limit corresponds to the regime where the Mukhanov-Sasaki variable oscillates in time. Oppositely, the superhorizon limit corresponds to the regime where the Mukhanov-Sasaki variable stops oscillating. It could still be time dependent, as there may be a growing mode, but in the present work superhorizon perturbations are found to freeze, meaning that either the constant mode dominates over a decaying mode or that the growing mode is marginal (e.g., at most logarithmic). Deep in the contraction, $1/k \sim -\tau$ gives a good approximation of the corresponding 'horizon scale', which is also proportional to the comoving Hubble radius $1/(a|H|)$ — in fact, it is very well approximated by the Hubble radius when $\epsilon \sim 3$. More precisely, by inspection of (3.34) the horizon is controlled by $\sqrt{|z''|/z}/c_{\rm s}$, and through the bounce phase, this may significantly differ from the Hubble radius.

where $\Gamma(\nu)$ is the Gamma function, and where we ignored the decaying term as $k\tau \to 0^-$. For subsequent use, it is also useful to keep the next-to-leading-order term to compute the time derivative,

$$\zeta_k'(\tau) \overset{-k\tau\ll 1}{\simeq} \left(\frac{\pi}{\Gamma\left(\frac{1}{2}-\frac{1}{\bar\epsilon-1}\right)} - i\Gamma\left(\frac{1}{2}+\frac{1}{\bar\epsilon-1}\right)\sin\left(\frac{\pi\bar\epsilon}{\bar\epsilon-1}\right)\right)\frac{(-\tau_{\rm b-})^{\frac{1}{\bar\epsilon-1}}}{\sqrt{2\pi\bar\epsilon}M_{\rm Pl}}\left(\frac{k}{2}\right)^{\frac{1}{2}-\frac{1}{\bar\epsilon-1}}(-\tau)^{-\frac{2}{\bar\epsilon-1}}\,. \tag{3.40}$$

To zeroth order in $1/\bar\epsilon \ll 1$, this tells us that ζ_k on superhorizon scales will enter the bounce phase with an amplitude and a time derivative of the order of

$$\zeta_k(\tau_{\rm b-}) \simeq \frac{i}{2M_{\rm Pl}\sqrt{\bar\epsilon k}}\,, \qquad \zeta_k'(\tau_{\rm b-}) \simeq -ik\zeta_k(\tau_{\rm b-}) \simeq \frac{1}{2M_{\rm Pl}}\sqrt{\frac{k}{\bar\epsilon}}\,, \tag{3.41}$$

respectively. In fact, $|\zeta_k'|^2 \simeq k^2|\zeta_k|^2$ holds at all times in the limit $1/\bar\epsilon \to 0^+$ (the Minkowski limit), so the kinetic and gradient terms in the second-order on-shell action for ζ contribute equally. In other words, $z''/z \to 0^-$ when $\bar\epsilon \to \infty$, so the solution for v_k is approximately that of a simple harmonic oscillator. Then from (3.39), the dimensionless power spectrum on superhorizon scales can be expressed as

$$\mathcal{P}_\zeta(k) \equiv \frac{k^3}{2\pi^2}|\zeta_k|^2 \simeq \frac{1}{2\pi^3\bar\epsilon}\Gamma\left(\frac{\bar\epsilon-3}{2(\bar\epsilon-1)}\right)^2\frac{1}{(-M_{\rm Pl}\tau_{\rm b-})^2}\left(\frac{-k\tau_{\rm b-}}{2}\right)^{\frac{2\bar\epsilon}{\bar\epsilon-1}}\,. \tag{3.42}$$

The scalar power spectrum is usually parametrized as $\mathcal{P}_s(k) = A_s(k/k_\star)^{n_s-1}$, with an amplitude A_s, pivot scale $k_\star$, and spectral index n_s; in particular, $n_s - 1 = \frac{{\rm d}\ln\mathcal{P}_s}{{\rm d}\ln k}$. We thus note that in the large-$\bar\epsilon$ limit the power spectrum (3.42) is blue with spectral index

$$n_s - 1 = \frac{2\bar\epsilon}{\bar\epsilon-1} \simeq 2, \tag{3.43}$$

while $A_s \simeq 1/(8\pi^2\bar\epsilon)$ at $k_\star = M_{\rm Pl}$. This is a very blue spectrum, which is clearly ruled out by observations [11]. It also means the adiabatic perturbations have essentially remained in their vacuum (as naively expected from the Minkowski limit $1/\bar\epsilon \to 0$). Thus, they should be unenhanced and unobservable on cosmological scales, unless something significant occurs in the bounce phase, but we will shortly see that, to a good approximation, perturbations remain constant through the bounce.

To see why they are not observable, let us assume that (3.42) describes the post bounce power spectrum and estimate the real-space 'amplitude' of the curvature perturbation, denoted $\bar\zeta$, on a relevant cosmological scale $\mathscr{R}$. One can find the variance of ζ coarse-grained over a region of comoving size $\mathscr{R}$ by smoothing[6] the power spectrum with a Window function,

$$\bar\zeta^2(\mathscr{R}) \equiv \int_0^\infty \frac{{\rm d}k}{k}\widetilde{\mathcal{W}}^2(k\mathscr{R})\mathcal{P}_\zeta(k) = \mathcal{P}_\zeta(1/\mathscr{R}) \simeq \frac{1}{8\pi^2\bar\epsilon}\left(\frac{1}{M_{\rm Pl}\mathscr{R}}\right)^2\,, \tag{3.44}$$

[6]For more details on smoothing the power spectrum, see, e.g., [196].

where the result of the integral follows by taking a delta function for the squared Fourier-space Window function,

$$\widetilde{\mathcal{W}}^2(k\mathscr{R}) = \delta(k\mathscr{R} - 1)\,, \tag{3.45}$$

and the last equality follows from (3.42) in the large-$\bar{\epsilon}$ limit. Considering the comoving cosmological scales $\mathscr{R}$ over which we observe correlations in ζ are many orders of magnitude larger than the Planck length $M_{\rm Pl}^{-1}$, we conclude that $|\bar{\zeta}(\mathscr{R})|$ is minute, as expected for 'vacuum fluctuations'. For example, if we take $\mathscr{R} = \text{Mpc}/0.05$, then we find $M_{\rm Pl}\mathscr{R} \sim 10^{58}$, hence $|\bar{\zeta}(\mathscr{R})|$ would be of the order of $10^{-59}/\sqrt{\bar{\epsilon}}$, which is further suppressed by the largeness of $\bar{\epsilon}$.

Let us move on to the discussion of the entropy perturbation χ. First, one has to specify the coupling function $F(\psi, X)$ between the ekpyrotic field ψ and the entropy field χ. As in [170, 186], we choose

$$F(\psi) = F_0 e^{-\sqrt{2\bar{\epsilon}(1+\delta)}\,\psi/M_{\rm Pl}}\,, \tag{3.46}$$

for some $F_0 > 0$ and some small parameter $0 < \delta \ll 1$. We do not assume any X dependence here. The EoM for χ_k in (3.27) can be expressed in terms of its Mukhanov-Sasaki variable $u_k \equiv M_{\rm Pl} z_\chi \chi_k$ as

$$u_k'' + \left(k^2 - \frac{z_\chi''}{z_\chi}\right) u_k = 0\,, \tag{3.47}$$

where we recall $z_\chi^2 = a^2 F$, so here we find

$$\frac{z_\chi''}{z_\chi} = \frac{\nu_s^2 - 1/4}{\tau^2}\,, \qquad \nu_s \simeq \frac{3}{2} + \frac{\bar{\epsilon}\delta}{2(\bar{\epsilon} - 1)}\,, \tag{3.48}$$

where the approximation is to leading order in small δ. The general solution for $u_k(\tau)$ will be of the same form as (3.36), with ν replaced by ν_s, and so imposing an adiabatic vacuum for u_k as in (3.37) yields the same expressions as (3.38) and (3.39), though again with ν replaced by ν_s:

$$u_k(\tau) = -\frac{\sqrt{-\pi\tau}}{2} \mathrm{H}^{(1)}_{\nu_s}(-k\tau) \overset{-k\tau \ll 1}{\simeq} i\frac{\Gamma(\nu_s)}{2}\sqrt{\frac{-\tau}{\pi}}\left(\frac{2}{-k\tau}\right)^{\nu_s} \tag{3.49a}$$

$$\Rightarrow \quad \begin{cases} \chi_k(\tau) \simeq i\Gamma(\nu_s)\left(\frac{(\bar{\epsilon}-1)^2}{\bar{\epsilon}-3}\right)^{\frac{2+\delta}{4}} \left(\frac{-\tau_{\rm b-}\sqrt{V_0}}{M_{\rm Pl}}\right)^{\delta/2} \sqrt{\frac{2V_0}{\pi F_0 M_{\rm Pl}^4 k^3}} \left(-\frac{k\tau_{\rm b-}}{2}\right)^{-\frac{\delta}{2}(1+\frac{1}{\bar{\epsilon}})} \\ \chi_k'(\tau) \simeq i\sqrt{\frac{\bar{\epsilon}V_0 k}{2F_0}}\,\frac{\tau}{M_{\rm Pl}^2} \end{cases}. \tag{3.49b}$$

Note that small-δ and large-$\bar{\epsilon}$ approximations are made to obtain the last expressions, and the next-to-leading-order term in the expansion of the Hankel function is kept to find the

time derivative. Correspondingly, the dimensionless power spectrum is

$$\begin{aligned}\mathcal{P}_\chi(k) &\equiv \frac{k^3}{2\pi^2}|\chi_k|^2 \\ &\simeq \frac{1}{\pi^3}\left(\frac{(\bar\epsilon-1)^2}{\bar\epsilon-3}\right)^{1+\frac{\delta}{2}}\Gamma\Big(\frac{3}{2}+\frac{\delta}{2}\Big(1+\frac{1}{\bar\epsilon}\Big)\Big)^2\left(\frac{-\tau_{\rm b-}\sqrt{V_0}}{M_{\rm Pl}}\right)^{\delta}\frac{V_0}{F_0M_{\rm Pl}^4}\left(-\frac{k\tau_{\rm b-}}{2}\right)^{-\delta(1+\frac{1}{\bar\epsilon})}.\end{aligned} \tag{3.50}$$

We can read the spectral index in this case to be

$$n_s - 1 \simeq -\delta \tag{3.51}$$

in the limit of large $\bar\epsilon$. By taking the small-δ limit, we can also estimate the amplitude as

$$A_s \simeq \frac{\bar\epsilon V_0}{4\pi^2F_0M_{\rm Pl}^4}\,. \tag{3.52}$$

Since δ is a small number, the spectral index (3.51) can be in accordance with the observations, which indicate $n_s \sim 0.97$ (e.g., [11]), provided these isocurvature perturbations convert to curvature perturbations nearly one-to-one (this will be rediscussed later). Similarly, the amplitude can be compatible with the current observational estimates of $A_s \sim 10^{-9}$ (e.g., [11]), given some parameter values for $\bar\epsilon$, V_0, and F_0 (e.g., $\bar\epsilon \sim 100$, $V_0 \sim 10^{-9}\,M_{\rm Pl}^4$, $F_0 \sim 1$). This also tells us that the amplitude of the real-space perturbation χ is of the order of 10^{-5}. Finally, let us note from (3.49b) and (3.52) that

$$\chi_k(\tau_{\rm b-}) \simeq i\pi\sqrt{\frac{2A_s}{k^3}}\,, \qquad \chi_k'(\tau_{\rm b-}) \simeq i\pi\sqrt{2A_sk}\,\tau_{\rm b-}\,, \tag{3.53}$$

by the time superhorizon perturbations enter the bounce phase.

Massless case Let us move on to the discussion of the second model, i.e., the one where the scalar field ψ is massless. As we said, this can either be viewed as the kination phase that often follows ekpyrotic contraction or it can be viewed as a toy model of its own. In such a case, the background scaling solution early in the contracting phase generally reads

$$a(\tau) = \left(\frac{\tau}{\tau_{\rm b-}}\right)^{1/2}, \qquad \psi(\tau) = \psi_0 \pm \sqrt{\frac{3}{2}}\,M_{\rm Pl}\ln\left(\frac{\tau}{\tau_0}\right), \qquad \tau < \tau_{\rm b-} < 0\,, \tag{3.54}$$

where $\psi_0 = \psi(\tau_0)$ is an integration constant at a time $\tau_0 < \tau_{\rm b-} < 0$, and $\psi'(\tau_0) = \pm\sqrt{3/2}\,M_{\rm Pl}/\tau_0$ is taken in accordance with the Friedmann constraint equation, which reads $3M_{\rm Pl}^2(a'/a)^2 \simeq \psi'^2/2$ when ψ dominates [which implies $\tau_0 \ll \tau_{\rm b-}$ with our choice of calibration for $a(\tau)$]. If such a phase is to follow a phase of ekpyrosis [cf. (3.32)], then we pick the plus sign in the expression for $\psi'(\tau_0)$ [so that $\psi'(\tau_0) \propto 1/\tau_0 < 0$] and ψ_0 is fixed at the

transition[7] point; otherwise, if there is no ekpyrosis, the sign of $\psi'(\tau_0)$ and the value of ψ_0 are completely arbitrary. From (3.54), we note

$$\mathcal{H} = \frac{1}{2\tau}\,, \qquad \epsilon = 3\,, \tag{3.55}$$

so the energy density of the scalar field scales as $\rho_\psi \propto a^{-6}$, and the EoS in this case is referred to as stiff. This is marginally stable against the growth of anisotropies, which also scale as a^{-6}, and if anisotropies are diluted in an earlier phase of ekpyrotic contraction, then they should not be an issue closer to the bounce either.

The calculation of the perturbations is similar to the previous one, where now $z(\tau) = \sqrt{6}\,a(\tau)$, so $z''/z = a''/a = (\nu^2 - 1/4)/\tau^2$ with $\nu = 0$, so the solution for $v_k = M_{\rm Pl} z \zeta_k$ is

$$\begin{aligned} v_k(\tau) &= -\frac{\sqrt{-\pi\tau}}{2}\mathrm{H}_0^{(1)}(-k\tau) \overset{-k\tau\ll 1}{\simeq} i\sqrt{\frac{-\tau}{\pi}}\left(\ln(2) - \gamma_{\rm E} + \frac{i\pi}{2} - \ln(-k\tau)\right) \\ \Rightarrow\quad \zeta_k(\tau) &\simeq \frac{i}{M_{\rm Pl}}\sqrt{\frac{-\tau_{\rm b-}}{6\pi}}\left(\ln(2) - \gamma_{\rm E} + \frac{i\pi}{2} - \ln(-k\tau)\right) \quad\Rightarrow\quad \zeta_k'(\tau) \simeq -\frac{i}{M_{\rm Pl}}\sqrt{\frac{-\tau_{\rm b-}}{6\pi}}\frac{1}{\tau}\,, \end{aligned} \tag{3.56}$$

where $\gamma_{\rm E}$ is the Euler-Mascheroni constant. Note that as $\tau \to 0^-$, the log term represents a (slowly) growing mode, as is evident from the time derivative. The power spectrum is

$$\begin{aligned} \mathcal{P}_\zeta(k,\tau) &\simeq \frac{-\tau_{\rm b-}k^3}{12\pi^3 M_{\rm Pl}^2}\left|\ln(2) - \gamma_{\rm E} + \frac{i\pi}{2} - \ln(-k\tau)\right|^2 \\ \Rightarrow\quad \mathcal{P}_\zeta(k_{\rm p},t) &\simeq \frac{k_{\rm p}(t)^3}{24\pi^3 M_{\rm Pl}^2|H(t)|}\left|\ln(2) - \gamma_{\rm E} + \frac{i\pi}{2} - \ln\left(\frac{k_{\rm p}(t)}{2|H(t)|}\right)\right|^2\,, \end{aligned} \tag{3.57}$$

where $k_{\rm p}(t) \equiv k/a(t)$ represents the physical wavenumber on the second line. Since k is a superhorizon mode above, the deep-IR limit implies that the log term should dominate over the constant coefficients inside the absolute values, so we may approximately write the power spectrum right at the end of the standard contracting phase as

$$\mathcal{P}_\zeta(k_{\rm p},t_{\rm b-}) \simeq \frac{k_{\rm p}(t_{\rm b-})^3}{24\pi^3 M_{\rm Pl}^2|H_{\rm b-}|}\left(\ln\left[\frac{k_{\rm p}(t_{\rm b-})}{2|H_{\rm b-}|}\right]\right)^2\,, \tag{3.58}$$

where $|H_{\rm b-}|$ represents the maximal Hubble scale that is cosmologically reached (in absolute value). Note that the spectral index of the above power spectrum runs, i.e., it is not simply a constant as before, but the log term is only a slight correction to a deeply blue $n_s - 1 = 3$ spectrum. Indeed, one finds

$$n_s - 1 = 3 + \frac{2}{\ln\left(\frac{k}{2a|H|}\right)}\,, \qquad \alpha_s \equiv \frac{\mathrm{d}n_s}{\mathrm{d}\ln k} = -\frac{2}{\left(\ln\left[\frac{k}{2a|H|}\right]\right)^2}\,. \tag{3.59}$$

[7] The goal is not to model such a transition, so we omit details. A refined treatment would consider a more elaborate potential function that asymptotes zero — recall footnote 4.

The fact that the spectrum is deeply blue implies that it is not cosmologically observable, as was the case for the ζ spectrum in ekpyrosis. Indeed, even if one takes a very small bouncing energy scale, say $|H_{\rm b-}| \sim 10^{-15}\, M_{\rm Pl} \sim$ TeV, then for $k \sim 10^{-58}\, M_{\rm Pl} \sim 0.04\,{\rm Mpc}^{-1}$, (3.58) yields $\mathcal{P}_\zeta \sim 10^{-158}$, hence in real space we expect $\bar{\zeta} \sim 10^{-79}$.

Turning to the entropy perturbation, we once again have to specify what is the coupling $F(\psi, X)$ between the background (massless) scalar ψ and the perturbation χ. If we think of the regime where ψ is massless as following ekpyrosis, then it is sensible to take the same coupling function as (3.46), except[8] with $\bar{\epsilon} = 3$:

$$F(\psi) = F_0 e^{-\sqrt{6(1+\delta)}\,\psi/M_{\rm Pl}}\,. \tag{3.60}$$

From $\psi(\tau) = \psi_0 + \sqrt{3/2}\,M_{\rm Pl}\ln(\tau/\tau_0)$, this implies $F \propto (-\tau)^{-3(1+\delta/2)}$ in the small-δ limit. Note, though, that we can just as well choose something like

$$F(X) = \left(\frac{X}{\Lambda^2 M_{\rm Pl}^2}\right)^{1+\delta/2} \tag{3.61}$$

for some new energy scale Λ, which also yields $F \propto (-\tau)^{-3(1+\delta/2)}$ and thus achieves the same power spectrum for χ_k. The latter was the approach of [172] — they essentially had a term of the form $(\partial_\mu\psi\partial^\mu\psi)(\partial_\nu\chi\partial^\nu\chi)$ in their Lagrangian leading to $F \sim (H/\Lambda)^2$ in the prebounce phase — and it is better motivated if there is no ekpyrosis, i.e., if ψ is fundamentally massless, since the coupling function (3.61) is shift symmetric. It is, however, a noncanonical higher-dimensional operator if one wishes to obtain the desired red tilt in the scalar power spectrum (i.e., when $\delta \gtrless 0$). A term of the form $e^{\lambda\psi}(\partial_\mu\chi)^2$ in the Lagrangian, for some real constant λ, might find motivation in string theory (where ψ and χ would be akin to the dilaton and the axion[9], respectively), though it is beyond the scope of this chapter to derive (3.60) from a more fundamental theory (see, e.g., [197–199] for comments about this) — we have in mind either (3.60) or (3.61) from a phenomenological point of view.

The EoM for $u_k = M_{\rm Pl} z_\chi \chi_k$ is the same as before [recall (3.47)], except this time $z_\chi''/z_\chi = (\nu_s^2 - 1/4)/\tau^2$ with

$$\nu_s \simeq \frac{3}{2} + \frac{3\delta}{4} \tag{3.62}$$

[8] Again, a proper model that transitions from ekpyrosis to kination would have a smooth potential that asymptotes 0 and correspondingly ϵ would smoothly transition from a large constant to 3. The point here is that, in principle, even if we treat $V(\psi)$ as being exactly vanishing, we can use the same couplings and solutions as in ekpyrosis and just replace $\bar{\epsilon}$ by 3.

[9] Note that this is essentially the setup of pre-big bang cosmology [157], a string cosmology scenario where the dilaton acts as a massless scalar driving an Einstein-frame contracting phase with stiff EoS and where the axion is responsible for generating a nearly scale-invariant scalar power spectrum of isocurvature perturbations [158, 159].

in the small-δ limit, though this is really the same as (3.48) with $\bar{\epsilon} = 3$. The solution for $u_k(\tau)$ is then as in (3.49a), with the expression for ν_s replaced by (3.62), so in the end, we have [taking the coupling (3.61) for concreteness, but similar expressions can be derived from (3.60) as the coupling]

$$\begin{aligned}
\chi_k(\tau) &\overset{-k\tau\ll 1}{\simeq} i\frac{2^{\frac{3+\delta}{2}}}{3^{\frac{2+\delta}{4}}\sqrt{\pi}}\Gamma\Big[\frac{3}{2}\Big(1+\frac{\delta}{2}\Big)\Big]\frac{\Lambda}{M_{\rm Pl}}(-\tau_{\rm b-}\Lambda)^{\delta/2}\left(\frac{2}{-k\tau_{\rm b-}}\right)^{3\delta/4} k^{-3/2} \overset{\delta\ll 1}{\simeq} i\sqrt{\frac{2}{3}}\,\frac{\Lambda}{k^{3/2}M_{\rm Pl}}\,,\\
\chi_k'(\tau) &\overset{-k\tau\ll 1}{\simeq} i\frac{2^{\frac{1+\delta}{2}}}{3^{\frac{2+\delta}{4}}\sqrt{\pi}}\Gamma\Big[\frac{1}{2}\Big(1+\frac{3\delta}{2}\Big)\Big]\frac{\Lambda}{M_{\rm Pl}}(-\tau_{\rm b-}\Lambda)^{\delta/2}\left(\frac{2}{-k\tau_{\rm b-}}\right)^{3\delta/4} \sqrt{k}\,\tau \overset{\delta\ll 1}{\simeq} i\sqrt{\frac{2}{3}}\,\frac{\Lambda\sqrt{k}\,\tau}{M_{\rm Pl}}\,.
\end{aligned} \tag{3.63}$$

Then, the power spectrum is

$$\mathcal{P}_\chi(k) \simeq \frac{2^{2+\delta}\Gamma\left[\frac{3}{2}\left(1+\frac{\delta}{2}\right)\right]^2}{3^{1+\delta/2}\pi^3}\left(\frac{\Lambda}{M_{\rm Pl}}\right)^2(-\tau_{\rm b-}\Lambda)^\delta\left(\frac{2}{-k\tau_{\rm b-}}\right)^{3\delta/2}\,, \tag{3.64}$$

and the spectral index reads

$$n_s - 1 \simeq -\frac{3}{2}\delta\,. \tag{3.65}$$

This is not exactly the same as (3.51) — it simply means that if the same model is used through subsequent phases of ekpyrosis and kination, the spectral index will simply change by a factor of 3/2 from one phase to the other. If the model consists purely of a massless field, then certainly (3.65) can be in agreement with $n_s \sim 0.97$ (e.g., $\delta \sim 0.02$). The amplitude of (3.64) in the small-δ limit can be estimated as

$$A_s \simeq \frac{\Lambda^2}{3\pi^2 M_{\rm Pl}^2}\,, \tag{3.66}$$

hence the amplitude is controlled by $\Lambda/M_{\rm Pl}$ (e.g., $\Lambda \approx 2.67\times 10^{-4}\,M_{\rm Pl}$ achieves $A_s \approx 2.4\times 10^{-9}$). From this and (3.63), we obtain

$$\chi_k(\tau_{\rm b-}) \simeq i\pi\sqrt{\frac{2A_s}{k^3}}\,, \qquad \chi_k'(\tau_{\rm b-}) \simeq i\pi\sqrt{2A_s k}\,\tau_{\rm b-}\,, \tag{3.67}$$

by the time the superhorizon perturbations enter the bounce phase. Unsurprisingly, these expressions are the same as (3.53).

Bounce phase, conversion, and reheating

So far, we have reviewed two models of contracting cosmology in which the matter field ψ dominates: when ψ has an ekpyrotic potential and when ψ is massless. In both cases, adiabatic perturbations (ζ) have a blue spectrum [$n_s \sim 3$ (3.43) and $n_s \sim 4$ (3.59), respectively],

with ζ_k being approximately constant on superhorizon scales [recall (3.39) and (3.56), respectively; there is a slight logarithmic growth in the latter case]. In both cases, though, an appropriate kinetic coupling to a spectator field χ generates a (nearly) scale-invariant power spectrum of entropy perturbations, where χ_k is approximately constant on superhorizon scales [recall (3.49b) and (3.63)].

Three things still need to happen at this point: the background cosmology must bounce, i.e., there must be a transition from contraction to expansion; the isocurvature perturbations must be converted into curvature perturbations, i.e., χ must source ζ, so that before the radiation era, it is the ζ power spectrum that becomes nearly scale invariant; and finally the primordial matter field ψ must also decay into standard model fields, i.e., the universe must reheat, so that radiation-dominated expansion (standard hot big bang cosmology) may begin. Our focus will be on modeling and studying the bounce carefully; the conversion of isocurvature perturbations into curvature perturbations and reheating have been the subject of other studies (e.g., [170, 181, 182, 200–210]), whose results we will simply quote when needed. However, it is still worth noting that some of the details of the conversion process have important implications for perturbations through the bounce. First, we need to distinguish two possibilities (e.g., as studied in [182]): if the conversion occurs before the bounce or after. If the latter prevails, then we need to understand how both ζ (which has a blue spectrum) and χ (which has a scale-invariant spectrum) evolve through the bounce, matching them to their prebounce properties. In the former possibility, i.e., if χ gets converted into ζ perturbations before the bounce, then we simply need to study ζ (which has acquired the prebounce properties of the χ field, so it has a scale-invariant power spectrum) through the bounce.[10]

With that in mind, we now focus on the nonsingular transition from contraction to expansion, i.e., the bounce phase, which is driven by the Cuscuton field. We shall present specific examples of the resulting cosmological dynamics using numerical computations in a subsequent subsection, but before that, it is also useful to provide some analytical approximations and some intuitive picture regarding the general dynamics of the background and of the perturbations through the bounce phase.

First, the Cuscuton bounce phase is a regime in which the geometric null convergence condition[11] (which in FLRW says that $\dot{H} \leq 0$) is violated, so $\dot{H} > 0$ for some time interval. This ensures the bounce is nonsingular. In other words, $H(t)$ never blows up; instead, it

[10] We shall always assume that the conversion process is fully efficient, in other words that negligible isocurvature perturbations are left after sourcing the curvature perturbations. This is a good approximation according to [170, 181, 182, 201, 203–206, 209, 210].

[11] The geometric null convergence condition states that the contraction of the Ricci tensor with any null vector twice should be nonnegative.

crosses 0, i.e., we have a smooth transition from contraction ($H < 0$) to expansion ($H > 0$). The bounce phase begins when $\dot{H}$ crosses zero for the first time, which we denote by the subscript 'b−', and it ends the second time $\dot{H}$ crosses zero, which we denote by the subscript 'b+'. As this phase is usually short, the scale factor is roughly constant to leading order ($a \approx 1$ in our normalization), hence conformal time and physical time are approximately equal in the bounce phase. The following is often a good approximate parameterization of a bounce phase,

$$a(\tau) \overset{\tau\approx 0}{=} 1 + \mathcal{O}(\tau^2)\,, \qquad \mathcal{H}(\tau) \overset{\tau\approx 0}{=} \Upsilon\tau + \mathcal{O}(\tau^3)\,, \qquad \mathcal{H}'(\tau) \overset{\tau\approx 0}{=} \Upsilon + \mathcal{O}(\tau^2)\,, \tag{3.68}$$

for some dimension-2 constant $\Upsilon > 0$, where we choose the origin of the time axis, $\tau = 0$, to correspond to the bounce point.[12] In fact, (3.68) has been derived in [99] as a series solution for a generic Cuscuton bounce. Thus, (3.68) applies to the interval $\tau_{\rm b-} < \tau < \tau_{\rm b+}$. When a variable is evaluated at $\tau = 0$, we denote it by a subscript 'b'. Note that the parameter Υ and the bounce duration $\tau_{\rm b+} - \tau_{\rm b-}$ control the energy scale of the bounce phase.

With such a parameterization, we can find approximate solutions to the mode functions (3.27). Focusing on deep IR modes ($k \to 0$), which is the relevant regime of observational CMB modes as they cross a nonsingular bounce, and using (3.25), the equations (which include the contributions from the Cuscuton) can be expressed as

$$\zeta_k'' + \frac{(z^2)'}{z^2}\zeta_k' \overset{k\to 0}{\simeq} 0\,, \qquad z^2 \overset{k\to 0}{\simeq} \frac{2\epsilon_\psi a^2}{1+\sigma/3}\,, \tag{3.69}$$

and

$$\chi_k'' + 2\frac{(z_\chi^2)'}{z_\chi^2}\chi_k' \overset{k\to 0}{\simeq} 0\,. \tag{3.70}$$

To be more quantitative, the regime of validity of the IR approximation for the ζ_k and χ_k EoMs above are, respectively,

$$kc_{\rm s} \ll \sqrt{\frac{|z''|}{z}}\,, \qquad k \ll \sqrt{\frac{|z_\chi''|}{z_\chi}}\,, \tag{3.71}$$

where we recall from (3.25) that $c_{\rm s}$ and z are complicated background- and k-dependent quantities, while $z_\chi = a\sqrt{F}$ is simply a power-law in τ. Therefore, the first expression of (3.71) is a somewhat involved time-dependent nonlinear inequality in k. This may be simplified, in principle, but for the sake of simplicity here, we assume that the approximation

[12] Naturally, since the series expansion is performed about the bounce where $\tau = 0$, the approximation becomes less accurate the farther away from the bounce point at $\tau = 0$. The same is true if the bounce duration (scale) is too long (high) — it holds best as long as $\Upsilon\tau^2 \lesssim \mathcal{O}(1)$. Nevertheless, it shall capture the leading-order behavior in what follows and hence remain valid as an order-of-magnitude approximation.

is respected. For a range of IR modes of cosmological relevance, the approximation has been confirmed numerically. The time-dependent solutions to (3.69) and (3.70) are

$$\zeta_k'(\tau) \overset{k\to 0}{\simeq} \zeta_k'(\tau_{\rm b-})\frac{a(\tau_{\rm b-})^2}{a(\tau)^2}\frac{1-\frac{\epsilon(\tau)}{\epsilon_\psi(\tau)}+\frac{3}{\epsilon_\psi(\tau)}}{1-\frac{\epsilon(\tau_{\rm b-})}{\epsilon_\psi(\tau_{\rm b-})}+\frac{3}{\epsilon_\psi(\tau_{\rm b-})}}\,, \qquad \chi_k'(\tau) \overset{k\to 0}{\simeq} \chi_k'(\tau_{\rm b-})\frac{a(\tau_{\rm b-})^2F(\tau_{\rm b-})}{a(\tau)^2F(\tau)}\,, \tag{3.72}$$

where the 'initial' time derivatives $\zeta_k'(\tau_{\rm b-})$ and $\chi_k'(\tau_{\rm b-})$ can be estimated by matching the solutions to the end of the prebounce contracting phase. The equations also admit the usual superhorizon constant solutions, $\zeta_k' \simeq 0$ and $\chi_k' \simeq 0$, but by knowing the approximate value of the perturbations and their first derivatives at the onset of the bounce, the above captures the correct time evolution of the perturbations through the bounce.

Assuming for now that the conversion of isocurvature perturbations into curvature perturbations occurs only after the bounce, let us solve the χ_k equation, which is controlled by the time dependence of F, itself a function of ψ. Let us assume that the matter field ψ is massless in the bounce phase in any situation; recall the earlier discussion according to which even if ψ is an ekpyrotic field, its potential should go to zero by the time of the bounce. Then, its background EoM is $\psi'' = -2\mathcal{H}\psi'$, hence from (3.68) one finds

$$\psi(\tau) \overset{\tau\approx 0}{=} \psi_{\rm b} + \psi_{\rm b}'\tau + \mathcal{O}(\tau^3)\,. \tag{3.73}$$

Recalling the Friedmann constraint (3.7a) in conformal time and substituting $V(\psi) = 0$ for the duration of the bounce, we have $3M_{\rm Pl}^2\mathcal{H}^2 = \psi'^2/2 + a^2U(\bar\varphi)$, so at the bounce point

$$\psi_{\rm b}'^2 = -2U_{\rm b}\,, \tag{3.74}$$

where the Cuscuton potential at the bounce point must be negative (a well-known fact; see [98, 99]), i.e., $U_{\rm b} \equiv U(\bar\varphi = 0) < 0$. Considering the case where the coupling function is given by (3.61) when ψ is fundamentally massless, and in the $\delta \ll 1$ limit, we find $a(\tau)^2F(\tau) \approx \psi'(\tau)^2/(2\Lambda^2M_{\rm Pl}^2)$, which near $\tau = 0$ can be estimated as $a(\tau)^2F(\tau) \approx$ constant $+ \mathcal{O}(\tau^2)$. Hence, substituting this estimate into (3.72) for the near-bounce regime implies $\chi_k'(\tau) \approx \chi_k'(\tau_{\rm b-})$ and

$$\chi_k(\tau) \approx \chi_k(\tau_{\rm b-}) + \chi_k'(\tau_{\rm b-})(\tau - \tau_{\rm b-})\,. \tag{3.75}$$

In the case where ψ is an ekpyrotic field, we may be tempted to use (3.46) or (3.60) for F, but as argued in, e.g., [181, 182], these functions should really be modified to reach unity by the time of the bounce, i.e., one assumes that the Lagrangian for χ approaches that of a massless canonical scalar (which has decoupled from ψ). In that case, we are left with $\chi_k'' + 2\mathcal{H}\chi_k' \simeq 0$ for the IR modes, and hence we recover (3.75) again, provided $a \approx$ constant through the bounce.

If we use the expected value of χ_k and χ_k' at the onset of the bounce phase [recall (3.67)] and assume that the background dynamics of the bounce phase is symmetric, so that $\tau_{\rm b+} = -\tau_{\rm b-}$, then we may infer from (3.75) the value of the perturbation after the bounce phase as

$$\chi_k(\tau_{\rm b+}) \simeq \chi_k(\tau_{\rm b-}) \left(1 - 2k^2\tau_{\rm b+}^2\right) . \tag{3.76}$$

For an IR mode, $k^2\tau_{\rm b+}^2 \ll 1$ is an extremely small correction. Therefore, it is well justified to treat χ_k as constant through the bounce phase.

Moving on to the description of ζ_k through the bounce, we see from (3.72) that we need to know the behavior of ϵ and ϵ_ψ. First, recall that prebounce, when the Cuscuton is negligible, we have $\epsilon \approx \epsilon_\psi \approx 3$. It is exactly 3 for a massless scalar, and we expect the ekpyrotic potential to go to zero by the time of the bounce, so we assume that this approximation holds by the time the bounce starts at $\tau_{\rm b-}$. We are thus left with evaluating (treating the scale factor as approximately constant again)

$$\zeta_k'(\tau) \approx \zeta_k'(\tau_{\rm b-}) \left(1 - \frac{\epsilon(\tau)}{\epsilon_\psi(\tau)} + \frac{3}{\epsilon_\psi(\tau)}\right) . \tag{3.77}$$

Gathering (3.8), (3.68), and (3.73), we note

$$\frac{1}{\epsilon_\psi} = \frac{2M_{\rm Pl}^2\mathcal{H}^2}{\psi'^2} \overset{\tau\approx 0}{=} \frac{2M_{\rm Pl}^2\Upsilon^2\tau^2}{\psi_{\rm b}'^2} + \mathcal{O}(\tau^4) , \tag{3.78}$$

while

$$\frac{\epsilon}{\epsilon_\psi} = -\frac{2M_{\rm Pl}^2\dot{H}}{\dot{\psi}^2} \overset{\tau\approx 0}{=} -\frac{2M_{\rm Pl}^2\Upsilon}{\psi_{\rm b}'^2} + \mathcal{O}(\tau^2) , \tag{3.79}$$

so to leading order in small τ we only have

$$\zeta_k'(\tau) \approx \zeta_k'(\tau_{\rm b-}) \left(1 - \frac{\epsilon(\tau)}{\epsilon_\psi(\tau)}\right) . \tag{3.80}$$

At this point, we note that by combining equations (3.7b) and (3.9), we may generally write

$$M_{\rm Pl}^2 \left(\frac{3\mu^4 M_{\rm Pl}^2}{U_{,\varphi\varphi}(\bar{\varphi})} - 2\right) \dot{H} = \dot{\psi}^2 , \tag{3.81}$$

so that near the bounce we have the relation

$$\Upsilon = \frac{\psi_{\rm b}'^2}{M_{\rm Pl}^2(3\mu^4 M_{\rm Pl}^2/m_{\rm b}^2 - 2)} = -\frac{2U_{\rm b}}{M_{\rm Pl}^2(3\mu^4 M_{\rm Pl}^2/m_{\rm b}^2 - 2)} , \tag{3.82}$$

where we defined $m_{\rm b}^2 \equiv U_{,\varphi\varphi}(\bar{\varphi} = 0)$ and used (3.74). This relation tells us the Hubble rate of change approximately throughout the bounce phase (or exactly at the bounce point) solely

as a function of the potential parameters. It is also useful since — combining (3.79), (3.81), and (3.82) — we find

$$\frac{\epsilon_\psi}{\epsilon} = 1 - \frac{3\mu^4 M_{\rm Pl}^2}{2U_{,\varphi\varphi}(\bar{\varphi})} \qquad \Rightarrow \qquad \frac{\epsilon}{\epsilon_\psi} \overset{\tau\approx 0}{\approx} \left(1 - \frac{3\mu^4 M_{\rm Pl}^2}{2m_{\rm b}^2}\right)^{-1} . \tag{3.83}$$

Following [99], we may thus define

$$\widetilde{M}_{\rm Pl}^2 \equiv M_{\rm Pl}^2 \left(1 - \frac{3\mu^4 M_{\rm Pl}^2}{2m_{\rm b}^2}\right) , \tag{3.84}$$

hence we are left with

$$\zeta_k'(\tau) \approx \zeta_k'(\tau_{\rm b-}) \left(1 - \frac{M_{\rm Pl}^2}{\widetilde{M}_{\rm Pl}^2}\right) \qquad \Rightarrow \qquad \zeta_k(\tau) \approx \zeta_k(\tau_{\rm b-}) + \zeta_k'(\tau_{\rm b-}) \left(1 - \frac{M_{\rm Pl}^2}{\widetilde{M}_{\rm Pl}^2}\right)(\tau - \tau_{\rm b-}) . \tag{3.85}$$

According to [99], if the sound speed is to remain close to unity at all times and on all scales[13], then we should set $m_{\rm b} \approx \mu^2 M_{\rm Pl}$, which would result in $\widetilde{M}_{\rm Pl}^2 \approx -M_{\rm Pl}^2/2$. As explained in [99], the negativity of the effective background gravitational coupling about the bounce point, $\widetilde{M}_{\rm Pl}^2$, can be understood as the fact that the effective NEC is violated there.[14] Finally, this would mean $1 - M_{\rm Pl}^2/\widetilde{M}_{\rm Pl}^2 \approx 3$, and thus $\zeta_k(\tau) \approx \zeta_k(\tau_{\rm b-}) + 3\zeta_k'(\tau_{\rm b-})(\tau - \tau_{\rm b-})$.

We note that, because we are doing an expansion about $\tau = 0$ and only keeping leading-order terms, the solution when extrapolated to $\tau_{\rm b\pm}$ may be quite approximate. In particular, we note that we lost continuity of $\zeta_k'(\tau)$ at $\tau_{\rm b-}$ in going from (3.77) to (3.80). While this may appear bad, it is actually very close to the reality. This is because the behavior of $\zeta_k'(\tau)$ is controlled by the ratios $3/\epsilon_\psi$ and ϵ/ϵ_ψ according to (3.77), with $3/\epsilon_\psi \approx \epsilon/\epsilon_\psi \approx 1$ before the bounce and $3/\epsilon_\psi \approx 0$ and $\epsilon/\epsilon_\psi \approx M_{\rm Pl}^2/\widetilde{M}_{\rm Pl}^2 \approx -2$ during the bounce, and it turns out that these ratios change abruptly at $\tau_{\rm b-}$ and $\tau_{\rm b+}$. In fact, in all fully numerical solutions we computed, we confirm that $\zeta_k'(\tau)$ abruptly jumps up and then back down at the beginning and end of the bounce phase, respectively, with a near constant value in between. In other words, $\zeta_k'(\tau)$ is well approximated by a step function in that interval. This agrees with our approximate solution (3.85), which suggests a discontinuous jump in $\zeta_k'(\tau)$ at $\tau_{\rm b-}$. We will show numerical examples confirming this in the next subsection (cf. Fig. 3.2).

[13] It is worth noting that in the IR limit, the term 'sound speed' is a bit misleading since the Wentzel-Kramers-Brillouin (WKB) approximation breaks down in that regime; thus, this quantity no longer reflects the waves' propagation speed.

[14] One can generally think of the Cuscuton as a modified gravity theory, which yields a set of modified Friedmann equations in FLRW. However, those equations can be recast as the standard Friedmann equations of GR, except with a modified Planck mass [86]. This is what is meant by the effective background gravitational coupling here.

If we consider adiabatic perturbations from an ekpyrotic phase entering the bounce, recalling (3.41), then we expect from (3.85)

$$\zeta_k(\tau_{\rm b+}) \simeq \zeta_k(\tau_{\rm b-})\left(1+2i\left(1-\frac{M_{\rm Pl}^2}{\widetilde{M}_{\rm Pl}^2}\right)k\tau_{\rm b+}\right), \tag{3.86}$$

assuming again $\tau_{\rm b+}=-\tau_{\rm b-}$. Given $k\tau_{\rm b+}\ll 1$ for IR modes, this implies ζ_k receives marginal growth through the bounce phase considering $1-M_{\rm Pl}^2/\widetilde{M}_{\rm Pl}^2\sim\mathcal{O}(1)$. If we consider adiabatic perturbations from a massless phase entering the bounce, recalling (3.56), then we expect

$$\zeta_k(\tau_{\rm b+}) \simeq \zeta_k(\tau_{\rm b-})\left(1+\frac{2}{-\ln(-k\tau_{\rm b-})}\left(1-\frac{M_{\rm Pl}^2}{\widetilde{M}_{\rm Pl}^2}\right)\right), \tag{3.87}$$

where we further assumed $-k\tau_{\rm b-}$ is sufficiently small so that $-\ln(-k\tau_{\rm b-})\gg|\ln(2)-\gamma_{\rm E}+i\pi/2|\approx 1.575$. This implies that ζ_k may grow through the bounce, but only by a limited amount, small relative to its size when it entered the bounce. Finally, in the situation isocurvature perturbations from the prebounce phase source curvature perturbations before the bounce starts, we may expect [e.g., from (3.67)] that $\zeta_k(\tau_{\rm b-})\simeq i\pi\sqrt{2A_s/k^3}$ and $\zeta_k'(\tau_{\rm b-})\simeq i\pi\sqrt{2A_sk}\,\tau_{\rm b-}$. Thus, (3.85) implies that

$$\zeta_k(\tau_{\rm b+}) \simeq \zeta_k(\tau_{\rm b-})\left(1-2\left(1-\frac{M_{\rm Pl}^2}{\widetilde{M}_{\rm Pl}^2}\right)k^2\tau_{\rm b+}^2\right). \tag{3.88}$$

From this, we conclude again that ζ_k remains more or less constant throughout the bounce phase. This matches the expectation that ζ_k can only receive limited growth through the bounce, as derived in [99].

3.2.3 Numerical examples

In the previous subsection, we provided general details for different contracting models that then undergo a bounce due to the Cuscuton, and we presented how the different perturbation modes evolve across the different phases of evolution. Various approximations were made along the way to isolate the different regimes of interest and to derive analytical results. However, everything can be computed numerically to verify the validity of these approximations. We present such examples in this subsection for both the background and linear perturbations.

Since our focus is on testing the physics of the Cuscuton bouncing phase and since treating ψ as an ekpyrotic field or a fundamentally massless one does not affect the physics of the bounce phase significantly, we only present here numerical results for the case where ψ is massless. Provided a Cuscuton potential $U(\varphi)$, this means that we can numerically solve

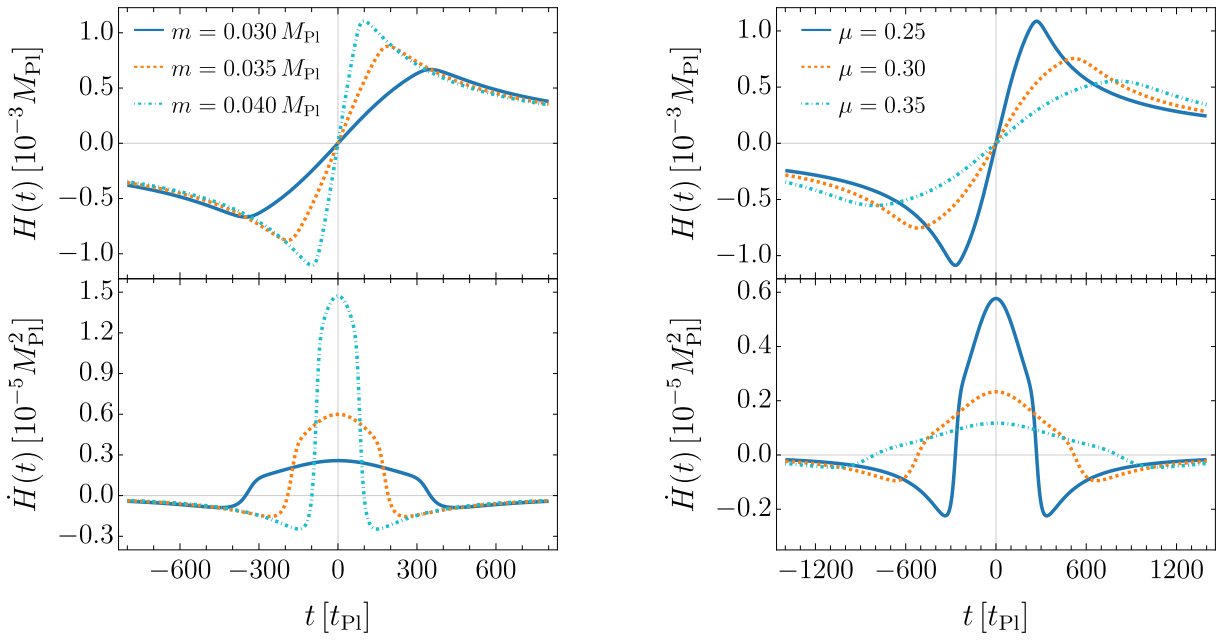


Figure 3.1: Numerical background solutions for the Hubble parameter H and its time derivative $\dot{H}$ as functions of the physical time t. The Cuscuton parameters m and μ are varied in the left- and right-hand plots, respectively. In the left panel, we fix $\mu = 0.2$ and $\varphi_\infty = 0.1\,M_{\rm Pl}$. In the right panel, we fix $m = 0.04\,M_{\rm Pl}$ and $\varphi_\infty = 0.2\,M_{\rm Pl}$.

(3.12) for $\bar{\varphi}(t)$, and then any other background quantities can be derived from the solution for $\bar{\varphi}(t)$. For instance, recalling (3.7d), we have $H(t) = U_{,\varphi}(\bar{\varphi}(t))/(3\mu^2 M_{\rm Pl}^2)$. In the spirit of following up on [98, 172], we choose the following Cuscuton potential in our numerical tests,

$$U(\varphi) = \frac{1}{2}m^2(\varphi^2 - \varphi_\infty^2) - \frac{m^4}{4}\left(e^{2(\varphi^2-\varphi_\infty^2)/m^2} - 1\right) , \tag{3.89}$$

for some constants $m, \varphi_\infty > 0$. For more details on why this particular form of the potential works, we refer the reader to [98, 172], but in short, provided $\varphi_\infty^2 \gg m^2$, one has $U_{\rm b} = U(0) \simeq -m^2\varphi_\infty^2/2 < 0$ and $m_{\rm b}^2 = U_{,\varphi\varphi}(0) \simeq m^2 > 0$ as desired. This choice of potential is in no way unique, but the numerical results presented here with this choice of potential are qualitatively quite generic.

A set of background solutions for different Cuscuton potential parameters is shown in Fig. 3.1. The time range is chosen as to put the emphasis mostly on the bouncing phase since this is the nontrivial part of the background evolution. For the prebounce contracting phase, the numerical solutions approach an $\epsilon = 3$ constant EoS, with $H(t) \simeq 1/(3t)$ as $-t$ becomes large (recall that we assume ψ is massless here). Note that the same scaling is found at large t (i.e., after the bounce), because we have not implemented reheating in this example — ψ is the only matter field ever present in this toy model. All examples shown exhibit a

bounce, that is, H crosses 0 and $\dot{H}$ has a regime where it is positive; hence, the effective NEC is violated. We can quantitatively verify from the solutions that $\dot{H}\Delta t^2 \lesssim \mathcal{O}(1)$, where Δt indicates the duration of the bounce phase [defined as the time interval where $\dot{H}(t) > 0$]. This supports the approximation made in (3.68) and throughout Sec. 3.2.2 that the scale factor is roughly constant to leading order through the bounce phase, so that in this regime $t \approx \tau$, $H \approx \mathcal{H}$, $\dot{H} \approx \mathcal{H}'$, etc. Moreover, the linear behavior of H as a function of time is manifest about the bounce point $t = 0$, supporting the approximations in (3.68), which were made to solve the perturbation equations analytically about the bounce point. Of course, $\dot{H}$ is not exactly constant in the bounce phase in the examples of Fig. 3.1 — it is more like a flattened bell curve — but nevertheless, this shows $\dot{H}$ is dominated by a more or less constant value once $\dot{H}$ has sharply transitioned from negative to positive values.

Given numerical solutions for the background, we may numerically solve the equations for the perturbations without any approximations, i.e., the equations of (3.27) with all the coefficients described in (3.25) and (3.26). From this point on, for all the numerical solutions shown as examples, we choose the following model parameters: $m = 0.04\, M_{\rm Pl}$, $\varphi_\infty = 0.1\, M_{\rm Pl}$, and $\mu = 0.2$; the corresponding background is the cyan dot-dashed curve in the left panel of Fig. 3.1. The initial conditions for the perturbations are set such that we have an adiabatic vacuum on subhorizon scales in the far past (at a prebounce time $\tau \ll -1/k$ for a given k mode). For modes of observational interest, the transition from sub- to superhorizon scales occurs far away from the bounce phase, in a regime where the contributions from the Cuscuton are completely negligible (a fact that we checked numerically). Correspondingly, we confirm that the numerical solutions are nearly identical to the analytical solutions expressed in terms of the Hankel function of the first kind shown in the previous subsection. As such, we may use the analytical solutions (in terms of the Hankel function) to set the initial conditions on superhorizon scales, as long as it is sufficiently far away from the bounce, and numerically evolve from there until after the bounce phase. Since the computational cost associated with numerically solving perturbations from the subhorizon regime to the superhorizon regime is high, it is advantageous to initiate the analysis using the Hankel function on superhorizon scales. This approach is particularly beneficial given that the primary interest lies in the physics of superhorizon modes throughout the bounce phase.

The numerical solutions for the absolute value of the curvature perturbations, $|\zeta_k(t)|$, are shown in Fig. 3.2 in the case where ψ is massless. The colored dashed curves show the numerical solutions, where the color coding represents the wavenumber k of the perturbations (we choose a range that roughly corresponds to the observable scales in the CMB). The black vertical dotted lines indicate the moments $\dot{H}$ crosses zero, so in the left plot the focus of the logarithmic time scale is on the contracting phase and the beginning of the bounce phase, while in the right plot the focus of the linear time scale is on the whole bounce phase. In the

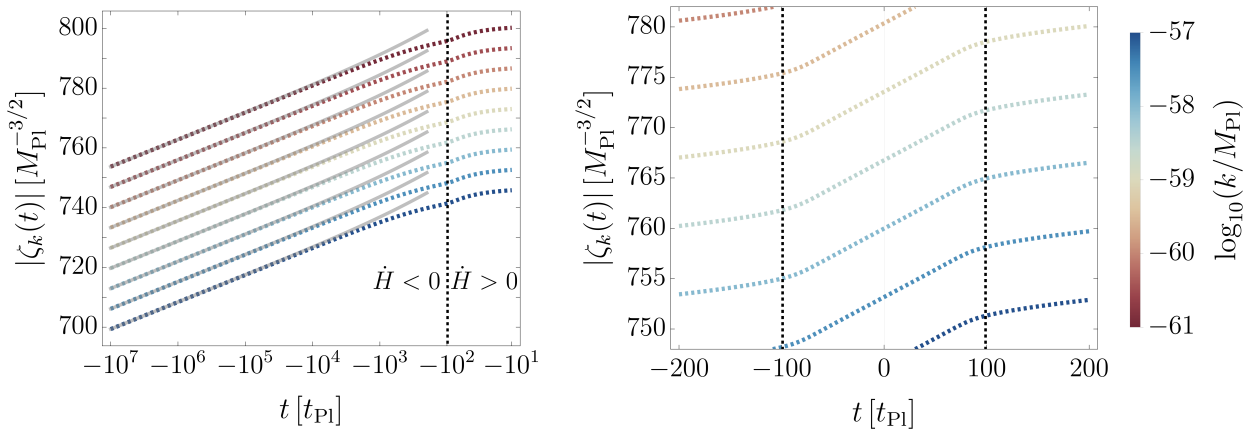


Figure 3.2: Numerical solutions for the absolute value of the curvature perturbations $|\zeta_k(t)|$ on superhorizon scales. The left panel shows the contracting phase alongside the beginning of the bounce phase on a logarithmic time scale. Dashed colored curves show the numerical solutions, and the grey solid lines show the corresponding analytic approximation. The color represents the wavenumber k according to the color bar on the far right. The vertical black dotted lines indicate when $\dot{H}=0$, representing the start and end of the bounce phase. The right panel shows a zoomed-in version of the whole bouncing phase on a linear time scale.

left plot, the faint gray solid lines show the analytical approximation (3.56), which should hold in the contracting phase well before the bounce phase, for the same set of k modes as the numerical solutions. As we can see, the numerical and analytical curves match almost perfectly for $t/t_{\rm Pl} \lesssim -10^4$. Even by $t/t_{\rm Pl} \sim -10^{2.2}$ (where we cut the gray lines for visual purposes, though the bounce phase really starts at $t/t_{\rm Pl} \approx -99$), the relative difference between the analytical and numerical curves remains very small ($\sim 1\%$). This deviation is due to the Cuscuton becoming important in the approach to the bounce — indeed, the bounce is triggered because of the Cuscuton, so its effect should become manifest in that regime.

The plot on the left of Fig. 3.2 shows that superhorizon adiabatic perturbations grow logarithmically in a contracting phase dominated by a massless scalar (note that the horizontal axis is logarithmic, while the vertical axis is linear). Similarly, we see that the k dependence is very weak, again logarithmic in fact: the amplitude of $|\zeta_k|$ at any fixed time changes by less than 10% across 4 orders of magnitude in k. While the plots show that the dimensionful quantity $|\zeta_k|$ is nearly independent of k, we can infer that the dimensionless power spectrum $\sim k^3|\zeta_k|^2$ is strongly blue, as anticipated analytically.

The right plot of Fig. 3.2 shows that the numerical solutions match the analytical expectation for $\zeta_k(t)$ also through the bounce, which is that it grows linearly as a function of time, with a rate that is an $\mathcal{O}(1)$ factor times its rate when it entered the bounce phase; recall

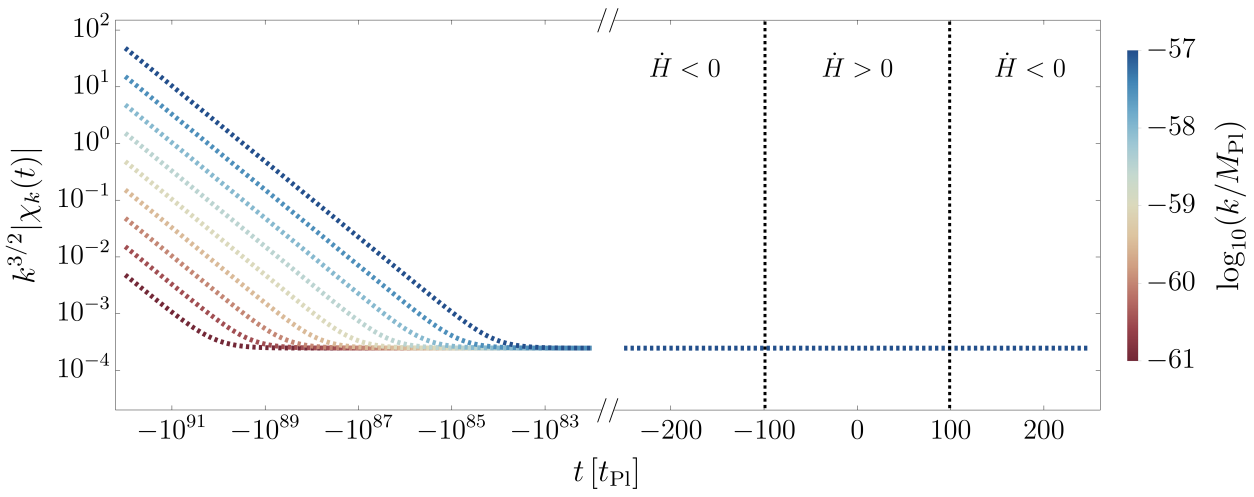


Figure 3.3: Behavior of $k^{3/2}|\chi_k(t)|$ from the numerical solutions for the entropy perturbations, $\chi_k(t)$. The plot shows part of the subhorizon evolution, horizon crossing, and part of the superhorizon evolution (on a logarithmic time scale to the left of the break), and it shows the evolution through the bounce phase (on a linear time scale to the right of the break). The colors indicate the wavenumber k. The vertical black dotted lines show when $\dot{H} = 0$, indicating the beginning and end of the bounce phase.

(3.85). Indeed, we see quite neatly the change of slope in $\dot{\zeta}_k(t)$ (roughly equivalent to $\zeta_k'(\tau)$ in the bounce phase) going from the contracting phase to the bouncing phase. Moreover, the amplification $|\zeta_k|$ receives passing through the bounce is relatively small compared to its amplitude, again as anticipated analytically in (3.87).

Moving on to entropy perturbations, we take (3.61) as the choice of coupling function, with $\Lambda = 10^{-3.5}\, M_{\rm Pl}$ and $\delta = 0$. This will yield an exactly scale-invariant power spectrum, which is sufficient for the sake of presenting the general behavior of χ_k, but the computation is straightforward to repeat with a slight $\delta > 0$. We could also take (3.60) instead of (3.61) for the coupling function. However, they have the same time dependence when the Cuscuton is negligible. Therefore, they produce identical results for χ_k during the contracting phase and expectedly through the bounce, as discussed in the previous subsection. We have confirmed this numerically.

The numerical solutions for the entropy perturbation $\chi_k(t)$ are shown in Fig. 3.3, where we plot the dimensionless quantity $k^{3/2}|\chi_k(t)|$, essentially the square root of the dimensionless power spectrum as a function of time. Curves of different colors represent different k modes as in the previous figure. The time axis is broken as to put the emphasis on two processes happening in very different time regimes: when the modes exit the horizon; and when they pass through the bounce phase (in between the two black vertical dotted lines). The

decrease in $|\chi_k(t)|$ when the modes are subhorizon is due to the fact that[15] $z_\chi = a\sqrt{F} \sim 1/(-\tau)$, so then $|\chi_k(\tau)| \propto |u_k|/z_\chi(\tau) \sim (-\tau)$ if u_k is in its adiabatic vacuum state ($|u_k| \sim 1/\sqrt{2k}$). Once the modes cross the horizon, they freeze and χ_k becomes nearly constant; in fact, the numerical results match the analytical approximation (3.63) according to which $\chi_k'(\tau)/\chi_k(\tau) \simeq k^2\tau$, which means that the growth of $|\chi_k(\tau)|$ in the supper-Hubble regime as $k\tau \to 0^-$ is very much suppressed. Figure 3.3 also shows that as the modes of different k freeze on superhorizon scales, they all approach the same value for $k^{3/2}|\chi_k| \sim 2 \times 10^{-4}$ — this is the behavior to be expected for an exact scale-invariant dimensionless power spectrum (recall that we took $\delta = 0$) of amplitude $\sim 10^{-9}$. The plots further show that the numerical solutions remain essentially constant as they pass through the bounce phase, which matches the analytical approximation given in (3.76) — the possible amplification through the bounce phase is IR-suppressed as $(k\tau_{\mathrm{b}-})^2 \ll 1$. The lesson here is that the bounce phase does not affect the scale-invariant spectrum. Note that if the χ_k isocurvature perturbations are converted into curvature perturbations ζ_k right before the bounce phase, then the same conclusion would be reached according to the discussion around (3.88).

3.3 Beyond the linear regime

Since this section covers multiple aspects, we provide a brief guideline on how to read it. The main objective is to compute the third-order action, determine the strong coupling scale, and analyze the bispectrum to quantify non-Gaussianities. We carry out this analysis both analytically and numerically across different phases and scales. We begin in Sec. 3.3.1 with a brief literature review and theoretical expectations. In Sec. 3.3.2, we derive the third-order action. Section 3.3.3 focuses on the strong coupling scale, considering linear perturbations and interactions, analyzed separately for the prebounce and bounce phases across IR and UV scales. Section 3.3.4 then presents the bispectrum and non-Gaussianity analysis, addressing the same cases as before. The strong coupling scale is revisited using bispectrum calculations and compared to previous results. Finally, we end the section with numerical results that support the analytical estimates.

3.3.1 Theoretical insights and expectations

So far, we have seen how the Cuscuton may yield a nonsingular bouncing cosmology at the background level, and we have seen examples of cosmological perturbations that may

[15] A word on notation: $\sim$ should be read as 'approximately scales as' here (similar to $\simeq$, but up to proportionality constants, ignoring numerical factors). Elsewhere, it may mean 'of the order of'. Its interpretation should be clear from context.

evolve in such a background. Specifically, a large-scale nearly scale-invariant 2-point function of entropy perturbations may be generated either during an ekpyrotic or stiff contracting phase, together with unamplified adiabatic perturbations. Provided that the isocurvature perturbations are efficiently converted into curvature perturbations, the models could match the current observational constraints from the CMB.

However, these findings are based on the study of linear perturbations and their corresponding 2-point statistics, but a detailed comparison with CMB constraints requires advancing the analysis beyond the linear regime. In particular, nonlinearities should not compromise the results. For example, the existence of strong coupling issues or the production of (very large) large non-Gaussianities would represent severe drawbacks to the theory.

Naturally, when the Cuscuton contribution is negligible (i.e., deep in the contracting phase before the bounce) and before the conversion of isocurvature to curvature perturbations, we should recover the results of the literature, according to which the perturbations should remain nearly perfectly Gaussian (see, e.g., [169, 170, 181] and references therein). This makes intuitive sense: during that phase, ζ is effectively unexcited and remains very close to its Gaussian vacuum state, and χ has no cubic interaction term in the action (since, in particular, it has no potential), hence, it has a vanishing 3-point correlation function. To better understand the rationale behind this intuition, we note that in this regime, the theory is very close to GR with canonical scalar fields, so in particular the scalars' sound speed is very close to unity, and initial conditions are taken according to a Lorentzian adiabatic vacuum; hence non-Gaussianities of equilateral and orthogonal shape should not get generated significantly (see, e.g., [45, 211]).[16] Furthermore, local non-Gaussianities are similarly not generated, as can be assessed through the EoMs for ζ and χ. That is, if we expand the perturbations to second order at the level of their EoMs, we find no source for the entropy mode at quadratic order since there is no potential for χ and since there is no turning of the trajectory in $\{\psi, \chi\}$-space [169, 170, 181]. In other words, ζ and χ are decoupled at the perturbative level.

Therefore, no intrinsic local non-Gaussianities are generated for the entropy perturbations. Similarly, if we expand the curvature perturbation to second order at the level of its EoM on large scales, one finds that it is sourced to second order according to (in the superhorizon limit; see again, e.g., [169, 170, 181])

$$\dot{\zeta}^{(2)} \simeq \frac{HV_{,\psi}}{\dot{\psi}^2} s \left(\frac{1}{2} \frac{F_{,\psi}}{F} s - \frac{\dot{s}}{\dot{\psi}} \right) = - \frac{HV_{,\psi} F}{\dot{\psi}^3} \chi \dot{\chi} \,, \tag{3.90}$$

where $s \equiv \sqrt{F}\chi$. However, we know that the superhorizon solution for χ is constant in time

[16] The intuition really comes from inflation, where $\epsilon < 1$. Since $\epsilon \gtrsim \mathcal{O}(1)$ for contracting models, the non-Gaussianity results from inflation are not directly applicable. Still, they may provide insight.

to leading order [i.e., $\dot{\chi} \simeq 0$, recall (3.49b) or (3.63)], hence $s \sim \sqrt{F}$ and $\dot{\zeta}^{(2)} \simeq 0$. Therefore, the curvature perturbation is not sourced by entropy modes at this point, indicating once more no deviations from Gaussianity.[17] We note that these statements remain true even to third order [181], hence we expect $f_{\rm NL}^{\rm (local)} \approx 0$ and $g_{\rm NL}^{\rm (local)} \approx 0$. These quantities are the coefficients of the expansion of the full real-space perturbation $\zeta(\boldsymbol{x})$ in powers of a Gaussian variable $\zeta_{\rm G}(\boldsymbol{x})$, schematically $\zeta(\boldsymbol{x}) = \zeta_{\rm G}(\boldsymbol{x}) + f_{\rm NL}^{\rm (local)}\zeta_{\rm G}^2(\boldsymbol{x}) + g_{\rm NL}^{\rm (local)}\zeta_{\rm G}^3(\boldsymbol{x}) + \cdots$. We will provide a more formal definition when we discuss the computation of the bispectrum in a subsequent subsection.

Two things can still generate non-Gaussianities at this point: the conversion of isocurvature perturbations into curvature perturbations; and the Cuscuton's contribution. The former has already been considered quite extensively in the literature [170, 181, 182, 201, 203–206]. The conversion can be modeled in different ways, but the general idea is that there must exist a potential barrier that induces a rotation in $\{\psi, \chi\}$-space. If the rotation rate is $\dot{\theta} = V_{,s}/\dot{\psi}$, where we recall $s(\psi, \chi) = \sqrt{F(\psi)}\chi$, then curvature perturbations are sourced on large scales following $\dot{\zeta} \simeq 2H\dot{\theta}s/\dot{\psi}$, after which ζ roughly acquires the amplitude and spectral dependence of χ. This conversion generates local non-Gaussianities of the order of $f_{\rm NL}^{\rm (local)} = \mathcal{O}(1-10)$, depending on the precise modeling of the conversion process [170, 181, 182, 201, 203–206]. This is very close to current observational bounds, e.g., $f_{\rm NL}^{\rm (local)} = -0.9 \pm 5.1$ (65% CL) from *Planck* [27]. This tells us that if there is any Cuscuton contribution to $f_{\rm NL}^{\rm (local)}$, it should not exceed that constraint either, that is, not much larger than $\mathcal{O}(1)$. Future observational constraints on $f_{\rm NL}$ will certainly constrain this even more (see, e.g., [212–217] and references therein).

This naturally leads us to compute the bispectrum with the Cuscuton included to determine both that the theoretical analysis of the model is sound and that the model is not excluded by observations. In particular, the behavior of the bispectrum going through the bounce will be crucial since it is precisely the presence of the Cuscuton that enables the nonsingular transition. Two things could go wrong: the first one, as already pointed out, would be that large non-Gaussianities could be generated through the bounce such that observational bounds rule out the models; and the second one would be that the theory becomes strongly coupled, meaning in this context that the perturbative expansion would essentially break down, indicating that the linear perturbations would be untrustworthy. Many nonsingular bounces (using different physics from Cuscuton) are known to suffer from either issue, if not both (e.g., [112, 160, 218, 219]). We will show that the Cuscuton bounce is robust

[17] Subleading effects could still generate small local non-Gaussianities at this point, such as the decaying contribution to χ. Also, more refined models (especially of ekpyrosis) would have a slightly time-dependent background EoS, which would also have the effect of sourcing ζ to second order. Such effects remain tiny, though [169].

in that respect: no sizable non-Gaussianities are produced through the bounce, whatsoever, and the theory remains well under perturbative control throughout. We estimate the strong coupling scale along the way and show that it is well above the scale of the background.

3.3.2 Third-order action

The strategy we shall employ to investigate strong coupling and compute bispectra is to use the action expanded to third order in perturbations. We show the details of how this is obtained in this subsection.

From this point on, we only consider the case where ψ is massless at the level of the interactions. This would not be applicable deep in the contracting phase for ekpyrosis, but we already know that the perturbations remain Gaussian in that regime. Thus, to investigate the effects of the Cuscuton and the nonsingular bounce, it is fair to set the potential for ψ to zero from the start. We then expand the action given in (3.4) to third order in α, β, ζ, $\delta\varphi$, and χ defined in (3.5) and (3.6), always performing some integration by parts and using the background EoMs for simplifications. For better illustration, we classify the third-order action in three parts:

$$S^{(3)} = M_{\rm Pl}^2 \int \mathrm{d}^3\boldsymbol{x}\,\mathrm{d}t\,a^3 \left(\mathcal{L}^{(3)}_{(\chi)} + \mathcal{L}^{(3)}_{(\delta\varphi)} + \mathcal{L}^{(3)}_{(\rm rest)}\right) . \tag{3.91}$$

Note that $\mathcal{L}^{(3)}_{(\chi)}$ is not meant to represent the third-order expansion of $\mathcal{L}_\chi$ as defined in (3.3) and likewise for the other terms. Instead, we have in mind expanding the full action to third order and gathering all the cubic terms involving two χs first in $\mathcal{L}^{(3)}_{(\chi)}$, then all the remaining cubic terms involving at least one $\delta\varphi$ in $\mathcal{L}^{(3)}_{(\delta\varphi)}$, and finally all the remaining cubic terms in $\mathcal{L}^{(3)}_{(\rm rest)}$ (before the substitution of any constraint equations for α, β, and $\delta\varphi$).[18] This yields

$$\mathcal{L}^{(3)}_{(\chi)} = -\frac{1}{2}F(\psi,X)\left((\alpha-3\zeta)\dot\chi^2 + (\alpha+\zeta)\frac{(\partial_i\chi)^2}{a^2} + 2\dot\chi\frac{\partial^i\beta\partial_i\chi}{a^2}\right) - XF_{,X}\alpha\left(\dot\chi^2 - \frac{(\partial_i\chi)^2}{a^2}\right) , \tag{3.92a}$$

$$\begin{aligned}\mathcal{L}^{(3)}_{(\delta\varphi)} = {}& 9\mu^2(\dot\zeta - H\alpha)\zeta\,\delta\varphi + \frac{3\mu^4\epsilon(\alpha+3\zeta)}{4\sigma}\delta\varphi^2 - \frac{\mu^6\epsilon_\psi\epsilon(2\epsilon-\eta-6)}{8\sigma^3 H}\delta\varphi^3 + \mu^2\zeta\frac{\partial^i\beta\partial_i\delta\varphi}{a^2} \\ & + \frac{\mu^4}{4\sigma H^2}\left(2\alpha+\zeta-\frac{\mu^2\dot{\delta\varphi}}{2\sigma H^2}\right)\frac{(\partial_i\delta\varphi)^2}{a^2} ,\end{aligned} \tag{3.92b}$$

$$\mathcal{L}^{(3)}_{(\rm rest)} = (3-\epsilon_\psi)H^2\alpha^3 + 3(\epsilon_\psi-3)H^2\alpha^2\zeta - 6H\alpha^2\dot\zeta + 18H\alpha\zeta\dot\zeta + 3\alpha\dot\zeta^2 - 9\zeta\dot\zeta^2 + 2H\zeta\frac{\partial^i\alpha\partial_i\beta}{a^2}$$

[18] Note also that the 'Lagrangian density' as defined by $S^{(3)} = M_{\rm Pl}^2\int \mathrm{d}^3\boldsymbol{x}\,\mathrm{d}t\,a^3\mathcal{L}^{(3)}$ does not have the appropriate dimensions given the factorization of $M_{\rm Pl}^2$, i.e., henceforth $\mathcal{L}$ has dimensions of mass squared. Still, we abuse notation given the practicality of this factorization.

$$+(\zeta-\alpha)\frac{(\partial_i\zeta)^2}{a^2}-2\zeta\frac{\partial^i\beta\partial_i\dot\zeta+\alpha\partial^2\zeta}{a^2}+2(H\alpha-\dot\zeta)\alpha\frac{\partial^2\beta}{a^2}+\alpha\frac{(\partial^2\beta)^2-(\partial_i\partial_j\beta)^2}{2a^4}$$
$$+\frac{\partial^2\beta\partial^i\beta\partial_i\zeta}{a^4}+\frac{3\partial^i\beta\partial^j\beta\partial_i\partial_j\zeta}{2a^4}\,. \tag{3.92c}$$

Note that partial derivatives are always meant to act on immediate variables only. Upon substituting the solutions for the perturbed lapse and shift (3.15) and recalling the definition of Q in (3.16), the action can be recast in the form of (3.91), except now with different $\mathcal{L}$s. First, the terms that couple to χ are

$$\begin{aligned}\mathcal{L}^{(3)}_{(\chi)}=&-\frac{F(\psi,X)}{2H}\Big(\dot\zeta-3H\zeta-\frac{\mu^2}{2}\delta\varphi\Big)\dot\chi^2-\frac{F(\psi,X)}{2H}\Big(\dot\zeta+H\zeta-\frac{\mu^2}{2}\delta\varphi\Big)\frac{(\partial_i\chi)^2}{a^2}\\&+F(\psi,X)\left(\frac{\partial_i\zeta}{aH}-a\epsilon_\psi\Big(\partial_i\partial^{-2}\dot\zeta-\frac{\mu^2}{2}\partial_i\partial^{-2}\delta\varphi\Big)\right)\frac{\partial^i\chi}{a}\dot\chi\\&-\frac{XF_{,X}}{H}\left(\dot\zeta-\frac{\mu^2}{2}\delta\varphi\right)\left(\dot\chi^2-\frac{(\partial_i\chi)^2}{a^2}\right)\,.\end{aligned} \tag{3.93}$$

Then, terms that still include the Cuscuton contribution can be divided as $\mathcal{L}^{(3)}_{(\delta\varphi)}=\mathcal{L}^{(3)}_{(\delta\varphi)^3}+\mathcal{L}^{(3)}_{(\delta\varphi)^2}+\mathcal{L}^{(3)}_{(\delta\varphi)^1}$ with

$$\mathcal{L}^{(3)}_{(\delta\varphi)^3}=-\frac{\mu^6}{8\sigma^2a^2H^4}\Big(2\sigma H\,\delta\varphi+\delta\dot\varphi\Big)(\partial_i\delta\varphi)^2-\frac{\mu^6\epsilon_\psi C_3}{16\sigma^3H}\delta\varphi^3+\frac{\mu^6\epsilon_\psi^2}{16H}(\partial_i\partial_j\partial^{-2}\delta\varphi)^2\,\delta\varphi\,, \tag{3.94a}$$

$$\begin{aligned}\mathcal{L}^{(3)}_{(\delta\varphi)^2}=&\frac{\mu^4}{4\sigma a^2H^3}(2\dot\zeta+H\zeta)(\partial_i\delta\varphi)^2+\frac{\mu^4}{a^2H^2}\delta\varphi\,\partial^i\zeta\partial_i\delta\varphi-\frac{\mu^4\epsilon_\psi}{4a^2H^2}\delta\varphi^2\,\partial^2\zeta\\&+\frac{\mu^4\epsilon_\psi}{8\sigma H}\Big(6(\sigma+3)H\zeta+\big(\sigma(2+3\epsilon_\psi)+6\big)\dot\zeta\Big)\delta\varphi^2+\frac{\mu^4\epsilon_\psi}{4a^2H^2}\delta\varphi\,\partial^i\partial^j\zeta\partial_i\partial_j\partial^{-2}\delta\varphi\\&-\frac{\mu^4\epsilon_\psi^2}{8H}\Big(2\,\delta\varphi\,\partial^i\partial^j\partial^{-2}\delta\varphi\,\partial_i\partial_j\partial^{-2}\dot\zeta+\dot\zeta(\partial_i\partial_j\partial^{-2}\delta\varphi)^2\\&-H\partial^i\partial^{-2}\delta\varphi\,\big(2\,\delta\varphi\,\partial_i\zeta+3\partial^j\partial^{-2}\delta\varphi\,\partial_i\partial_j\zeta\big)\Big)\,,\end{aligned} \tag{3.94b}$$

$$\begin{aligned}\mathcal{L}^{(3)}_{(\delta\varphi)^1}=&\frac{\mu^2\epsilon_\psi}{4H}\Big((2-3\epsilon_\psi)\dot\zeta-12H\zeta\Big)\,\dot\zeta\,\delta\varphi+\frac{\mu^2}{2a^2H^2}\Big(H(1+\epsilon_\psi)(\partial_i\zeta)^2+2\big(H\zeta+(1+\epsilon_\psi)\dot\zeta\big)\partial^2\zeta\Big)\,\delta\varphi\\&+\frac{\mu^2}{4a^4H^3}\left((\partial_i\partial_j\zeta)^2-(\partial^2\zeta)^2\right)\delta\varphi+\frac{\mu^2\epsilon_\psi^2}{4H}\partial^i\partial^j\partial^{-2}\dot\zeta\Big(\delta\varphi\,\partial_i\partial_j\partial^{-2}\dot\zeta+2\dot\zeta\partial_i\partial_j\partial^{-2}\delta\varphi\Big)\\&-\frac{\mu^2\epsilon_\psi}{2a^2H^2}\Big(\partial^i\partial^j\zeta\big(\delta\varphi\,\partial_i\partial_j\partial^{-2}\dot\zeta-2H\partial_i\partial^{-2}\delta\varphi\partial_j\zeta\big)+\big(H\partial^i\zeta\partial^j\zeta+\dot\zeta\partial^i\partial^j\zeta\big)\partial_i\partial_j\partial^{-2}\delta\varphi\Big)\\&-\frac{\mu^2\epsilon_\psi^2}{2}\Big(\delta\varphi\,\partial^i\zeta\partial_i\partial^{-2}\dot\zeta+\partial^i\partial^{-2}\delta\varphi\big(\dot\zeta\partial_i\zeta+3\partial^j\partial^{-2}\dot\zeta\partial_i\partial_j\zeta\big)\Big)\,.\end{aligned} \tag{3.94c}$$

In (3.94a) to lighten the (already very long) expression we defined

$$C_3=-\epsilon^3(2+\epsilon_\psi)+\epsilon_\psi^2(6+\epsilon_\psi(2+\epsilon_\psi))+\epsilon^2(10+3\epsilon_\psi(2+\epsilon_\psi))-3\epsilon(4+\epsilon_\psi(4+\epsilon_\psi(2+\epsilon_\psi)))-2\epsilon\eta\,. \tag{3.95}$$

Finally, the remaining third-order terms containing ζ are

$$\begin{aligned}\mathcal{L}^{(3)}_{(\text{rest})} = &\frac{\epsilon_\psi(\epsilon_\psi-2)}{2H}\dot\zeta^3 + 3\epsilon_\psi\zeta\dot\zeta^2 - \frac{\epsilon\zeta(\partial_i\zeta)^2}{a^2} - \frac{\epsilon_\psi\dot\zeta^2\partial^2\zeta}{a^2H^2} + \frac{3\epsilon\partial^2\zeta(\partial_i\zeta)^2}{4a^4H^2}\\ &+\frac{\epsilon_\psi^2}{2H}\left(H\partial^i\partial^{-2}\dot\zeta\left(2\dot\zeta\partial_i\zeta + 3\partial^j\partial^{-2}\dot\zeta\partial_i\partial_j\zeta\right) - \dot\zeta(\partial_i\partial_j\partial^{-2}\dot\zeta)^2\right)\\ &+\frac{\epsilon_\psi}{a^2H^2}\partial^i\partial^j\partial^{-2}\dot\zeta\left(H\partial_i\zeta\partial_j\zeta + \dot\zeta\partial_i\partial_j\zeta\right)\,. \end{aligned} \tag{3.96}$$

In principle at this point, the solution to the constraint $\delta\varphi$, (3.17), could be substituted to further obtain an action that depends solely on the two dynamical degrees of freedom χ and ζ. However, the resulting expression is tremendously long and is not particularly instructive. One could hope that integration by parts further simplifies the action as is the case in GR [220], but we have not found this to be the case here (more comments about this below). Still, in what follows, we will find that the third-order action as expressed above is already useful.

An important observation here is about the GR limit. If one sets the Cuscuton and the entropy field to zero (μ, U, and F all equivalently 0), then the theory reduces to GR with a free (noninteracting, massless) scalar. In the third-order action above, this means only $\mathcal{L}^{(3)}_{(\text{rest})}$ [eq. (3.96)] survives with $\epsilon_\psi = \epsilon = 3$. In such a case, one can verify that the action is indeed equivalent to what one expects in GR with a free scalar (see, e.g., [32, 220]), namely, after integrating by parts and using the GR EoM for ζ,

$$\mathcal{L}^{(3)}_{(\text{GR})} = 9\zeta\dot\zeta^2 + 9\zeta\frac{(\partial_i\zeta)^2}{a^2} - \frac{9}{2}\dot\zeta\partial_i\zeta\partial^i\partial^{-2}\dot\zeta + \frac{27}{4}\partial^2\zeta(\partial_i\partial^{-2}\dot\zeta)^2\,. \tag{3.97}$$

When we have the Cuscuton, the 'GR part' of our *full* third-order action, which we called $\mathcal{L}^{(3)}_{(\text{rest})}$, is generally different from $\mathcal{L}^{(3)}_{(\text{GR})}$. Indeed, one could contemplate extracting the GR terms by performing integration by parts and using the full linear EoM for ζ as in the standard calculation (e.g., [32, 220]). However, the second-order action for ζ, (3.20), does not straightforwardly factorize into a 'GR part' and a 'Cuscuton part', and the nonlocality of the corresponding EoM for ζ and of the constraint equation for $\delta\varphi$ makes it practically impossible to further simplify (3.96) [and (3.94) for that matter]. One implication of this is that, when the Cuscuton is important, (3.96) has higher-order operators of the form $\nabla^4\zeta^3/H^2$, while in the GR limit we only expect operators of the form $\nabla^2\zeta^3$ because higher-order operators like $\nabla^4\zeta^3/H^2$ amount to boundary terms. [Henceforth, we use ∇ to schematically denote the power counting of derivatives (spatial or temporal).] Yet, formally in the UV limit $k\to\infty$, the Cuscuton reduces to GR [86], so one could contemplate the possibility that terms of the form $\nabla^4\zeta^3/H^2$ may be removed from the third-order Cuscuton action by integration by parts perturbatively in $1/k$. We checked and confirm that, before integration by parts, the third-order action of the Cuscuton and of GR match to leading order in $1/k$, and they are of the form $\nabla^4\zeta^3/H^2$. More comments on this will follow later.

3.3.3 Toward estimating the strong coupling scale

Strategy

The first thing we wish to do with the knowledge of the third-order action is to compare the 'size' of its terms to those of the second-order action. If we express these actions as, say, $S^{(3)} = M_{\rm Pl}^2 \int \mathrm{d}^3\boldsymbol{x}\,\mathrm{d}t\,a^3\mathcal{L}^{(3)}$ and $S^{(2)} = M_{\rm Pl}^2 \int \mathrm{d}^3\boldsymbol{x}\,\mathrm{d}t\,a^3\mathcal{L}^{(2)}$, respectively, then the ratio $\mathcal{L}^{(3)}(t,\boldsymbol{x})/\mathcal{L}^{(2)}(t,\boldsymbol{x})$ is informative of the perturbative expansion's regime of validity and of the theory's strong coupling scale (e.g., [137, 141, 143, 144, 221]). Indeed, in order for ζ and χ to be meaningful perturbation variables and for the theory to be controllable, one should have

$$\mathcal{L} = \mathcal{L}^{(0)} + \mathcal{L}^{(2)} + \mathcal{L}^{(3)} + \dots \tag{3.98}$$

with every term smaller than the previous one. (The expansion generally includes $\mathcal{L}^{(1)}$, but it is often omitted since up to boundary terms it vanishes on shell, i.e., upon applying the background EoMs.) As well explained in [137], two things could go wrong: for subhorizon fluctuations, a ratio $|\mathcal{L}^{(3)}/\mathcal{L}^{(2)}| \gtrsim 1$ is indicative of a quantum strong coupling problem, with unsuppressed loop corrections and potential violation of unitarity; while for superhorizon fluctuations, a ratio $|\mathcal{L}^{(3)}/\mathcal{L}^{(2)}| \gtrsim 1$ is indicative of large classical nonlinearities, which may spoil the analysis of the linear perturbations and perhaps even of the background. In the same spirit, $|\mathcal{L}^{(2)}/\mathcal{L}^{(0)}| \gtrsim 1$ would signify a breakdown of perturbativity. While this would have nothing to do with interactions being strongly coupled, we nevertheless abuse terminology and refer to scales where $|\mathcal{L}^{(2)}/\mathcal{L}^{(0)}| \sim 1$ as strong coupling scales. A proper analysis of the quantum strong coupling scale in cosmological contexts involves the computation of loops or nonlinear corrections to various observables to determine at what point they become substantial. As a sensible alternative, one might instead check whether these corrections are significant within the S-matrix (e.g., [112, 138–140]; see [222] for yet another alternative involving entanglement), but a rough estimate can already be found by estimating $\mathcal{L}^{(2)}/\mathcal{L}^{(0)}$ and especially $\mathcal{L}^{(3)}/\mathcal{L}^{(2)}$ (e.g., [136, 140–147]). Moreover, since our primary goal is to test the physics of the Cuscuton in the bounce phase, we will be mostly concerned with the classical strong coupling problem for superhorizon perturbations (which are of observational interest), but we will comment on the subhorizon regime too. Generally speaking, the goal is to find the regime of validity of the model at hand, which from an effective field theory perspective may come with UV (and potentially IR) cutoffs.

A comparison of perturbed Lagrangians as described above can only make sense when performed with real-space perturbations, because cosmological perturbations (through the perturbative expansion of the metric and the matter fields) are defined in real space; cf. (3.5) and (3.6). Let us denote either ζ or χ by the variable ξ. Note that since these variables are

technically quantum operators, $\hat{\xi}$, they are not deterministic functions of $\boldsymbol{x}$ in real space; instead their expectation values manifest themselves as stochastic fields. However, defining $\bar{\xi}(t;\mathscr{R})$ as the standard deviation of the coarse-grained $\hat{\xi}$ with some Window function $\mathcal{W} * \hat{\xi}(t,\boldsymbol{x})$ over some radius $\mathscr{R}$ enables us to quantify the real-space 'size' (i.e., amplitude) of the coarse-grained $\hat{\xi}$ (and likewise for the derivatives thereof). Recalling (3.44), we can evaluate $\bar{\xi}(t;\mathscr{R})$ in terms of the Fourier-space mode function ξ_k and Fourier transform of $\mathcal{W}$ as[19]

$$\bar{\xi}(t;\mathscr{R}) \equiv \sqrt{\int_{\boldsymbol{k}} \widetilde{\mathcal{W}}^2(k\mathscr{R})|\xi_k(t)|^2}\,, \qquad \bar{\dot{\xi}}(t;\mathscr{R}) \equiv \sqrt{\int_{\boldsymbol{k}} \widetilde{\mathcal{W}}^2(k\mathscr{R})|\dot{\xi}_k(t)|^2}\,. \tag{3.99}$$

Ignoring $\mathcal{O}(1)$ numerical factors, estimating the square of the Window function in Fourier space using a delta function as in (3.45), and denoting the smoothing comoving scale at which we evaluate the 'size' of the perturbation ξ as $\mathscr{R} = 1/k_{\mathscr{R}}$, we have

$$\xi(t;1/k_{\mathscr{R}}) \sim k_{\mathscr{R}}^{3/2}|\xi_{k_{\mathscr{R}}}(t)|\,, \qquad \dot{\xi}(t;1/k_{\mathscr{R}}) \sim k_{\mathscr{R}}^{3/2}|\dot{\xi}_{k_{\mathscr{R}}}(t)|\,, \tag{3.100}$$

where to simplify the notation the bar indicating the coarse-grained estimates is omitted from this point onward. From this, we can similarly estimate the size of spatial derivatives as

$$\partial\xi(t;1/k_{\mathscr{R}}) \sim k_{\mathscr{R}}\xi(t;1/k_{\mathscr{R}})\,, \tag{3.101}$$

and so on. For instance, if ξ_k is related to its Mukhanov-Sasaki variable v_k through a dimensionless function f_ξ as $v_k = M_{\rm Pl} a f_\xi \xi_k$, then for $k_{\mathscr{R}}$ in the far UV we expect for an adiabatic vacuum[20] with Lorentzian dispersion relation $\omega_k \sim k$ ($v_k \sim k^{-1/2}$ and $\dot{v}_k \sim \omega_k v_k \sim k^{1/2}$)

$$\xi \sim \frac{k_{\mathscr{R}}/a}{M_{\rm Pl} f_\xi}\,, \qquad \dot{\xi}^2 \sim \frac{(\partial_i\xi)^2}{a^2} \sim \frac{(k_{\mathscr{R}}/a)^4}{M_{\rm Pl}^2 f_\xi^2}\,. \qquad \text{(UV)} \tag{3.102}$$

As another example, when ξ represents χ which has a scale-invariant power spectrum on superhorizon scales in the prebounce phase [as in (3.50) and (3.64) in the limit $\delta = 0$], then for $k_{\mathscr{R}}$ in the far IR we find

$$\xi \sim \sqrt{A_s}\,, \qquad \dot{\xi} \sim \frac{(k_{\mathscr{R}}/a)^2}{|H|}\sqrt{A_s}\,, \qquad \frac{\partial\xi}{a} \sim \frac{k_{\mathscr{R}}}{a}\sqrt{A_s}\,. \qquad \text{(IR)} \tag{3.103}$$

[19] Since acting on a function that depends only on the magnitude of $\boldsymbol{k}$, note

$$\int_{\boldsymbol{k}} \widetilde{\mathcal{W}}^2(k\mathscr{R})|\xi_k(t)|^2 = \int_0^\infty \frac{\mathrm{d}k\ 4\pi k^2}{(2\pi)^3}\,\widetilde{\mathcal{W}}^2(k\mathscr{R})|\xi_k(t)|^2 = \int_0^\infty \frac{\mathrm{d}k}{k}\,\widetilde{\mathcal{W}}^2(k\mathscr{R})\mathcal{P}_\xi(k,t)\,,$$

where we recall $\mathcal{P}_\xi(k,t) \equiv k^3|\xi_k(t)|^2/(2\pi^2)$.

[20] Note that the time variation of a and f_ξ should be such that the WKB approximation holds in the UV limit, so that a well-defined adiabatic vacuum for v_k in that regime exists. In most scenarios in the literature, this is ensured by construction using the generalized 'slow-roll conditions'; otherwise one may seek more sophisticated transformations to check if a proper canonical variable exists to set the adiabatic vacuum conditions [223]. Also, (3.102) really holds only if $f_\xi \neq 0$ and as long as it has little k dependence (subleading in the UV).

Such estimates will be used repeatedly in what follows.

Background

Let us start by estimating the relevant energy scales of the background cosmology (see, e.g., [112]),

$$E_{\rm back}(t) \sim \max\left\{ |H|\,, \left|\dot{H}\right|^{1/2}, \left|\frac{\ddot{\psi}}{\dot{\psi}}\right|, \cdots \right\}. \tag{3.104}$$

We find $E_{\rm back}(t) \sim |H(t)| \sim 1/|t|$ before the bounce phase, that is, for $t \lesssim t_{\rm b-} \sim 1/|H_{\rm b-}|$. Thus, the maximal energy scale before the bounce is $|H_{\rm b-}|$. Using the potential (3.89) as an example, one estimates[21]

$$|H_{\rm b-}| \approx \frac{m^2\varphi_\infty}{3\mu^2 M_{\rm Pl}^2} \approx \frac{\mu^2}{3}\varphi_\infty\,, \tag{3.105}$$

where the second approximation assumes $m \approx \mu^2 M_{\rm Pl}$ [more about this later, below (3.113)]. In order for the bounce to be semiclassical (sub-Planckian), we ask $|H_{\rm b-}| \ll M_{\rm Pl}$. Given that the potential (3.89) assumes $m/M_{\rm Pl} \leq \mu^2 \ll \varphi_\infty/M_{\rm Pl} < 1$ (see [98]), this is indeed the case. Under the same approximation, (3.82) tells us that[22]

$$\Upsilon \approx \frac{m^4}{M_{\rm Pl}^4}\frac{\varphi_\infty^2}{3\mu^4 - 2m^2/M_{\rm Pl}^2} \approx \mu^4\varphi_\infty^2 \approx 9H_{\rm b-}^2\,, \tag{3.106}$$

where the second approximation assumes $m \approx \mu^2 M_{\rm Pl}$ and the third uses (3.105). Hence, according to (3.68), we expect the relevant energy scale through the bounce, $|\dot{H}|^{1/2} \approx \Upsilon^{1/2}$, to remain roughly constant and of the same order as $|H_{\rm b-}|$.

Linear perturbations

Before the bounce Recalling (3.24) and below, deep in the contraction when $\epsilon \approx \epsilon_\psi \approx$ constant, we simply have $\mathcal{L}_\zeta^{(2)} \approx \epsilon_\psi(\dot{\zeta}^2 - a^{-2}(\partial_i\zeta)^2)$ and $\mathcal{L}_\chi^{(2)} = (F/2)(\dot{\chi}^2 - a^{-2}(\partial_i\chi)^2)$, and this is true on all scales. One can check the estimated coarse-grained size of $\mathcal{L}_\zeta^{(2)}$ and $\mathcal{L}_\chi^{(2)}$ as outlined in Sec. 3.3.3. First in the IR, if we have ekpyrosis, then $|\mathcal{L}_\zeta^{(2)}| \sim k_{\rm p}^4/M_{\rm Pl}^2$ [using

[21] The points where $\dot{H} = 0$ can be found in terms of $\bar{\varphi}$ by finding the roots of $U_{,\varphi\varphi}$ [recall (3.9)]. For the potential (3.89), those points can be expressed in terms of the Lambert W function as

$$\bar{\varphi} = \pm\frac{m}{2}\sqrt{-1 + 2W\left(\frac{1}{2}\exp\left[\frac{1}{2} + \frac{2\varphi_\infty^2}{m^2}\right]\right)}\,.$$

Then, the value of H at those points is found through (3.7d), and we assume $\varphi_\infty \gg m$.

[22] This also shows that a minimal requirement for a successful bounce with $0 < \Upsilon < \infty$ is generally that $0 < m < \sqrt{3/2}\,\mu^2 M_{\rm Pl}$; see [98, 99].

(3.39) and (3.40) in the large $\bar{\epsilon}$ limit, as well as (3.100) and (3.101); also recall $k_{\rm p} \equiv k/a$ is the physical momentum]. If we have kination (ψ is massless), then using (3.56)

$$|\mathcal{L}_\zeta^{(2,\rm IR)}| \sim \frac{k_{\rm p}^3 H^2}{M_{\rm Pl}^2 |H_{\rm b-}|} \,. \tag{3.107}$$

And in either case in the IR using (3.103) we have

$$|\mathcal{L}_\chi^{(2,\rm IR)}| \sim \frac{k_{\rm p}^2 H^2}{M_{\rm Pl}^2} \,. \tag{3.108}$$

In all cases, this confirms that $|\mathcal{L}^{(2)}| \ll |\mathcal{L}^{(0)}| \sim H^2$ before the bounce phase in the IR. In the UV for both ζ and χ (collectively denoted ξ), one finds using (3.102)

$$|\mathcal{L}_\xi^{(2,\rm UV)}| \sim \frac{k_{\rm p}^4}{M_{\rm Pl}^2} \,, \tag{3.109}$$

which upon comparing with $|\mathcal{L}^{(0)}|$ suggests a strong coupling scale of the order of $\Lambda_{\rm strong} \sim \sqrt{|H| M_{\rm Pl}}$. Note that this is, in essence, only a statement about semiclassical GR since we assume that the Cuscuton is negligible before the bounce phase. However, as it is the Cuscuton that sets the maximal Hubble scale $|H_{\rm b-}|$ right at the onset of the bounce phase, this is already informative of the Cuscuton's strong coupling scale; using (3.105), it suggests

$$\Lambda_{\rm strong}^{(\rm Cuscuton)} \sim \sqrt{|H_{\rm b-}| M_{\rm Pl}} \sim \frac{m}{\mu} \sqrt{\frac{\varphi_\infty}{M_{\rm Pl}}} \sim \mu M_{\rm Pl} \sqrt{\frac{\varphi_\infty}{M_{\rm Pl}}} \,, \tag{3.110}$$

where we use $m \approx \mu^2 M_{\rm Pl}$ in the last estimate. Unsurprisingly, the resulting scale is proportional to $\mu M_{\rm Pl} = M_{\rm L}$, which is the new mass scale that the Cuscuton introduced from the start in (3.1). If we set $\varphi_\infty \sim M_{\rm Pl}$ for the sake of simplicity, then the strong coupling scale is solely controlled by $M_{\rm L}$ as $\Lambda_{\rm strong}^{(\rm Cuscuton)} \sim M_{\rm L}$. Generally we want $\mu^2 \ll \varphi_\infty/M_{\rm Pl} < 1$, so assuredly $\Lambda_{\rm strong}^{(\rm Cuscuton)} < M_{\rm Pl}$, but the strong coupling scale is not necessarily orders of magnitude smaller than the Planck scale at this point. However, if $|H_{\rm b-}|$ is sufficiently small compared to $M_{\rm Pl}$, then we have the hierarchy $|H| \leq |H_{\rm b-}| \ll \Lambda_{\rm strong}^{(\rm Cuscuton)} \sim \sqrt{|H_{\rm b-}| M_{\rm Pl}} \ll M_{\rm Pl}$. This is indeed assured whenever $\sqrt{m/M_{\rm Pl}} \lesssim \mu \ll \sqrt{\varphi_\infty/M_{\rm Pl}} < 1$.

Bounce phase, isocurvature perturbations Let us now focus on the bounce phase, starting with χ. We again have $\mathcal{L}_\chi^{(2)} = (F/2)(\dot{\chi}^2 - a^{-2}(\partial_i \chi)^2)$, but now F essentially becomes a constant through the bounce phase [recall the discussion below (3.74)]. As we saw around (3.76), χ_k also remains more or less constant through the bounce phase on superhorizon scales, so one can estimate the size of $\mathcal{L}_\chi^{(2)}$ in the IR from its value at the onset of the bounce. Thus, we infer from (3.108) that $|\mathcal{L}_\chi^{(2,\rm IR)}| \sim (k_{\rm p}/M_{\rm Pl})^2 H_{\rm b-}^2$. Since $|\mathcal{L}^{(0)}| \sim H_{\rm b-}^2$, this confirms $|\mathcal{L}_\chi^{(2,\rm IR)}| \ll |\mathcal{L}^{(0)}|$ through the bounce in the IR. Note that this is true regardless of

whether the prebounce phase was ekpyrosis or kination. Likewise in the UV, $|\mathcal{L}_\chi^{(2)}|$ is the same as at the onset of the bounce, (3.109), and the resulting strong coupling scale is given by (3.110), which remains constant through the bounce.

Bounce phase, curvature perturbations Moving on to ζ through the bounce, let us start with the analysis on subhorizon scales. Since the action for ζ in the UV reduces to the usual 'canonical' form [recall (3.29)], one finds the same result for ζ as for χ regarding the UV strong coupling scale derived above. One may be concerned about the point where $H = 0$ [recall the discussion after (3.29)], but by doing the limits carefully, we confirm that the result is unchanged. In fact, taking $H \to 0$ before $k \to \infty$ in (3.25), one finds $z \simeq \sqrt{2/\sigma}\,(k/H)$, so the amplitude of the UV fluctuations at the bounce point is not given by (3.102), but rather by

$$|\zeta^{(\mathrm{UV})}| \overset{H\approx 0}{\sim} \frac{\sqrt{\sigma}\,|H|}{M_{\mathrm{Pl}}}\,, \tag{3.111}$$

where we recall that $\sqrt{\sigma}\,|H|$ is $\mathcal{O}(H^0)$. Combining this and (3.30), which is the UV action about $H = 0$, one recovers (3.109). In passing, we have to recall from (3.83) and (3.84) that $\sigma/\epsilon_\psi = 1 - \epsilon/\epsilon_\psi \approx 1 - M_{\mathrm{Pl}}^2/\widetilde{M}_{\mathrm{Pl}}^2$ is $\mathcal{O}(1)$ about $H \approx 0$.

Away from the large-k limit, the Lagrangian for ζ is generally quite complicated in the bounce phase, so let us explore the opposite regime, i.e., the IR limit given by (3.28), which is relevant for superhorizon modes of observational relevance. To approximate $\mathcal{L}_\zeta^{(2,\mathrm{IR})}$, we make use of the fact that we have previously carried out the series expansion for most of the required background quantities around the bounce point, as presented in (3.68), (3.73), (3.78), and (3.79). Now, let us evaluate the series expansion for the other necessary 'slow-roll' parameters:

$$\epsilon \overset{t\approx 0}{\simeq} -\frac{1}{\Upsilon t^2} + \mathcal{O}(t^0)\,, \qquad \eta \overset{t\approx 0}{\simeq} -\frac{2}{\Upsilon t^2} + \mathcal{O}(t^0)\,, \qquad \eta_\psi \overset{t\approx 0}{\simeq} -\frac{2}{\Upsilon t^2} + \mathcal{O}(t^0)\,. \tag{3.112}$$

Applying this to (3.28), one finds to leading order

$$\begin{aligned}\mathcal{L}_\zeta^{(2,\mathrm{IR})} &\overset{H\approx 0}{\simeq} \frac{3\dot{\psi}_{\mathrm{b}}^2}{2M_{\mathrm{Pl}}^2\Upsilon + \dot{\psi}_{\mathrm{b}}^2}\dot{\zeta}^2 - \frac{4M_{\mathrm{Pl}}^2\Upsilon - \dot{\psi}_{\mathrm{b}}^2}{2M_{\mathrm{Pl}}^2\Upsilon + \dot{\psi}_{\mathrm{b}}^2}\frac{(\partial_i\zeta)^2}{a^2} \\ &\approx \left(3 - \frac{2m^2}{\mu^4 M_{\mathrm{Pl}}^2}\right)\left(\dot{\zeta}^2 - \frac{2m^2 - \mu^4 M_{\mathrm{Pl}}^2}{3\mu^4 M_{\mathrm{Pl}}^2 - 2m^2}\frac{(\partial_i\zeta)^2}{a^2}\right) \approx \dot{\zeta}^2 - \frac{(\partial_i\zeta)^2}{a^2}\,,\end{aligned} \tag{3.113}$$

where the first approximation is for IR perturbations about $t \approx 0$ (equivalently $H \approx 0$), the second approximation assumes the potential (3.89) with $m/M_{\mathrm{Pl}} \leq \mu^2 \ll \varphi_\infty/M_{\mathrm{Pl}}$ [and uses (3.74) and (3.82)], and the last approximation is when $m \approx \mu^2 M_{\mathrm{Pl}}$. Note that (3.113) makes it clear why $m \approx \mu^2 M_{\mathrm{Pl}}$ is a sensible assumption since the Lagrangian becomes canonical with unit sound speed; indeed, [99] showed that the theory behaves best when $m \approx \mu^2 M_{\mathrm{Pl}}$,

that is, the sound speed remains closest to unity across time and k-space. In [99], it was also observed that to maintain a positive coefficient for the gradient term (i.e., the sound speed[23] squared) and to ensure that the sound speed remains subluminal, the condition

$$\frac{1}{2} < \frac{m^2}{\mu^2 M_{\rm Pl}^2} \leq 1 \tag{3.114}$$

must be satisfied. Therefore, [99] anticipated that similar constraints should likely be applied to prevent strong coupling issues.

From (3.113), if the superhorizon near-scale-invariant χ perturbations are converted to ζ perturbations just before the onset of the bounce, then as was the case for χ before, we conclude that $|\mathcal{L}_\zeta^{(2,\rm IR)}| \sim (k_{\rm p}/M_{\rm Pl})^2 H_{\rm b-}^2$ on superhorizon scales through the bounce. If the conversion occurs after the bounce one finds on superhorizon scales through the bounce $|\mathcal{L}_\zeta^{(2,\rm IR)}| \sim k_{\rm p}^4/M_{\rm Pl}^2$ or $|\mathcal{L}_\zeta^{(2,\rm IR)}| \sim k_{\rm p}^3 |H_{\rm b-}|/M_{\rm Pl}^2$ depending on whether the prebounce phase is ekpyrosis or kination, respectively. In all cases, $|\mathcal{L}_\zeta^{(2,\rm IR)}| \ll |\mathcal{L}^{(0)}| \sim H_{\rm b-}^2$ for $k_{\rm p} \ll |H_{\rm b-}|$.

Interactions

Moving on to evaluating $\mathcal{L}^{(3)}/\mathcal{L}^{(2)}$, we first notice from the derivation of the third-order action in Sec. 3.3.2 that $\mathcal{L}^{(3)}$ separates into terms that represent $\zeta\chi^2$ interactions [what we called $\mathcal{L}^{(3)}_{(\chi)}$ in (3.93); equivalently $\mathcal{L}_{\zeta\chi\chi}$, $\mathcal{L}_{\zeta\chi^2}$ or $\mathcal{L}_{\chi\chi\zeta}$ henceforth] and terms that represent ζ^3 interactions [which corresponds to $\mathcal{L}^{(3)}_{(\delta\varphi)} + \mathcal{L}^{(3)}_{(\rm rest)}$ of (3.94) and (3.96); $\mathcal{L}_{\zeta^3}$ or $\mathcal{L}_{\zeta\zeta\zeta}$ henceforth]. Therefore, there are two meaningful ratios to evaluate: the ratio of the most dominant terms in $\mathcal{L}_{\zeta\chi\chi}$ over the most dominant terms in $\mathcal{L}^{(2)}_\chi$ ($\mathcal{L}_{\chi\chi}$ henceforth); and the ratio of the most dominant terms in $\mathcal{L}_{\zeta\zeta\zeta}$ over the most dominant terms in $\mathcal{L}^{(2)}_\zeta$ ($\mathcal{L}_{\zeta\zeta}$ henceforth).

Before the bounce Before the bounce, the Cuscuton is negligible, so from (3.93) we have as an order-of-magnitude estimate of the size of the terms involved in $\mathcal{L}_{\zeta\chi\chi}$ (ignoring signs and numerical constants),

$$\mathcal{L}_{\zeta\chi\chi} \sim \left(\left(1 - \frac{XF_{,X}}{F}\right)\frac{\dot\zeta}{H} + \zeta\right)\mathcal{L}_{\chi\chi} + F\left(\frac{\partial_i\zeta}{aH} - a\partial_i\partial^{-2}\dot\zeta\right)\frac{\partial^i\chi}{a}\dot\chi\,, \tag{3.115}$$

and likewise $\mathcal{L}_{\zeta\zeta\zeta}$ is approximately given by (3.97). In the UV, one can verify that $|\mathcal{L}_{\zeta\zeta\zeta}/\mathcal{L}_{\zeta\zeta}| \sim |\zeta|$ and therefore the strong coupling scale at which $|\mathcal{L}_{\zeta\zeta\zeta}| \sim |\mathcal{L}_{\zeta\zeta}|$ corresponds to the Planck scale (when $|\zeta| \sim 1$). Meanwhile, $|\mathcal{L}_{\zeta\chi\chi}/\mathcal{L}_{\chi\chi}|$ is dominated by terms such as

[23] While a negative sound speed squared would not necessarily be a sign of a gradient instability since this is in the IR as opposed to the UV, it may still be an undesirable feature. We will in fact see later that $m \to 0^+$ could lead to the production of arbitrarily large non-Gaussianities through the bounce.

$\dot\zeta/H \sim k_{\rm p}^2/(|H|M_{\rm Pl})$, so the strong coupling scale, similar to the case of $\mathcal{L}^{(2)}/\mathcal{L}^{(0)}$ discussed after (3.109), is expected to be of the order of $\sqrt{|H|M_{\rm Pl}}$. In the IR, using (3.56) for ζ, we estimate

$$|\zeta| \sim \frac{k_{\rm p}^{3/2}}{M_{\rm Pl}\sqrt{|H_{\rm b-}|}} \ln\left(\frac{|H|}{k_{\rm p}}\right), \qquad |\dot\zeta| \sim \frac{k_{\rm p}^{3/2}|H|}{M_{\rm Pl}\sqrt{|H_{\rm b-}|}}, \qquad \frac{|\partial\zeta|}{a} \sim k_{\rm p}|\zeta|, \tag{3.116}$$

and similarly using (3.63) for χ, (3.103) follows. This implies $|\mathcal{L}_{\zeta\chi\chi}/\mathcal{L}_{\chi\chi}| \sim |\mathcal{L}_{\zeta\zeta\zeta}/\mathcal{L}_{\zeta\zeta}| \sim |\zeta| \ll 1$ as desired.

Through the bounce in the IR The IR limit ($k/a \ll |H|$) of the Cuscuton constraint (3.17) gives

$$\delta\varphi \overset{\rm (IR)}{\simeq} \frac{2\sigma}{\mu^2(\sigma+3)}\dot\zeta\,. \tag{3.117}$$

Upon substituting in (3.94) and (3.96), the dominant terms in $\mathcal{L}_{\zeta\zeta\zeta}$ are of the form $\zeta\dot\zeta^2$ and $\dot\zeta^3$, so one finds

$$\frac{\mathcal{L}_{\zeta\zeta\zeta}}{\mathcal{L}_{\zeta\zeta}} \overset{\rm (IR)}{\simeq} \frac{9H(2\epsilon^2-\epsilon(12+5\epsilon_\psi)+3(6+5\epsilon_\psi+\epsilon_\psi^2))\zeta - 2(2\epsilon^2+9(3+\epsilon_\psi)-\epsilon(15+\eta))\dot\zeta}{6H(3-\epsilon+\epsilon_\psi)^2}\,. \tag{3.118}$$

Near the beginning or end of the bounce phase when $|H| \approx |H_{\rm b-}|$, $\epsilon \approx 0$, and $\sqrt{\epsilon_\psi} \sim \mathcal{O}(1)$, this results in $|\mathcal{L}_{\zeta\zeta\zeta}/\mathcal{L}_{\zeta\zeta}| \sim |\zeta| \sim \sqrt{A_s} \ll 1$, assuming χ converts to curvature perturbations before the bounce. If the conversion occurs after the bounce, it results in

$$\left|\frac{\mathcal{L}_{\zeta\zeta\zeta}}{\mathcal{L}_{\zeta\zeta}}\right| \overset{\rm (IR)}{\sim} |\zeta| \sim \frac{k_{\rm p}^{3/2}}{M_{\rm Pl}\sqrt{|H_{\rm b-}|}} \ln\left(\frac{|H_{\rm b-}|}{k_{\rm p}}\right), \tag{3.119}$$

which is also very small in the IR. In both cases, we make use of the fact that ζ receives marginal growth through the bounce phase in accordance with (3.88) and (3.87), respectively. Furthermore, if entropy perturbations pass through the bounce before being converted, then they also receive marginal growth according to (3.76), and we find $|\mathcal{L}_{\zeta\chi\chi}/\mathcal{L}_{\chi\chi}| \sim |\zeta|$, which is suppressed according to (3.119).

Note that (3.118) is not applicable too close to the bounce point where $H \approx 0$ since, firstly, it appears to diverge in that region and since, secondly, superhorizon IR modes (while still satisfying $k \ll E_{\rm back}$ [see (3.104)]) have crossed into the $k/a \gtrsim |H|$ regime. Therefore, as mentioned before, extra caution is necessary when taking any limits in the vicinity of $H = 0$. Recalling the expansions (3.18) and (3.19) near the bounce point, substituting them in (3.92), using (3.112) and (3.78), and taking an IR limit yields instead

$$\mathcal{L}^{\rm (IR)}_{\zeta\zeta\zeta} \overset{H\approx 0}{\simeq} \frac{9\dot\psi_{\rm b}^2(4M_{\rm Pl}^2\Upsilon+3\dot\psi_{\rm b}^2)}{2(2M_{\rm Pl}^2\Upsilon+\dot\psi_{\rm b}^2)^2}\zeta\dot\zeta^2 \approx \left(\frac{27}{2}-\frac{12m^2}{\mu^4M_{\rm Pl}^2}+\frac{2m^4}{\mu^8M_{\rm Pl}^4}\right)\zeta\dot\zeta^2\,. \tag{3.120}$$

Upon comparing with the dominant term of (3.113), we conclude that

$$\left|\frac{\mathcal{L}_{\zeta\zeta\zeta}}{\mathcal{L}_{\zeta\zeta}}\right| \sim \left(\frac{9}{2} - \frac{m^2}{\mu^4 M_{\rm Pl}^2}\right)|\zeta| \ll 1 \qquad (H \approx 0,\ \text{IR}) \tag{3.121}$$

holds when H passes through zero.

Through the bounce in the UV Similar to the IR, we have to be careful when taking limits through the bounce phase in the UV. For example, as long as $H \neq 0$, the UV limit of the Cuscuton constraint (3.17) gives $\mu^2\delta\varphi \simeq 2\sigma H\zeta$, so the dominant terms in the respective third-order Lagrangians are schematically of the form

$$\mathcal{L}_{\zeta\zeta\zeta} \overset{\text{(UV)}}{\sim} \frac{\epsilon_\psi}{H^2}\nabla^4\zeta^3\,, \qquad \mathcal{L}_{\zeta\chi\chi} \overset{\text{(UV)}}{\sim} \frac{F}{|H|}\nabla^3(\zeta\chi^2)\,, \tag{3.122}$$

where we recall ∇ indicates either a spatial or time derivative here (as they scale equally in the UV). For $\mathcal{L}_{\zeta\chi^2}$, such terms can be seen in (3.93), and for $\mathcal{L}_{\zeta\zeta\zeta}$ they can be seen in (3.94c) and (3.96), although recall that in this case they may amount to boundary terms in the formal UV limit $k \to \infty$. Since these expressions are invalid in the vicinity of the bounce point (indeed, they appear to diverge as $H \to 0$), they are only applicable at the very beginning (or very end) of the bounce phase, when $\sqrt{\epsilon_\psi}$ is of order unity and $|H| \approx |H_{\rm b-}|$. Correspondingly,

$$\left|\frac{\mathcal{L}_{\zeta\zeta\zeta}}{\mathcal{L}_{\zeta\zeta}}\right| \overset{\text{(UV)}}{\sim} \frac{|\nabla^2\zeta|}{H_{\rm b-}^2} \sim \frac{k_{\rm p}^3}{H_{\rm b-}^2 M_{\rm Pl}}\,, \qquad \left|\frac{\mathcal{L}_{\zeta\chi\chi}}{\mathcal{L}_{\chi\chi}}\right| \overset{\text{(UV)}}{\sim} \frac{|\nabla\zeta|}{|H_{\rm b-}|} \sim \frac{k_{\rm p}^2}{|H_{\rm b-}|M_{\rm Pl}}\,. \tag{3.123}$$

The latter ratio yields the same strong coupling scale as (3.110), but the former ratio points toward a smaller strong coupling scale of the order of

$$\Lambda_{\rm strong}^{\rm (Cuscuton)} \sim (H_{\rm b-}^2 M_{\rm Pl})^{1/3} \sim \left(\frac{m}{\mu M_{\rm Pl}}\right)^{4/3}\left(\frac{\varphi_\infty}{M_{\rm Pl}}\right)^{2/3} M_{\rm Pl} \sim \left(\frac{M_{\rm L}}{M_{\rm Pl}}\right)^{4/3} M_{\rm Pl}\,, \tag{3.124}$$

where the last estimate uses $m \sim \mu^2 M_{\rm Pl}$, $\varphi_\infty \sim M_{\rm Pl}$, and $\mu = M_{\rm L}/M_{\rm Pl}$. While $(H_{\rm b-}^2 M_{\rm Pl})^{1/3} < (|H_{\rm b-}|M_{\rm Pl})^{1/2}$, we still have $|H_{\rm b-}| \ll (H_{\rm b-}^2 M_{\rm Pl})^{1/3}$ provided $|H_{\rm b-}|$ is sufficiently small compared to $M_{\rm Pl}$, so such a strong coupling scale would not necessarily represent an immediate strong coupling issue for the model.

Examining the vicinity of $H \approx 0$ as discussed before [using (3.18), (3.112), and (3.78) in (3.92)], one finds no divergent terms in $\mathcal{L}_{\zeta\zeta\zeta}$ and $\mathcal{L}_{\zeta\chi\chi}$, and a full estimate with the same approximations yields

$$\begin{aligned}
\left|\frac{\mathcal{L}_{\zeta\zeta\zeta}}{\mathcal{L}_{\zeta\zeta}}\right| &\sim \frac{\mu^6 M_{\rm Pl}^3 k_{\rm p}^2}{m\varphi_\infty(2m^2 + 3\mu^4 M_{\rm Pl}^2)\sqrt{3\mu^4 M_{\rm Pl}^2 - 2m^2}}\,, \qquad (H \approx 0,\ \text{UV}) \\
\left|\frac{\mathcal{L}_{\zeta\chi\chi}}{\mathcal{L}_{\chi\chi}}\right| &\sim \frac{\mu^2 M_{\rm Pl} k_{\rm p}^2}{m\varphi_\infty\sqrt{3\mu^4 M_{\rm Pl}^2 - 2m^2}}\,. \qquad (H \approx 0,\ \text{UV})
\end{aligned} \tag{3.125}$$

The corresponding strong coupling scales are

$$\Lambda_{\rm strong}^{\rm (Cuscuton)} \sim \sqrt{\frac{m\varphi_\infty}{\mu^6 M_{\rm Pl}^3}(3\mu^4 M_{\rm Pl}^2 + 2m^2)\sqrt{3\mu^4 M_{\rm Pl}^2 - 2m^2}}\,,$$
$$\Lambda_{\rm strong}^{\rm (Cuscuton)} \sim \sqrt{\frac{m\varphi_\infty}{\mu^2 M_{\rm Pl}}\sqrt{3\mu^4 M_{\rm Pl}^2 - 2m^2}}\,, \tag{3.126}$$

and using (3.106), they can be further expressed as

$$\frac{\Lambda_{\rm strong}^{\rm (Cuscuton)}}{\sqrt{\Upsilon}} \sim \sqrt{\frac{3\mu^4 M_{\rm Pl}^2 + 2m^2}{\mu^6 M_{\rm Pl} m^3 \varphi_\infty}}\,(3\mu^4 M_{\rm Pl}^2 - 2m^2)^{3/4}\,,$$
$$\frac{\Lambda_{\rm strong}^{\rm (Cuscuton)}}{\sqrt{\Upsilon}} \sim \sqrt{\frac{M_{\rm Pl}}{\mu^2 m^3 \varphi_\infty}}\,(3\mu^4 M_{\rm Pl}^2 - 2m^2)^{3/4}\,. \tag{3.127}$$

Since $E_{\rm back} \sim \sqrt{\Upsilon}$ when $H \approx 0$, we must ask for the above ratios to be very large for there to be no strong coupling issue. This cannot be achieved in the limit $m^2 \to (3/2)\mu^4 M_{\rm Pl}^2$, but in the same limit no bounce can occur at the background level in the first place — recall either Refs. [98, 99], eq. (3.106) or footnote 22. Therefore, if we instead demand $m^2 \leq \mu^4 M_{\rm Pl}^2$ as before, we find in both cases

$$\frac{\Lambda_{\rm strong}^{\rm (Cuscuton)}}{\sqrt{\Upsilon}} \gtrsim \frac{1}{\mu}\sqrt{\frac{M_{\rm Pl}}{\varphi_\infty}} \gg 1\,, \tag{3.128}$$

hence there is no strong coupling issue provided $\mu \ll \sqrt{\varphi_\infty/M_{\rm Pl}} < 1$ as before. In fact making use of (3.105), the above implies a strong coupling scale of the same order as (3.110) again, which goes as $\Lambda_{\rm strong}^{\rm (Cuscuton)}/|H_{\rm b-}| \sim \sqrt{M_{\rm Pl}/|H_{\rm b-}|}$.

Summary To summarize this subsection, we showed that whenever the Cuscuton parameters satisfy $\sqrt{m/M_{\rm Pl}} \leq \mu \ll \sqrt{\varphi_\infty/M_{\rm Pl}} < 1$, the Cuscuton model remains properly weakly coupled in the IR throughout time, notably through the bounce phase. The same is true in the UV ($k_{\rm p} \gg |H|$): the strong coupling scale is always larger than the background energy scale, as desired ($|H| \ll k_{\rm p} \ll \Lambda_{\rm strong}$). Indeed, in the contracting phase, $\Lambda_{\rm strong} \sim \sqrt{|H|M_{\rm Pl}}$, which by the time the Cuscuton becomes important at the onset of the bounce is $\sqrt{|H_{\rm b-}|M_{\rm Pl}} \sim M_{\rm L}$. Through the bounce, though, the additional interactions the Cuscuton provides lower the strong coupling scale to $\Lambda_{\rm strong}^{\rm (Cuscuton)} \sim (H_{\rm b-}^2 M_{\rm Pl})^{1/3} \sim (M_{\rm L}/M_{\rm Pl})^{4/3} M_{\rm Pl}$. However, as commented above, provided that $|H_{\rm b-}|$ is relatively small compared to $M_{\rm Pl}$, such a strong coupling scale does not imply an immediate strong coupling problem for the model. Additionally, as discussed in the previous subsection, the operators yielding this lower strong coupling scale might be removable by integration by parts perturbatively in $1/k$ in the UV. Nevertheless, this scale could be taken as a conservative lower bound, while $\Lambda_{\rm strong} \sim \sqrt{|H_{\rm b-}|M_{\rm Pl}}$ might be more robust.

3.3.4 Bispectrum calculations

Setup

The strong coupling order-of-magnitude estimate from the previous subsection was rough in the sense that we used the 2-point correlation function to approximate the coarse-grained amplitude of the real-space fluctuations, and we used that to compare coefficients of the second- and third-order Lagrangians. We can improve this estimate by performing an actual computation of the higher correlation functions. Specifically in this subsection, we will be interested in the 3-point function, represented as the bispectrum in Fourier space.

At the same time as being of theoretical relevance for the strong coupling issue, the bispectrum (like other correlation functions in Fourier space) is an observable that we can measure or constrain through cosmological observations (see, e.g., [27, 224]), which serves as additional motivation for its computation. A nonzero 3-point function is synonymous with a deviation from Gaussianity, which if local in real space can parametrized as (we omit the time dependence here)

$$\begin{aligned}\zeta(\boldsymbol{x}) &= \zeta_{\rm G}(\boldsymbol{x}) + f^{\zeta}_{\rm NL}\left(\zeta_{\rm G}(\boldsymbol{x})^2 - \langle\zeta_{\rm G}(\boldsymbol{x})^2\rangle\right) + C^{\zeta}_{\rm NL}\zeta_{\rm G}(\boldsymbol{x})\chi_{\rm G}(\boldsymbol{x}) + \dots\,,\\ \chi(\boldsymbol{x}) &= \chi_{\rm G}(\boldsymbol{x}) + f^{\chi}_{\rm NL}\left(\chi_{\rm G}(\boldsymbol{x})^2 - \langle\chi_{\rm G}(\boldsymbol{x})^2\rangle\right) + C^{\chi}_{\rm NL}\chi_{\rm G}(\boldsymbol{x})\zeta_{\rm G}(\boldsymbol{x}) + \dots\,,\end{aligned} \tag{3.129}$$

where the subscript 'G' indicates that this is the Gaussian part of the full perturbation (assuming no interaction) and where the ellipses denote further (third-order) deviations from Gaussianity (which would start appearing at the level of the four-point function). Note that there are different conventions in the literature, which may differ from (3.129), the more standard one being of the form $\zeta = \zeta_{\rm G} - (3/5)f^{\zeta}_{\rm NL}\zeta^2_{\rm G} - \dots$ (e.g., [30, 225]). In the above, $f^{\zeta}_{\rm NL}$, $f^{\chi}_{\rm NL}$, $C^{\zeta}_{\rm NL}$, and $C^{\chi}_{\rm NL}$ have the interpretation of local non-Gaussianity parameters, but shortly we will consider non-Gaussianities beyond the locality assumption. From (3.129), one can compute the various 3-point functions in Fourier space involving ζ and χ (see Appendix A for a derivation). We define the general bispectrum B and the dimensionless bispectrum $\tilde{B}$ according to

$$\begin{aligned}\langle\xi^I_{\boldsymbol{k}_1}\xi^J_{\boldsymbol{k}_2}\xi^K_{\boldsymbol{k}_3}\rangle &= (2\pi)^3\delta^{(3)}(\boldsymbol{k}_1+\boldsymbol{k}_2+\boldsymbol{k}_3)B_{\xi^I\xi^J\xi^K}(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3)\,,\\ \tilde{B}_{\xi^I\xi^J\xi^K}(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3) &= \frac{k_1^3k_2^3k_3^3}{k_1^3+k_2^3+k_3^3}B_{\xi^I\xi^J\xi^K}(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3)\,,\end{aligned} \tag{3.130}$$

where $\xi^I = (\zeta,\chi)$, $I,J,K \in \{1,2\}$. We then use the calculation of the 3-point functions for local non-Gaussianities to define the general non-Gaussianity parameters according to

$$f^{\zeta}_{\rm NL}(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3) = \frac{B_{\zeta\zeta\zeta}(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3)}{2\left(P_\zeta(k_1)P_\zeta(k_2)+P_\zeta(k_2)P_\zeta(k_3)+P_\zeta(k_1)P_\zeta(k_3)\right)}\,,$$

$$
\begin{aligned}
C^{\zeta}_{\rm NL}(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3) &= \frac{B_{\zeta\zeta\chi}(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3)}{(P_\zeta(k_1)+P_\zeta(k_2))\,P_\chi(k_3)}\,,\\
f^{\chi}_{\rm NL}(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3) &= \frac{B_{\chi\chi\chi}(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3)}{2\,(P_\chi(k_1)P_\chi(k_2)+P_\chi(k_2)P_\chi(k_3)+P_\chi(k_1)P_\chi(k_3))}\,,\\
C^{\chi}_{\rm NL}(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3) &= \frac{B_{\chi\chi\zeta}(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3)}{(P_\chi(k_1)+P_\chi(k_2))\,P_\zeta(k_3)}\,,
\end{aligned}
\tag{3.131}
$$

where

$$
\langle \xi^I_{\boldsymbol{k}}\xi^J_{\boldsymbol{q}}\rangle = (2\pi)^3\delta^{(3)}(\boldsymbol{k}+\boldsymbol{q})\delta^{IJ}P_{\xi^I}(k)\,,\qquad P_{\xi^I}(k)=|\xi^I_k|^2=\frac{2\pi^2}{k^3}\mathcal{P}_{\xi^I}(k)\,,
\tag{3.132}
$$

are the relations between the 2-point function in Fourier space, the power spectrum, and its dimensionless form. Note that $f^{\zeta}_{\rm NL}$ above reduces to the standard local expression, up to a factor of $-3/5$ due to the convention [225].

The standard approach to compute the bispectrum at some time $\tau_{\rm e}$ is to use the in-in formalism with the interaction picture, according to which to leading order (recall our unconventional dimensions for Lagrangian densities)

$$
\langle \xi^I_{\boldsymbol{k}_1}(\tau_{\rm e})\xi^J_{\boldsymbol{k}_2}(\tau_{\rm e})\xi^K_{\boldsymbol{k}_3}(\tau_{\rm e})\rangle = -2M_{\rm Pl}^2{\rm Im}\langle \xi^I_{\boldsymbol{k}_1}(\tau_{\rm e})\xi^J_{\boldsymbol{k}_2}(\tau_{\rm e})\xi^K_{\boldsymbol{k}_3}(\tau_{\rm e})\int_{-\infty(1+i\varepsilon)}^{\tau_{\rm e}}{\rm d}\tau\int{\rm d}^3\boldsymbol{x}\,a^4\mathcal{L}^{(3)}_{\xi^I\xi^J\xi^K}\rangle\,;
\tag{3.133}
$$

see, e.g., [31, 32] for reviews. The notation above makes it explicit that, in order to compute, e.g., $\langle\zeta^3\rangle$, only terms in $\mathcal{L}_{\zeta\zeta\zeta}$ should matter; likewise to compute $\langle\zeta\chi^2\rangle$, we only need $\mathcal{L}_{\zeta\chi^2}$. The underlying reason is that, when expanding the quantum operators into mode functions and raising/lowering operators,

$$
\hat\xi^I(\tau,\boldsymbol{x}) = \int_{\boldsymbol{k}}\hat\xi^I_{\boldsymbol{k}}(\tau)e^{i\boldsymbol{k}\cdot\boldsymbol{x}} = \int_{\boldsymbol{k}}\left(\xi^I_k(\tau)\hat{\mathfrak{a}}^I_{\boldsymbol{k}}e^{i\boldsymbol{k}\cdot\boldsymbol{x}}+\xi^I_k(\tau)^*\hat{\mathfrak{a}}^{I\dagger}_{\boldsymbol{k}}e^{-i\boldsymbol{k}\cdot\boldsymbol{x}}\right)\,,
\tag{3.134}
$$

the only nonvanishing commutators are

$$
[\hat{\mathfrak{a}}^I_{\boldsymbol{k}},\hat{\mathfrak{a}}^{J\dagger}_{\boldsymbol{q}}] = (2\pi)^3\delta^{IJ}\delta^{(3)}(\boldsymbol{k}-\boldsymbol{q})\,.
\tag{3.135}
$$

We temporarily used hats to explicitly denote quantum operators in the last two equations, but we omit them elsewhere.

As our third-order Lagrangian only contains ζ^3 and $\zeta\chi^2$ interactions, the only nonvanishing 3-point correlators will be $\langle\zeta^3\rangle$ and $\langle\zeta\chi^2\rangle$; in other words, $\langle\chi^3\rangle = \langle\zeta^2\chi\rangle = 0$. One could wonder how robust the vanishing of those bispectra is. As long as χ solely enters the action at the perturbation level with no contribution to the background evolution, with a standard kinetic term, and as long as no potential is generated for χ, this seems to be robust, because the perturbation of the χ Lagrangian to nth order will always be of the form $\mathcal{L}^{(n)}_\chi\sim\mathcal{O}(\zeta^{n-2}\chi^2)$. However, the stability of χ to higher-order corrections and whether

a potential term could appear due to, e.g., quantum loop corrections (which are intrinsically linked to the question of stability) is beyond the scope of this work. Conservatively, it appears $B_{\chi\chi\chi} \approx B_{\zeta\zeta\chi} \approx 0$ (hence $f^{\chi}_{\rm NL} \approx C^{\zeta}_{\rm NL} \approx 0$) hold at least to leading order in our approximations, so our focus shall be on evaluating $B_{\zeta\zeta\zeta}$ and $B_{\chi\chi\zeta}$ (and the corresponding $f^{\zeta}_{\rm NL}$ and $C^{\chi}_{\rm NL}$ functions).[24]

One can intuitively understand how computing the remaining nonvanishing bispectra assesses strong coupling by inspecting the local-type non-Gaussianities as (3.129), which with $f^{\chi}_{\rm NL} \approx C^{\zeta}_{\rm NL} \approx 0$ suggest

$$\frac{\zeta}{\zeta_{\rm G}} \sim 1 + f^{\zeta}_{\rm NL}\zeta_{\rm G}\,, \qquad \frac{\chi}{\chi_{\rm G}} \sim 1 + C^{\chi}_{\rm NL}\zeta_{\rm G}\,. \tag{3.136}$$

Therefore, remaining under perturbative control (being weakly coupled in other words) amounts to having $|f^{\zeta}_{\rm NL}\zeta_{\rm G}| \ll 1$ and $|C^{\chi}_{\rm NL}\zeta_{\rm G}| \ll 1$. If we use our definition of $f^{\zeta}_{\rm NL}$ and $C^{\chi}_{\rm NL}$ from (3.131) and evaluate them for $k_1 = k_2 = k_3 \equiv k$, then the ratios we wish to check can be expressed as

$$|f^{\zeta}_{\rm NL}\zeta_{\rm G}| \approx \frac{k^6|B_{\zeta\zeta\zeta}|}{24\pi^4\mathcal{P}_{\zeta}^{3/2}} = \frac{|\tilde{B}_{\zeta\zeta\zeta}|}{8\pi^4\mathcal{P}_{\zeta}^{3/2}}\,, \qquad |C^{\chi}_{\rm NL}\zeta_{\rm G}| \approx \frac{k^6|B_{\chi\chi\zeta}|}{8\pi^4\mathcal{P}_{\chi}\mathcal{P}_{\zeta}^{1/2}} = \frac{3|\tilde{B}_{\chi\chi\zeta}|}{8\pi^4\mathcal{P}_{\chi}\mathcal{P}_{\zeta}^{1/2}}\,, \tag{3.137}$$

where we approximated as before $|\zeta_{\rm G}| \approx \mathcal{P}_{\zeta}^{1/2}$.

Let us proceed with the calculation of $B_{\zeta\chi^2}(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3)$, where the Lagrangian that enters the calculation, $\mathcal{L}_{\zeta\chi^2}$, is given by (3.93). Expanding all the terms in the Lagrangian and using (3.17) to substitute $\delta\varphi$, the bispectrum as a function of the momenta $\boldsymbol{k}_n$, $n \in \{1,2,3\}$, and at time τ becomes

$$\begin{aligned} B_{\zeta\chi^2}(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3,\tau) = -M_{\rm Pl}^2\,{\rm Im}\Big[&\zeta_{k_1}(\tau)\chi_{k_2}(\tau)\chi_{k_3}(\tau) \\ &\times \int_{-\infty(1+i\varepsilon)}^{\tau} {\rm d}\tilde{\tau}\, a(\tilde{\tau})^2 F(\tilde{\tau})\Big(2\mathcal{I}^{\text{1-perm}}_{\zeta\chi^2}(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3, \tilde{\tau})^* + \mathcal{I}^{\text{2-perm}}_{\zeta\chi^2}(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3, \tilde{\tau})^*\Big)\Big]\,, \end{aligned} \tag{3.138}$$

where

$$\begin{aligned} \mathcal{I}^{\text{1-perm}}_{\zeta\chi^2}(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3, \tau) = {}& A_1^{\zeta\chi^2}(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3, \tau)\zeta_{k_1}(\tau)\chi'_{k_2}(\tau)\chi'_{k_3}(\tau) \\ &+ A_2^{\zeta\chi^2}(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3, \tau)\zeta'_{k_1}(\tau)\chi'_{k_2}(\tau)\chi'_{k_3}(\tau) \\ &+ A_3^{\zeta\chi^2}(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3, \tau)\zeta_{k_1}(\tau)\chi_{k_2}(\tau)\chi_{k_3}(\tau) \\ &+ A_4^{\zeta\chi^2}(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3, \tau)\zeta'_{k_1}(\tau)\chi_{k_2}(\tau)\chi_{k_3}(\tau)\,, \end{aligned}$$

[24] Let us stress that all of this applies ignoring the eventual conversion of isocurvature perturbations into curvature perturbations. We will comment on this separately later.

$$\mathcal{I}^{2\text{-perm}}_{\zeta\chi^2}(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3,\tau) = \Big[A_5^{\zeta\chi^2}(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3,\tau)\zeta_{k_1}(\tau)\chi_{k_2}(\tau)\chi'_{k_3}(\tau) + A_6^{\zeta\chi^2}(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3,\tau)\zeta'_{k_1}(\tau)\chi_{k_2}(\tau)\chi'_{k_3}(\tau)\Big] + [\boldsymbol{k}_2 \leftrightarrow \boldsymbol{k}_3]\,. \tag{3.139}$$

In the above, the $\boldsymbol{k}_n$- and τ-dependent coefficients are

$$\begin{aligned}
A_1^{\zeta\chi^2}(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3) &= 3\left(1+\sigma\frac{k_1^2}{\mathcal{K}_1^2}\right), & A_2^{\zeta\chi^2}(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3) &= -\frac{3}{\mathcal{H}}\left(1-\sigma\epsilon_\psi\frac{\mathcal{H}^2}{\mathcal{K}_1^2}\right),\\
A_3^{\zeta\chi^2}(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3) &= \frac{A_1^{\zeta\chi^2}(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3)}{3}\boldsymbol{k}_2\cdot\boldsymbol{k}_3\,, & A_4^{\zeta\chi^2}(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3) &= \frac{A_2^{\zeta\chi^2}(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3)}{3}\boldsymbol{k}_2\cdot\boldsymbol{k}_3\,,\\
A_5^{\zeta\chi^2}(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3) &= \frac{2A_2^{\zeta\chi^2}(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3)}{3}\boldsymbol{k}_1\cdot\boldsymbol{k}_2\,, & A_6^{\zeta\chi^2}(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3) &= \frac{\epsilon_\psi\mathcal{H}A_5^{\zeta\chi^2}(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3)}{k_1^2}\,,
\end{aligned} \tag{3.140}$$

where we define the momenta $\mathcal{K}_n^2 \equiv k_n^2 + (3+\sigma)\epsilon_\psi\mathcal{H}^2$ related to the Fourier transform of the modified Laplace operator defined in (3.17). Note that $\mathcal{K}_n$ is both k_n and τ dependent. In the above, we are assuming $XF_{,X} \approx F$ to simplify some of the coefficients, which is a good approximation taking, e.g., (3.61) as the coupling function when $\delta \ll 1$.

The calculation of $B_{\zeta^3}(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3)$ is found in a similar fashion, where the Lagrangian that enters the in-in formula (3.133) is now $\mathcal{L}_{\zeta^3}$, which corresponds to (3.96) together with (3.94) where $\delta\varphi$ is explicitly replaced by its ζ-dependent expression (3.17). The resulting bispectrum is of the same form as before,

$$B_{\zeta^3}(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3,\tau) = -2M_{\rm Pl}^2\,\mathrm{Im}\left[\zeta_{k_1}(\tau)\zeta_{k_2}(\tau)\zeta_{k_3}(\tau) \times\int_{-\infty(1+i\varepsilon)}^{\tau}\mathrm{d}\tilde{\tau}\,a(\tilde{\tau})^2\Big(6\mathcal{I}^{1\text{-perm}}_{\zeta^3}(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3,\tilde{\tau})^* + 2\mathcal{I}^{3\text{-perm}}_{\zeta^3}(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3,\tilde{\tau})^*\Big)\right], \tag{3.141}$$

with

$$\begin{aligned}
\mathcal{I}^{1\text{-perm}}_{\zeta^3}(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3,\tau) &= A_1^{\zeta^3}(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3,\tau)\zeta_{k_1}(\tau)\zeta_{k_2}(\tau)\zeta_{k_3}(\tau)\\
&\quad + A_2^{\zeta^3}(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3,\tau)\zeta'_{k_1}(\tau)\zeta'_{k_2}(\tau)\zeta'_{k_3}(\tau)\,,\\
\mathcal{I}^{3\text{-perm}}_{\zeta^3}(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3,\tau) &= \Big[A_3^{\zeta^3}(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3,\tau)\zeta_{k_1}(\tau)\zeta_{k_2}(\tau)\zeta'_{k_3}(\tau)\Big] + [\boldsymbol{k}_1\leftrightarrow\boldsymbol{k}_3] + [\boldsymbol{k}_2\leftrightarrow\boldsymbol{k}_3]\\
&\quad + \Big[A_4^{\zeta^3}(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3,\tau)\zeta_{k_1}(\tau)\zeta'_{k_2}(\tau)\zeta'_{k_3}(\tau)\Big] + [\boldsymbol{k}_1\leftrightarrow\boldsymbol{k}_2] + [\boldsymbol{k}_1\leftrightarrow\boldsymbol{k}_3]\,.
\end{aligned} \tag{3.142}$$

While the coefficients A^{ζ^3} are straightforward to compute from the Fourier transform of the Lagrangian, they are extremely long, so we omit them here. (They are also not particularly insightful.) Nevertheless, we use the full coefficients in our calculations below.

Analytical estimates

Before numerically computing the full bispectra outlined above, aiming to minimize approximations, let us examine a few regimes of interest. In these cases, we can obtain some analytical insight by applying analogous applicable approximations.

Before the bounce in the IR Starting with the prebounce phase in the IR regime, we can assume $\sigma \simeq 0$, $\epsilon_\psi \simeq 3$, $\mathcal{H} \simeq 1/(2\tau)$, $a^2 F \simeq 1/(4\pi^2 A_s M_{\rm Pl}^2 \tau^2)$, and use (3.56) and (3.63) to estimate $\tilde{B}_{\zeta\chi^2}$. Keeping only leading-order terms in the IR results in [the $A_3^{\zeta\chi^2}$, $A_4^{\zeta\chi^2}$, and $A_6^{\zeta\chi^2}$ terms turn out to dominate in (3.139)]

$$\tilde{B}_{\zeta\chi^2}(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3, \tau_{\rm b-}) \simeq -\frac{\pi^2 A_s k_1 (k_1^2 \boldsymbol{k}_2 \cdot \boldsymbol{k}_3 + 3k_2^2 \boldsymbol{k}_1 \cdot \boldsymbol{k}_3 + 3k_3^2 \boldsymbol{k}_1 \cdot \boldsymbol{k}_2)}{6M_{\rm Pl}^2 (k_1^3 + k_2^3 + k_3^3)}\,, \tag{3.143}$$

which is $\mathcal{O}(k^2/M_{\rm Pl}^2)$, hence highly suppressed.[25] Indeed, the deviation from Gaussianity can be quantified as in (3.137) by

$$|C_{\rm NL}^{\chi} \zeta_{\rm G}| \simeq \frac{7}{8M_{\rm Pl}} \sqrt{\frac{|H_{\rm b-}| k_{\rm p}}{6\pi}} \Bigg/ \left| \ln\left(\frac{k_{\rm p}}{2|H_{\rm b-}|}\right) \right|\,, \tag{3.144}$$

where IR modes of observational relevance have $k_{\rm p} \ll |H_{\rm b-}| \ll M_{\rm Pl}$, so this is a very small number. Let us stress that this is before any conversion of χ into ζ, hence $C_{\rm NL}^{\chi}$ is not straightforwardly mapped onto observational constraints on $f_{\rm NL}$. Nevertheless, since the correlation $\langle \zeta\chi^2 \rangle$ involves ζ (which is suppressed on cosmological scales [its spectrum is blue]), the contribution to the final $f_{\rm NL}$ will be severely IR suppressed, even though $C_{\rm NL}^{\chi}$ itself can be greater than one.

We can similarly evaluate $\tilde{B}_{\zeta^3}$ in the IR. In the equilateral limit, to leading order the integrand of (3.141) is $9i(-3\tau_{\rm b-}/2)^{1/2}(\gamma_{\rm E} + i\pi/2 + \ln[-k\tau/2])/(4\pi^{3/2} M_{\rm Pl}^3 \tau)$ — the Lagrangian is essentially dominated by its $\zeta\zeta'^2$ term — so we find

$$\tilde{B}_{\zeta^3}(k, \tau_{\rm b-}) \simeq \frac{k^6}{256\pi^2 M_{\rm Pl}^4 \mathcal{H}_{\rm b-}^2} \left(\ln\left[\frac{k}{2|\mathcal{H}_{\rm b-}|}\right] \right)^4\,. \tag{3.145}$$

Following (3.137), we can express the corresponding deviation from Gaussianity as

$$|f_{\rm NL}^{\zeta} \zeta_{\rm G}| \simeq \frac{(3k_{\rm p})^{3/2}}{64 M_{\rm Pl} \sqrt{2\pi^3 |H_{\rm b-}|}} \left| \ln\left(\frac{k_{\rm p}}{2|H_{\rm b-}|}\right) \right|\,, \tag{3.146}$$

which is again highly suppressed. In fact, the right-hand side of the above is the same as $|\zeta_{\rm G}|$ up to an $\mathcal{O}(1)$ number [cf. (3.116)], in agreement with the estimate $|\mathcal{L}_{\zeta^3}/\mathcal{L}_{\zeta^2}| \sim |\zeta|$

[25] This is to be contrasted with the dimensionless bispectrum corresponding to a scale-invariant perturbation (e.g., inflation or the matter bounce scenario), which depends on the angles of the $\boldsymbol{k}_n$ configuration but does not have an overall power-law dependence on the magnitude of the wavenumbers.

from Sec. 3.3.3. This shows that $|\mathcal{L}_{\zeta^3}/\mathcal{L}_{\zeta^2}| \sim |f_{\rm NL}^{\zeta}\zeta_{\rm G}|$, but we note that (3.144) implies $|\mathcal{L}_{\zeta\chi^2}/\mathcal{L}_{\chi^2}| \sim |\zeta| \nsim |C_{\rm NL}^{\chi}\zeta_{\rm G}|$. This simply indicates that the ratio of the third- and second-order Lagrangians can sometimes be a good proxy for non-Gaussianities (or, for that matter, corrections to any other observable due to any nonlinear interactions), but they are not necessarily (formally) the same quantities. The bispectrum and corresponding quantities such as $f_{\rm NL}$ and $C_{\rm NL}$ are observable (at least in principle), while the ratio of the Lagrangians is really only a rough *theoretical* estimate of the strong coupling scale. In particular, there can sometimes be cancellations in the full calculation of the bispectrum.

Before the bounce in the UV Previously, we calculated $\Lambda_{\rm strong}$ in Sec. 3.3.3 by comparing the size of the third- and second-order terms in the action. Here, we show how $\Lambda_{\rm strong}$ can be obtained through calculations of the bispectrum. Indeed, a more robust way to quantify this scale is to determine the scale where the non-Gaussianities become order unity. As mentioned previously, $|f_{\rm NL}^{\zeta}\zeta_{\rm G}|$ and $|C_{\rm NL}^{\chi}\zeta_{\rm G}|$ represent the level of non-Gaussianities. Focusing on $\langle\zeta^3\rangle$, we first compute the B_{ζ^3} bispectrum using (3.141) and (3.142) with all the terms from the third-order action [cf. (3.94) and (3.96)] for the equilateral shape. Assuming $\sigma \simeq 0$, $\epsilon_\psi \simeq 3$, and $\mathcal{H} \simeq 1/(2\tau)$ for the background contraction again and taking a UV limit, where in particular the mode function is approximated by $\zeta_k \simeq \sqrt{\tau_{\rm b-}/(3k\tau)}\,e^{-ik\tau}/(2M_{\rm Pl})$, one finds the integral in (3.141) to be dominated by

$$\frac{k^{5/2}}{2M_{\rm Pl}^3}\sqrt{\frac{-\tau_{\rm b-}}{3}}\int^{\tau_{\rm b-}} {\rm d}\tau\,(-\tau)^{3/2}e^{3ik\tau} = \frac{(-1)^{3/4}\sqrt{-\tau_{\rm b-}}}{54M_{\rm Pl}^3}\Gamma\Big(\frac{5}{2},-3ik\tau_{\rm b-}\Big) \overset{-k\tau_{\rm b-}\gg 1}{\simeq} -\frac{ik^{3/2}\tau_{\rm b-}^2 e^{3ik\tau_{\rm b-}}}{6\sqrt{3}M_{\rm Pl}^3}\,, \tag{3.147}$$

where Γ is the incomplete gamma function here, hence the dimensionless bispectrum by the end of the contracting phase is found to be (also keeping subleading terms)

$$|\tilde{B}_{\zeta^3}(k,\tau_{\rm b-})| \simeq \frac{k^6\tau_{\rm b-}^2}{1296M_{\rm Pl}^4}\left|1-\frac{14}{(k\tau_{\rm b-})^2}\right|\,. \tag{3.148}$$

Correspondingly,

$$|f_{\rm NL}^{\zeta}\zeta_{\rm G}| \simeq \frac{7}{18\sqrt{6}\,\pi}\left|\frac{k_{\rm p}^3}{56H_{\rm b-}^2M_{\rm Pl}}-\frac{k_{\rm p}}{M_{\rm Pl}}\right|\,, \tag{3.149}$$

in agreement with (3.123), suggesting here $|\mathcal{L}_{\zeta^3}/\mathcal{L}_{\zeta^2}| \sim |f_{\rm NL}^{\zeta}\zeta_{\rm G}|$. From (3.149), the strong coupling scale appears to be

$$\Lambda_{\rm strong}^{\rm (Cuscuton)} \simeq 2^{3/2}3^{5/6}\pi^{1/3}\left(H_{\rm b-}^2M_{\rm Pl}\right)^{1/3}\,, \tag{3.150}$$

again in agreement with (3.124), up to the numerical prefactor here that is around 10.35, so about an order of magnitude larger than the estimated $(H_{\rm b-}^2M_{\rm Pl})^{1/3}$. However, recall

that the higher-order operators yielding the above strong coupling scale could in principle be removed by integration by parts perturbatively in $1/k$ in the UV where it matches GR. Accordingly, the subleading term in (3.149) would actually imply a strong coupling scale at the Planck scale.

Performing a similar calculation for $B_{\zeta\chi^2}$ using (3.138), (3.139), and (3.140), together with the same approximations as before, where in particular in the UV now $\chi_k \simeq \sqrt{2A_s/k}(-\pi\tau)e^{-ik\tau}$, we find the integral in (3.138) to be dominated by

$$-\frac{3i\sqrt{-3k^3\tau_{\mathrm{b}-}}}{2M_{\mathrm{Pl}}^3}\int^{\tau_{\mathrm{b}-}} \mathrm{d}\tau\,\sqrt{-\tau}\,e^{3ik\tau} = \frac{(-1)^{3/4}\sqrt{-\tau_{\mathrm{b}-}}}{2M_{\mathrm{Pl}}^3}\Gamma\Big(\frac{3}{2},-3ik\tau_{\mathrm{b}-}\Big) \overset{-k\tau_{\mathrm{b}-}\gg 1}{\simeq} \frac{e^{3ik\tau_{\mathrm{b}-}}}{4\sqrt{3k}\,M_{\mathrm{Pl}}^3}(6k\tau_{\mathrm{b}-}+i+\ldots)\,. \tag{3.151}$$

We keep the subleading term since when computing (3.138), we have to multiply the above integral by $\zeta_k(\tau_{\mathrm{b}-})\chi_k(\tau_{\mathrm{b}-})^2$ — whose sole complex contribution goes as $e^{-3ik\tau_{\mathrm{b}-}}$ — and take the imaginary part of the result, hence only the subleading term in (3.151) survives. In the end, we find

$$|C_{\mathrm{NL}}^{\chi}\zeta_{\mathrm{G}}| \simeq \frac{k_{\mathrm{p}}}{8\sqrt{6}\,\pi M_{\mathrm{Pl}}}\,, \tag{3.152}$$

which becomes strongly coupled only near the Planck scale. We note that the leading-order term that canceled is the one that would otherwise have yielded $|C_{\mathrm{NL}}^{\chi}\zeta_{\mathrm{G}}| \sim k_{\mathrm{p}}^2/(|H_{\mathrm{b}-}|M_{\mathrm{Pl}})$ as initially expected in the discussion below (3.115). Thus, this is another situation where the leading-order terms in our previous order-of-magnitude estimate of the strong coupling scale cancel in the full calculation of the bispectrum, hence $|\mathcal{L}_{\zeta\chi^2}/\mathcal{L}_{\chi^2}| \nsim |C_{\mathrm{NL}}^{\chi}\zeta_{\mathrm{G}}|$.

Through the bounce in the IR We now evaluate the relevant bispectra $B_{\zeta\chi^2}$ and B_{ζ^3} in the IR at the time $\tau_{\mathrm{b}+}$, i.e., once the bounce phase has ended. In the case where all three modes $\boldsymbol{k}_1$, $\boldsymbol{k}_2$, and $\boldsymbol{k}_3$ exit the horizon in the prebounce phase, one can split the conformal-time integrations in (3.138) and (3.141) as $\int^{\tau_{\mathrm{b}-}} + \int_{\tau_{\mathrm{b}-}}^{\tau_{\mathrm{b}+}}$; the first integral was already evaluated in the IR calculation above, so the new contribution comes from the second integral. To get a rough approximation of the integral from $\tau_{\mathrm{b}-}$ to $\tau_{\mathrm{b}+}$, as before, we analytically expand and estimate the background and linear perturbations about the bounce point (i.e., about $\tau = 0$), according to what was derived in Sec. 3.2.2. Starting with the evaluation of $\langle\zeta\chi^2\rangle$, after substitution for those quantities we find

$$\tilde{B}_{\zeta\chi^2}(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3,\tau_{\mathrm{b}+}) \simeq -\frac{\pi^2 A_s k_1}{6M_{\mathrm{Pl}}^2(k_1^3+k_2^3+k_3^3)}\Bigg(-k_1^2\boldsymbol{k}_2\cdot\boldsymbol{k}_3\left[1-\frac{M_{\mathrm{Pl}}^2}{\widetilde{M}_{\mathrm{Pl}}^2}\left(2+\frac{\mu^4\varphi_\infty^2}{2\mathcal{H}_{\mathrm{b}-}^2}\right)\right] + \left(k_2^2\boldsymbol{k}_1\cdot\boldsymbol{k}_3+k_3^2\boldsymbol{k}_1\cdot\boldsymbol{k}_2\right)\left[3+\frac{m^2\varphi_\infty^2}{M_{\mathrm{Pl}}^2\mathcal{H}_{\mathrm{b}-}^2}\right]\Bigg)\,. \tag{3.153}$$

Upon comparing with (3.143), the overall $\mathcal{O}(k^2/M_{\rm Pl}^2)$ dependence before and after the bounce remains the same, but the coefficients of $k_1^2\boldsymbol{k}_2\cdot\boldsymbol{k}_3$ and $k_2^2\boldsymbol{k}_1\cdot\boldsymbol{k}_3+k_3^2\boldsymbol{k}_1\cdot\boldsymbol{k}_2$ now depend on the parameters of the bouncing phase (parameters of the Cuscuton potential), hence the precise shape changes. It is informative to compare the size of non-Gaussianities through $C^{\chi}_{\rm NL}\zeta_{\rm G}$ before and after the bounce as a ratio,

$$\left|\frac{C^{\chi}_{\rm NL}(\tau_{\rm b+})\zeta_{\rm G}(\tau_{\rm b+})}{C^{\chi}_{\rm NL}(\tau_{\rm b-})\zeta_{\rm G}(\tau_{\rm b-})}\right| \simeq \left|1+\frac{15}{7\gamma^2}-\frac{12}{7(3-2\gamma^2)}\right| , \tag{3.154}$$

where we made use of the first approximation in (3.105) and defined $\gamma\equiv m/(\mu^2 M_{\rm Pl})$. We first note that the above ratio diverges as $\gamma\to 0$ or as $\gamma\to\sqrt{3/2}$, indicating the theory becomes strongly coupled (and breaks down) in those limits. However, these correspond to the boundaries of the range over which the bounce phase can occur in the first place [recall again the footnotes 22 and 23, Refs. [98, 99], eq. (3.106), or the discussion below (3.127)]. Now, let us explore this ratio within that range. For $\gamma\approx 1$, the bounce phase does not generate any significant enhancement of non-Gaussianities from χ interactions since the above ratio is ≈ 1.43. In fact, the ratio (3.154) is $\lesssim 10$ for $0.471\lesssim\gamma\lesssim 1.196$, and it is $\lesssim 100$ for $0.147\lesssim\gamma\lesssim 1.221$. Thus, we confirm that the theory is not strongly coupled if $1/2<m/(\mu^2 M_{\rm Pl})\leq 1$; recall the discussion below (3.113). However, note that this is a conservative range as (3.143) and (3.153) indicate $|C^{\chi}_{\rm NL}(\tau_{\rm b+})\zeta_{\rm G}(\tau_{\rm b+})|$ is suppressed in the IR, so γ would have to be very close to 0 or $\sqrt{3/2}$ before it reaches the strong coupling limit.

The same calculation as above for $\langle\zeta^3\rangle$ yields

$$\left|\frac{f^{\zeta}_{\rm NL}(\tau_{\rm b+})\zeta_{\rm G}(\tau_{\rm b+})}{f^{\zeta}_{\rm NL}(\tau_{\rm b-})\zeta_{\rm G}(\tau_{\rm b-})}\right| \simeq 1+\frac{4(36+\gamma^2)}{3(3-2\gamma^2)\,|\ln(k_{\rm p}/|H_{\rm b-}|)|} , \tag{3.155}$$

indicating once again that the bounce phase has little effect on the deviation from Gaussianity in the IR provided γ is not close to $\sqrt{3/2}$. For $\gamma\lesssim 1$, the above ratio is between 1 and ~ 1.5. Once again one finds a divergence (and thus a strong coupling issue) if γ is close to $\sqrt{3/2}$, but then the same discussion as in the previous paragraph applies.

Lastly, if χ converts into ζ instantaneously at the onset of the bounce, so that ζ becomes scale invariant at $\tau_{\rm b-}$, then ζ evolves through the bounce according to (3.85), but with its initial conditions approximately set at $\tau_{\rm b-}$ as (3.53). Using that in the calculation of B_{ζ^3} through the bounce in the IR, we can quantify the enhancement of the non-Gaussianity from $\langle\zeta^3\rangle$ as

$$\left|\frac{f^{\zeta}_{\rm NL}(\tau_{\rm b+})\zeta_{\rm G}(\tau_{\rm b+})}{f^{\zeta}_{\rm NL}(\tau_{\rm b-})\zeta_{\rm G}(\tau_{\rm b-})}\right| \simeq 1+\frac{16(5-\gamma^2)}{(3-2\gamma^2)\,|\ln(k_{\rm p}/|H_{\rm b-}|)|}\sqrt{\frac{6\pi^7 A_s^3 k_{\rm p} M_{\rm Pl}^6}{|H_{\rm b-}|^7}} , \tag{3.156}$$

where we simply assumed $f^{\zeta}_{\rm NL}(\tau_{\rm b-})\zeta_{\rm G}(\tau_{\rm b-})\approx C^{\chi}_{\rm NL}(\tau_{\rm b-})\chi_{\rm G}(\tau_{\rm b-})$. Other than the same apparent strong coupling issue as before if $\gamma\to\sqrt{3/2}$, we note the additional k dependence in

the above enhancement indicating the spectral index of the bispectrum changes after conversion, through the bounce. This dependence implies that the bispectrum is either enhanced or suppressed through the bounce depending on whether $k_{\rm p}/|H_{\rm b-}|$ is greater or smaller than $(|H_{\rm b-}|/M_{\rm Pl})^6$. However, in magnitude, it remains the case that the contribution to $f_{\rm NL}^{\zeta}$ from the bounce phase is suppressed in the IR. The dominant contribution, if there is any, would then come from the conversion process itself, which was ignored in the above.

Numerical results

Let us confirm the above analytical IR estimates of $B_{\zeta\chi^2}$ and B_{ζ^3} by performing full numerical computations of the bispectra. We note that the three $\boldsymbol{k}_n$ modes of the bispectra relevant to observations first evolve from their subhorizon regime, cross the horizon, evolve on superhorizon scales in the prebounce regime, and then through the bounce. Thus, when evaluating the bispectra at late times (after the bounce phase), the integrals involved in computing (3.138) and (3.141) go through all of these regimes. Until the post horizon-crossing regime, our model resembles other contracting cosmologies, for which non-Gaussianities have been studied and shown to be negligible (e.g., [161, 226]) — this is in contrast to inflation, where contributions to the bispectrum mainly come from the horizon-crossing regime. Our numerical computations further confirm that the contribution from the superhorizon regime dominates over the preceding regimes. Thus, we focus on the superhorizon regime during the prebounce phase and then through the bounce phase. We evaluate the bispectrum using the exact numerical solutions presented in Sec. 3.2.3 for both the background and the mode functions $\chi_k(\tau)$ and $\zeta_k(\tau)$.

The numerical results for the strength of nonlinear effects expressed through the quantity $|C_{\rm NL}^{\chi}(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3,\tau)\zeta_{\rm G}|$ — recall (3.131) — are shown in Fig. 3.4. The pivot value of the first wavenumber is set to $k_1 = 10^{-58}M_{\rm Pl}$, and the value of $\zeta_{\rm G}$ is approximated by $\zeta_{\rm G}(\tau;1/k_1) \approx \mathcal{P}_{\zeta}(k_1)^{1/2}$ as before [recall (3.100)]. We initiate the integral of (3.138) deep in the contracting phase, but the precise starting point is unimportant, as the contribution from the lower integration bound is negligible. In the top-left panel of Fig. 3.4, we show the result when the integration is stopped at $\tau = \tau_{\rm b-}$, which is to say that the bounce phase is not included in the computation. The top-right panel comparatively shows the analytical result obtained from (3.143); the top-left and -right panels are nearly identical, thus demonstrating the validity of the many approximations used in deriving (3.143). The bottom-left panel then shows the result when the integration is stopped at $\tau = \tau_{\rm b+}$, which is to say that the bounce phase is included in the computation. In these three panels, note that $|C_{\rm NL}^{\chi}\zeta_{\rm G}|$ is shown on a logarithmic scale, and as such, it indicates that nonlinear effects are strongly suppressed, as expected from the analytical estimates.

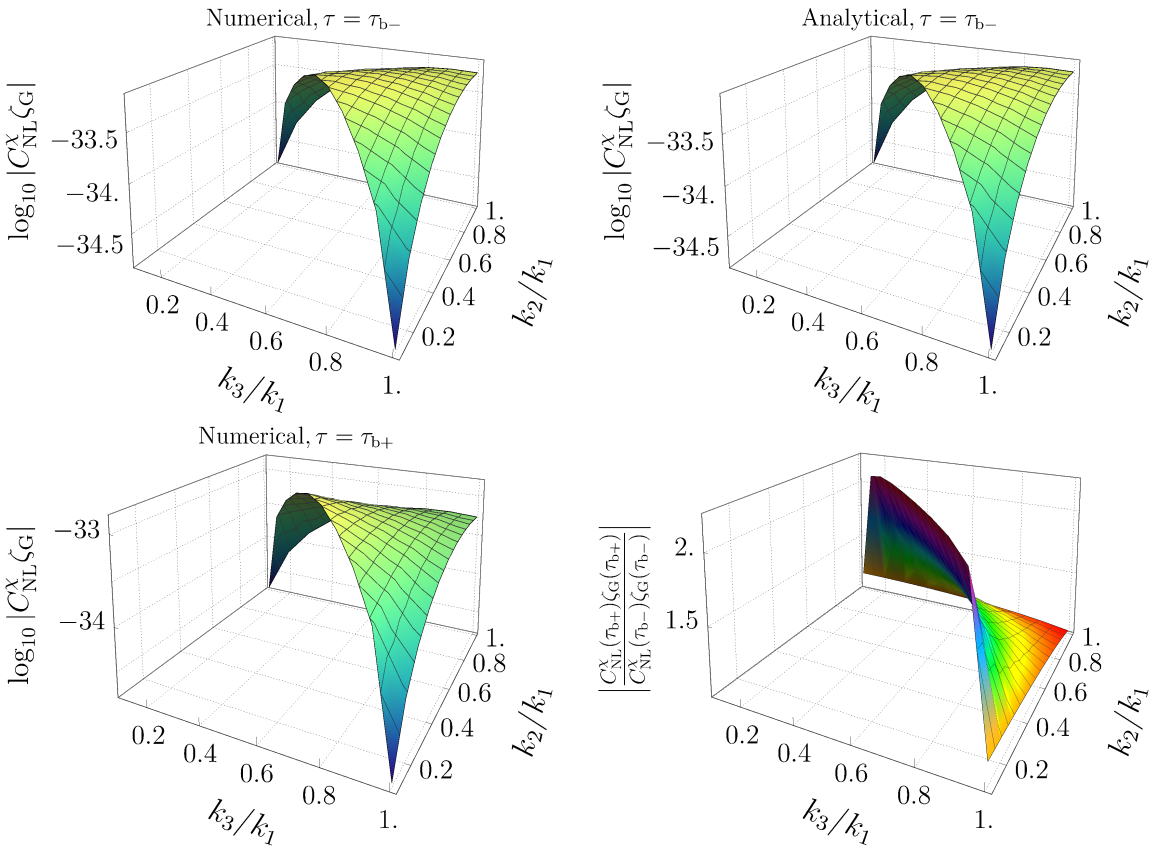


Figure 3.4: Results for the nonlinear corrections from $\zeta\chi^2$ bispectrum. The pivot value of the first wavenumber is set to be $k_1 = 10^{-58}M_{\rm Pl}$, and the value of $\zeta_{\rm G}$ is approximated by $|\zeta_{\rm G}| \approx \mathcal{P}_{\zeta}(k_1)^{1/2}$. The bispectrum is computed numerically in the two panels on the left, ending the integration at $\tau = \tau_{\rm b-}$ in the top-left plot and at $\tau = \tau_{\rm b+}$ in the bottom-left plot. In the top-right panel, the analytical estimate of (3.143) is plotted, which closely matches the corresponding numerical results in the top-left panel. The ratio of $|C^{\chi}_{\rm NL}\zeta_{\rm G}|$ numerical values at $\tau = \tau_{\rm b+}$ and $\tau = \tau_{\rm b-}$ is plotted in the bottom-right panel.

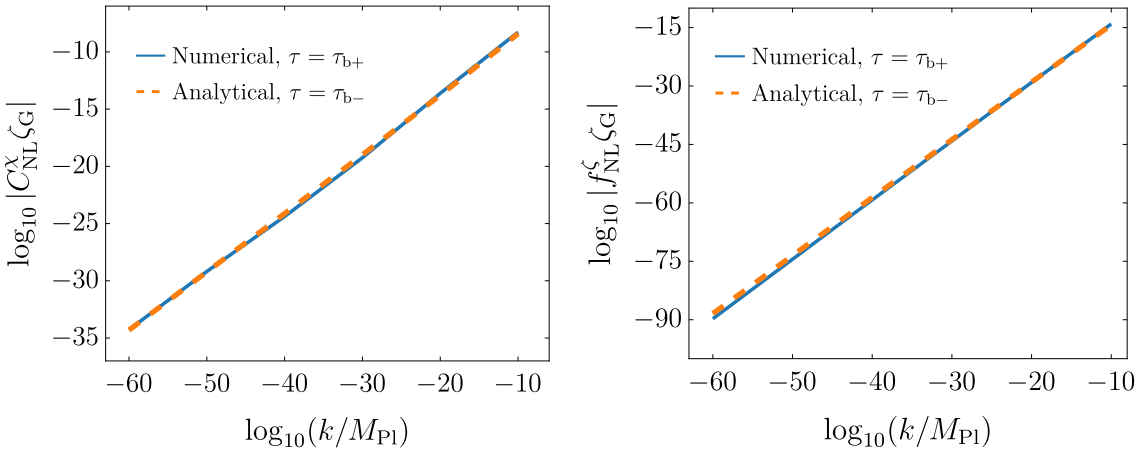


Figure 3.5: Numerical and analytical estimates of $|C^{\chi}_{\rm NL}\zeta_{\rm G}|$ (left) and $|f^{\zeta}_{\rm NL}\zeta_{\rm G}|$ (right) as functions of $k \equiv k_1 = k_2 = k_3$.

In the bottom-right panel of Fig. 3.4, we further plot the ratio of $|C^{\chi}_{\rm NL}\zeta_{\rm G}|$ after the bounce (at $\tau = \tau_{\rm b+}$) to its value before the bounce (at $\tau = \tau_{\rm b-}$). Again, we checked that this closely matches the ratio of the analytical estimates (3.153) and (3.143). The plot further shows that the bounce phase (with reasonable Cuscuton parameters such as those used in the numerical computations) has only a small effect on the size and shape of the bispectrum.

To examine the overall scale dependence of the $\zeta\chi^2$ and ζ^3 bispectra, we compute the quantities $|C^{\chi}_{\rm NL}\zeta_{\rm G}|$ and $|f^{\zeta}_{\rm NL}\zeta_{\rm G}|$ at $\tau_{\rm b+}$ for the simple case of $k_1 = k_2 = k_3 \equiv k$ over a range of k values. The results are respectively plotted in the left and right plots of Fig. 3.5. Additionally, the analytical estimates from (3.144) and (3.146) are included as dashed curves, which correspond to the bispectra evaluated at $\tau_{\rm b-}$. The plots thus show that the bounce has only a marginal effect on the size of non-Gaussianities [as expected from (3.154) and (3.155)] and that the analytical estimates closely match the full numerical computations. In both cases, non-Gaussianities are suppressed in the IR due to the 'blue' k dependence; indeed, ignoring logarithmic dependence, (3.144) and (3.146) imply $|C^{\chi}_{\rm NL}\zeta_{\rm G}| \sim k^{1/2}$ and $|f^{\zeta}_{\rm NL}\zeta_{\rm G}| \sim k^{3/2}$. Note that these power-laws are good approximations (coming from the IR limit) for the whole k range used in Fig. 3.5. In other words, k modes even as high as $10^{-10}M_{\rm Pl}$ safely remain superhorizon (as defined in footnote 5) for the time intervals of interest when computing the bispectra. This can be easily seen in the prebounce phase as the Hubble radius is at its smallest at $\tau_{\rm b-}$ and $|H_{\rm b-}| \sim 10^{-3}M_{\rm Pl}$ in the numerical example used. The same is true post bounce. Through the bounce, the horizon scale differs significantly from the Hubble radius, but (3.71) still holds for these modes.

In the previous plots, we demonstrated that the theory is not strongly coupled and that the perturbations remain highly Gaussian. However, the current observational constraints

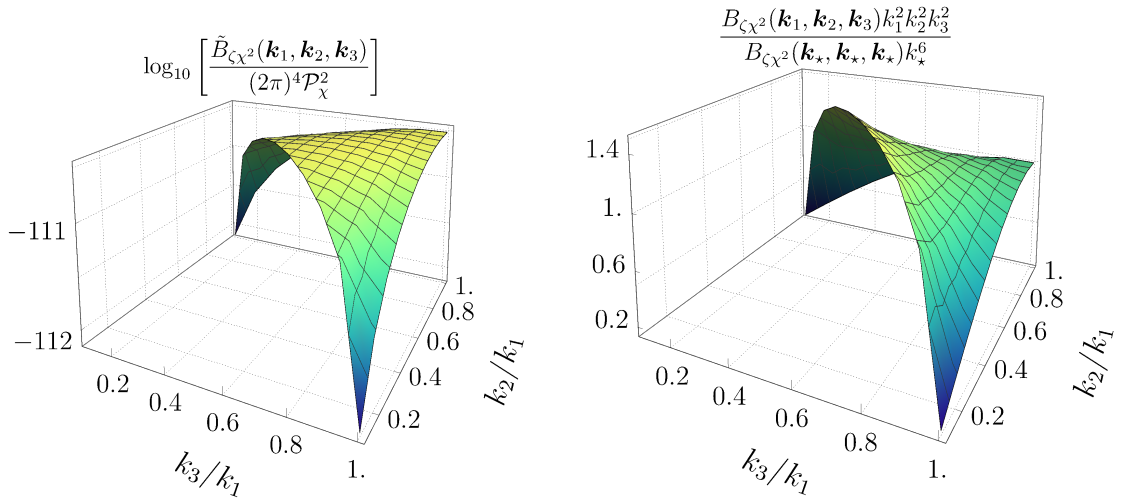


Figure 3.6: The left panel represents the level of non-Gaussianities, assuming that χ converted into ζ, ignoring potential enhancement due to nonlinear interactions needed for this process to happen. The right panel shows the shape function normalized by its value at $k_1 = k_2 = k_3 \equiv k_\star = 10^{-58} M_{\rm Pl}$.

on non-Gaussianities [27] are more stringent than the strong coupling constraints. In principle, if non-Gaussianities are observed in future experiments (e.g., [214, 215]), they can be used to constrain model parameters. However, the quantities shown in Fig. 3.5 cannot be directly mapped to CMB observables. Therefore, for completeness, we also estimate the non-Gaussianity 'amplitude' assuming that χ perturbations directly convert into ζ, ignoring any additional enhancements of non-Gaussianities from this conversion process. That is, we estimate $f_{\rm NL}^{\zeta}$ post conversion as $\tilde{B}_{\zeta\chi^2}(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3)/((2\pi)^4\mathcal{P}_\chi^2)$. This essentially corresponds to our definition of $f_{\rm NL}^{\zeta}$ provided in (3.131), with $B_{\zeta\zeta\zeta}$ replaced by $B_{\zeta\chi\chi}$ and P_ζ by P_χ. The numerical results are presented on a logarithmic scale in the left panel of Fig. 3.6, indicating extremely low values, substantially beneath current observational limits for observationally relevant modes. This agrees with our analytical estimate using (3.143), where we anticipate, for a fixed triangle shape, the overall amplitude to scale as $k^2/(M_{\rm Pl}^2 A_s)$, thereby resulting in significant suppression.

Finally, we end this section by presenting in the right panel of Fig. 3.6 the bispectrum's shape function, $B_{\zeta\chi^2}(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3)k_1^2k_2^2k_3^2$, normalized by its value at $k_1 = k_2 = k_3 \equiv k_\star = 10^{-58} M_{\rm Pl}$. The shape does not resemble any of the well-known shapes (equilateral, folded, or local); therefore, the bispectrum from the Cuscuton bounce, while highly suppressed, has a distinct shape of its own.

3.4 Discussion and conclusions

In this work, we revisited the Cuscuton bounce model of [98, 172], and besides the initial scenario where the matter field is massless, we also expanded our study to include the case where the matter field may be an ekpyrotic field. We provided new insight regarding the evolution of the perturbations through the bounce, reconfirming stability at the linear level, and extending both the numerical and analytical scope of the previous analyses (e.g., [98, 99, 172]).

The main novelty then lies in going beyond the linear level. We derived the full third-order action for the system consisting of the Cuscuton and a massless scalar. As far as we are aware, third-order actions involving the Cuscuton (or related theories) have only been derived in [227] under slow-roll approximations and in a related but different modified gravity based bouncing model in [228]. We used the third-order action with two goals in mind: to check the consistency of the models (i.e., find the range of validity for a perturbative approach and weak coupling) and to see if they produced sizable non-Gaussianities (which satisfy observational constraints or could lead to observable predictions).

Our strong coupling analysis showed that the Cuscuton bounce has a reasonable regime of validity. Two conditions are needed: the second derivative of the Cuscuton potential at the bounce point, $U_{,\varphi\varphi}(\bar{\varphi}=0)$, must not be larger than $M_{\rm L}^4/M_{\rm Pl}^2$ [see (3.114)]; and the mass coefficient of the Cuscuton's 'kinetic' term $M_{\rm L}$ must be much smaller than the Planck scale. Under these conditions, we find the strong coupling scale to be at least of the order $M_{\rm L}^{4/3}/M_{\rm Pl}^{1/3}$ [see (3.124)], while the background energy scale is at most of the order of $M_{\rm L}^2/M_{\rm Pl}$ [see (3.105)]. Such a hierarchy of scales encourages us that the Cuscuton bounce may remain weakly coupled over cosmological scales of relevance, something other nonsingular bouncing cosmologies can hardly realize (see again, e.g., [112, 113, 136]). While we reach this conclusion by comparing the third- and second-order actions and by computing the bispectra, it may be interesting to check its robustness using alternative methods in future work, such as through scattering amplitudes and positivity bounds (see again, e.g., [112, 139]).

Our analysis of the bispectra revealed non-Gaussianities are very suppressed, in contrast to other bouncing models such as [160, 168, 218, 219, 229–231] and notably [228], which has a Cuscuton-like modification to GR that enables the bounce [232] (although the mechanism for generating the perturbations before the bounce is rather different from ours since it is purely adiabatic). In particular, we found barely any enhancement of nonlinearities as perturbations pass through the bounce in the IR, provided the Cuscuton parameters are in the ranges described above (i.e., within the regime of validity of the model) and provided $U_{,\varphi\varphi}(\bar{\varphi}=0)$ is not too close to zero either [see (3.154)], something foreseen in [99].

Although finding no new observable non-Gaussian signature may be a disappointment, it is also a good thing in terms of the validity of the models under study and within the present observational bounds. Let us emphasize that this study focused on multifield models where a near-scale-invariant power spectrum was generated through an isocurvature perturbation.[26] As such, our models are not fully developed yet in the sense that there still needs to be some form of transition (e.g., through a reheating surface after the bounce phase) through which the isocurvature modes are converted into curvature perturbations. This has been analyzed in various other works such as [170, 181, 182, 201, 203–206, 209, 210], so we omitted this in our analysis. Still, from those references, we know that $\mathcal{O}(1-10)$ non-Gaussianities can be generated during the conversion process, already intriguingly close to current observational constraints such as from *Planck* [27]. Therefore, any additional significant non-Gaussianities from the preconversion phase would imply that these bouncing models are ruled out entirely. As we have shown in this work, this is not the case in Cuscuton bounce models since preconversion non-Gaussianities are extremely suppressed, but such $\mathcal{O}(1-10)$ non-Gaussianities from the conversion may be within observational reach by future cosmological surveys.

Another aspect of the Cuscuton bounce that we left aside in this work is primordial gravitational waves. Indeed, we ignored tensorial metric perturbations γ_{ij} from the start. At the level of the power spectrum, the evolution and spectrum of the tensor modes have been explored in [172], and they have been shown to be stable and nonsingular through the bounce phase. Their prebounce power spectrum can be understood quite intuitively. Tensor modes behave in a similar way to adiabatic perturbations (so schematically $|\gamma_{ij}| \sim |\zeta|$). In particular in the models of this work, this means that the tensor 2-point function $\langle\gamma_{ij}^2\rangle \sim \langle\zeta^2\rangle$ has a blue (vacuum) spectrum, which is highly suppressed on cosmologically relevant scales.[27] A similar statement is expected to hold for the 3-point functions, i.e., $\langle\gamma_{ij}^3\rangle \sim \langle\zeta^3\rangle$ and $\langle\gamma_{ij}\chi^2\rangle \sim \langle\zeta\chi^2\rangle$, and as such we would not expect any observable bispectra involving graviton legs. (For an alternative bouncing model that predicts an interesting signal in the tensor 3-point function,

[26] We did not consider purely adiabatic scenarios such as the matter bounce and varying equations of state [151–153, 168, 233, 234]. However, we believe our conclusions regarding the bouncing phase not contributing any significant changes to non-Gaussianities as the perturbations evolve through the bounce would still hold. This may be a good thing for the models that produce sizable non-Gaussianities on their own [137, 161, 168, 226].

[27] Typically, contracting models that produce a near-scale-invariant power spectrum through isocurvature perturbations, like ekpyrotic cosmology, do not forecast a detectable tensor-to-scalar ratio. This contrasts with models based on purely adiabatic perturbations, such as the matter bounce ($\epsilon = 3/2$) for which superhorizon modes grow significantly. The reason has to do with the fact that within contracting models for which $\epsilon \gtrsim 3$, superhorizon perturbations freeze (or at most grow logarithmically), ensuring that the model remains stable against anisotropies. However, as modes freeze out on relatively large physical scales during the contracting phase (in contrast to inflation), their amplitude is also very suppressed (see Sec. 4.1 of [235] for more details).

see [167].) Moreover, it is important to stress that the presence of the Cuscuton does not affect the action/EoM governing tensor perturbations at the linear level (so while the ζ evolution may be 'nontrivial' due to the Cuscuton, that of γ_{ij} would be exactly as in GR [172]), but it could be interesting to explore more carefully what happens when tensor perturbations are included at higher order.

In general, one of the holy grails for any early universe scenario would be to have a model that predicts sizable non-Gaussianities, yet within the observational bounds, at the same time as being distinguishable from other scenarios and detectable in the future. Although the bispectrum does not achieve this for the current Cuscuton bounce models, there could potentially be a distinguishable signature in the trispectrum (see, e.g., [181, 205] for trispectrum computations in ekpyrosis). Exploring this would involve going to next order in perturbations again, which would represent a significant computational challenge; hence we leave it for future work. Still, it would also provide further scrutinization in terms of the robustness and validity of the model to higher-order nonlinearities, and thanks to computational advances, we may be able to have interesting observational constraints on the trispectrum in the future (see, e.g., [236–238]). In passing, we wish to mention that while the overall size of the correlations may be relevant predictions for various models of primordial cosmology, specific features (such as shapes or oscillations in the amplitude) within the correlations could have an even greater constraining power in terms of having model-independent signatures of inflation versus bouncing alternatives (see, e.g., [171, 239–243]). Yet another observable signature of bouncing cosmology that deserves further exploration is the formation of primordial black holes (see, e.g., [244–250]). The Cuscuton bounce may provide an interesting framework for investigating primordial black holes in the context of bouncing cosmology.

Let us end by mentioning that the Cuscuton, which is still not widely studied, has revealed lots of interesting properties in recent years (e.g., [251–257]). While we treated the Cuscuton as an effective field theory as in [96], the interesting question remains whether this theory admits a UV completion, which still eludes us (see [85, 87, 258] for a few potential avenues, in particular the connection with Hořava-Lifshitz gravity [259]). Various directions could be explored in the future, such as those of [260, 261] for k-essence theories or discretization as alluded to in [85, 96]. Certainly, the Cuscuton still has lots of secrets to unveil.

Chapter 4

Large scale structure

In this chapter, we describe how primordial curvature perturbations evolve into large-scale structures observed in the universe today. Specifically, we explore how the complex statistics of dark matter halos and galaxies can be encoded in some bias parameters. However, this process involves numerous physical phenomena, making it difficult to cover all details comprehensively. Therefore, our discussion remains mostly qualitative, supplemented with general and simplified equations. The primary references for this discussion include [15, 262–267], which provides more detailed analyses for the interested reader.

4.1 Linear evolution; Transfer function

To study the post-inflation (or an alternative early phase scenario for generating initial perturbations) evolution of the perturbations, we can still follow the formalism in Section 2.2. However, since the universe has to inevitably heat up and transition into the radiation and matter dominated eras, the content of the universe can no longer be described by just a scalar field. Alternatively, during these follow-up phases, we can use the perfect fluid description (given in Eq (2.3)) to a good approximation[1] and perturb it to derive a new set of EoMs. To do so, we start by defining the density contrast as

$$\begin{aligned} \delta(\boldsymbol{x}, t) &\equiv \frac{\delta\rho(\boldsymbol{x}, t)}{\bar{\rho}(t)} \\ \delta_{\boldsymbol{k}}(t) &\equiv \frac{1}{V_u} \int \delta(\boldsymbol{x}, t) e^{-i\boldsymbol{k}\cdot\boldsymbol{x}} d^3\boldsymbol{x}\,, \end{aligned} \tag{4.1}$$

[1] In fact, the perfect fluid description might fail in some situations and instead, the distribution function in phase space should be used.

where $\delta\rho$ is the perturbations in the density, and $\bar{\rho}$ is the uniform density. The second line gives the Fourier transform of $\delta(\boldsymbol{x}, t)$, where V_u is the volume of a box assuming periodic boundary conditions. The set of δ, $\delta\rho$, and $\bar{\rho}$ could be separately defined for any component such as radiation, dark matter, and baryonic matter. The same quantity could be defined for perturbations in pressure, velocity field, and entropy. It is generally assumed that the initial conditions are isentropic and the evolution is adiabatic i.e. the entropy perturbations are zero in the beginning and remain zero throughout. However, in general, entropy perturbations can be generated both during inflation (or other scenarios) and through non-adiabatic early processes afterward. Although such perturbations may be produced during these early phases, observational constraints, such as those from the CMB, require that they decay sufficiently early to avoid impacting the CMB.

In the context of post-inflation cosmology, it is more convenient to use the Newtonian gauge, as defined in Eq. (2.21), especially since nonlinearities in the matter sector grow in the late universe, while perturbations in the gravitational field remain, to a good approximation, weak and linear on all scales. The primordial power spectrum given in Eq (2.59) was derived in the comoving gauge; however, it is straightforward to represent that quantity in the Newtonian gauge too. The gauge transformation between Φ and ζ is a function of background quantities; therefore, they do not change the scale dependence of the power spectrum. Hence, the power-law parametrization is also valid for Φ, i.e. $\mathcal{P}_\Phi(k) \propto k^{n_s-1}$, where $n_s \sim 1$. However, this relation is only valid immediately after the end of inflation. Then, the goal is to find the evolution of both Φ, and δ in the post-inflation regime. This complex evolution can be encoded in a function that relates the value of Φ at any time to its primordial value as follows

$$\Phi(k, t) = \Phi(k, t_{\rm i}) T(k) \frac{D(t)}{a(t)}\,, \tag{4.2}$$

where t is some time in the matter dominated era, $t_{\rm i}$ is when inflation is ended, and $T(k)$ is called the *Transfer Function*. $D(t)$ is called the growth factor and it can be shown that in the matter-dominated regime and for sub-Hubble modes, $D(t) \propto a(t)$.

In general, analytically deriving the transfer function, even in the linear regime, is highly complex, requiring numerical computations for precise results. However, there exists a characteristic scale that separates the transfer function into two distinct regimes, each exhibiting a different behavior. First, note that the modes that are cosmologically relevant to present time observations are super Hubble at the end of inflation. As the universe transitions into the radiation-dominated phase, the comoving Hubble horizon starts growing. Consequently, all modes gradually re-enter the horizon, with shorter-wavelength modes entering first. The universe then evolves further to become matter-dominated. Hence, there is a distinction between the modes that enter the horizon during the radiation-dominated phase and those

that enter during the matter-dominated phase. Therefore, the characteristic scale, defined as follows, is the mode that enters the horizon precisely at the moment of matter-radiation equality,

$$k_{\rm eq} \equiv \frac{2\pi}{\tau_{\rm eq}} \propto \Omega_{\rm m,0} h^2 \,, \tag{4.3}$$

where τ is the conformal time, $\Omega_{\rm m,0}$ is the density parameter for the matter at today (see below Eq (2.66) for the definition), and $h = H_0/(100 \text{ km/(s Mpc)})$ is the dimensionless Hubble constant.

It can be shown that, in the super Hubble regime, Φ remains almost constant. However, the sub-Hubble evolution depends on whether the universe is radiation or matter dominated. During the cold dark matter dominated phase, since the dominant component is pressureless with[2] $w = 0$, matter perturbations grow at the same rate as the expansion i.e., $\delta \propto a(t)$. Connecting density perturbations, δ, to the gravitational potential, Φ, requires solving the full set of EoMs. However, in the sub-Hubble regime, where gravity is effectively Newtonian, the Poisson equation holds as $k^2\Phi = -4\pi G\bar{\rho}a^2\delta$. Hence, in the matter dominated phase (note $\bar{\rho} \propto a(t)^{-3}$) the gravitational potential remains constant and $T(k) = 1$ for $k \ll k_{\rm eq}$. In contrast, for the modes that are sub-Hubble in the radiation dominated phase, Φ gets suppressed. To explain this, note that perturbations in radiation density can be shown to oscillate with a nearly constant amplitude[3], since the dominant component has non-zero pressure with $w = 1/3$. Therefore, for the duration that modes live in the sub-Hubble and radiation dominated phase, the gravitational potential decays as $\Phi \propto a(t)^{-2}$. Again, this can be seen from the Poisson equation where $\bar{\rho} \propto a(t)^{-4}$. In other words, Φ is suppressed by a factor of $(a_{\rm HC}/a_{\rm eq})^{-2}$ before entering the matter dominated phase and becoming constant again (HC refers to horizon crossing). The former factor can be shown to be equivalent to $(k/k_{\rm eq})^{-2}$, therefore, $T(k) \propto (k/k_{\rm eq})^{-2}$ for $k \gg k_{\rm eq}$. These effects collectively can be cast into a time evolution, parametrized in $D(t)$, and a scale dependence parametrized by the transfer function as

$$T(k) \begin{cases} = 1 & k \ll k_{\rm eq} \\ \propto (k/k_{\rm eq})^{-2} & k \gg k_{\rm eq} \end{cases} . \tag{4.4}$$

In practice, various effects can modify the transfer function, particularly on small scales. These include the presence of baryons, Silk damping in baryons, and damping due to free streaming in collisionless dark matter [268, 269]. Figure 4.1 shows the numerical calculations

[2]Note that w is the equation of state parameter given in Eq (2.7).

[3]To be more precise, if the initial conditions are isentropic, $\delta \propto \ln(a(t))$, and if they are isocurvature perturbations, $\delta \propto$ constant, however, we ignore the logarithmic growth here.

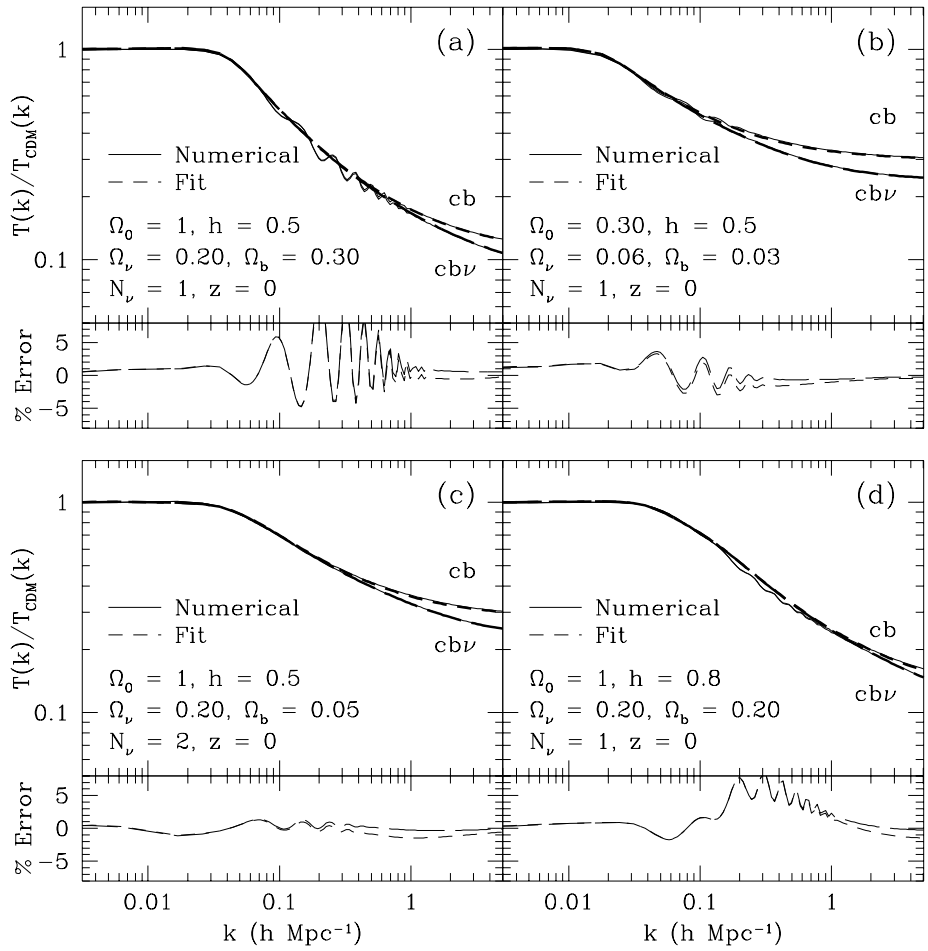


Figure 4.1: The transfer function for different cosmological models including dark matter and either baryons (denoted by cb), or baryons and neutrinos (denoted by cbν) normalized by a model that only contains dark matter. The figure is adopted from [268]. Note that the paper was written in 1997, before the discovery of dark energy, and thus does not include models incorporating it. However, the presence of dark energy does not significantly alter the shape of the transfer function, as its effects become important only at late times. Also, at that time, the constraints on neutrinos and the Hubble constant differed from their present-day values.

for the transfer function using *CMBfast* code [268]. The results are shown for different cosmological models, including dark matter and baryons (denoted by cb), or baryons and neutrinos (denoted by cbν) normalized by a model that contains only dark matter. It can be seen from the top-left panel that a high baryon fraction can cause acoustic oscillations in the transfer function. In addition, it can be shown that in a model consisting solely of baryonic matter (without cold dark matter), the transfer function dies off quickly due to Silk damping in the photon-baryon fluid. Such a suppressed transfer function contradicts observations, providing indirect evidence for the existence of dark matter.

As mentioned before, in the Newtonian regime, the density is related to the gravitational potential through the Poisson equation. In this regime, the density contrast itself, to the linear order, can be solved by the following differential equation

$$\ddot{\delta}_{\boldsymbol{k}} + 2H\dot{\delta}_{\boldsymbol{k}} = (4\pi G\bar{\rho} - \frac{k^2 c_s^2}{a^2})\delta_{\boldsymbol{k}}\,, \tag{4.5}$$

where $c_s = \sqrt{\partial P/\partial\rho}$ is the sound speed. Here, our goal is to show that there is a characteristic proper scale that separates two sets of solutions to this differential equation. This proper scale, known as the *Jeans Length*, is defined as

$$\lambda_{\rm J} \equiv \frac{2\pi a}{k_{\rm J}} = c_s\sqrt{\frac{\pi}{G\bar{\rho}}}\,. \tag{4.6}$$

Ignoring the Hubble drag force term, $2H\dot{\delta}_{\boldsymbol{k}}$, this differential equation has solutions of plane waves with either real frequency if $k > k_{\rm J}$, or imaginary frequency if $k < k_{\rm J}$. In the former case, the perturbations oscillate, which is a consequence of pressure in the fluid. In the latter case, the scales are large enough that pressure can no longer be effective and it can be shown that the perturbations grow. This is known as the Jeans instability, as the perturbations start growing and become unstable.

For baryonic matter after decoupling [4], the comoving Jeans length is about tens of kpc; therefore, perturbations with wavelength larger than this can grow. However, before decoupling, since baryons and photons are tightly coupled and the pressure term is much stronger for photons, the comoving Jeans length is significantly larger. As a result, the growth of perturbations is stalled for modes of cosmological size, about tens of Mpc. In addition, the existing perturbations in baryons decay away due to Silk damping, leaving baryons in the post-decoupling era without significant perturbations on those scales. However, since dark matter is effectively pressureless, its perturbations remain unaffected by Silk damping. After decoupling, baryons can then fall into potential wells associated with dark matter perturbations. This provides yet another indirect proof of the existence of dark matter.

[4] Decoupling refers to when CMB photons are decoupled from electrons.

4.2 Non-linear evolution; Halo and galaxy biases

The previous description of the evolution of perturbations through the transfer function is based on the assumption of linearity, which remains valid only as long as $\delta \ll 1$. However, there are regions in the universe with $\delta \geq 1$, therefore, a complete description of large scale structure formation requires solutions beyond linear ones. A full analytical solution to the non-linear system is almost impossible; hence, semi-analytical models are proposed, which can be tested and verified by N-body simulations. The very first models proposed for this nonlinear evolution are based on the spherical collapse model, the *Press–Schechter* (PS) formalism [270], and its extension, the excursion set theory [271]. Here, we briefly introduce each method.

Let us assume a spherically symmetric shell of matter with a total mass M enclosed in a sphere with radius r. In a universe without a cosmological constant, we have $\ddot{r} = -GM/r^2$, which is essentially equating the radial acceleration to the gravitational force. If M is assumed to be constant, the previous equation can be integrated, which gives the following constant of integration

$$E = \frac{1}{2}\dot{r}^2 - \frac{GM}{r}\,. \tag{4.7}$$

This can be thought of as the total energy of the system. If E is positive, the system expands forever, but if E is negative, it expands for some time and recollapses eventually, depending on the initial conditions. The initial conditions can be derived assuming that the region inherits its initial overdensity from the cosmological overdensity parameter, δ. Then, the system can be evolved and it can be shown that at some time, the radius becomes zero. In practice, the radius never becomes zero since, on small sizes, the spherical assumption breaks down, and the system creates a virialized object that has a stationary radius. These objects are called *dark matter Halo*, since the matter is mostly made of dark matter. Nevertheless, this collapsing limit provides a critical density contrast, $\delta_{\rm c}$, above which any region can be assumed to have formed a halo. A nice feature of the spherical collapse model is that it predicts an almost cosmology and time independent value of $\delta_{\rm c} \simeq 1.69$.

The spherical collapse model can help identify halos, however, it is limited and cannot predict the statistics of halos, such as their mass distribution and their clustering. Therefore, PS formalism builds on that and provides a semi-analytical model to evaluate the mass function of halos, $n(M)$, i.e., the number density of halos as a function of mass. Next, we briefly review the PS formalism. First, let us define a few new parameters. The overdensity field in the linear regime evolves as $\delta(\boldsymbol{x}, t) = \delta_0(\boldsymbol{x})D(t)$ where $\delta_0(\boldsymbol{x})$ is the overdensity field linearly extrapolated to the present time and $D(t)$, as defined before, is the growth factor.

Then the condition of $\delta(\boldsymbol{x},t) > \delta_{\rm c}$ translates into $\delta_0(\boldsymbol{x}) > \delta_{\rm c}/D(t) \equiv \delta_{\rm c}(t)$. Next, the smoothed density field on the scale of R is defined as

$$\delta_{\rm s}(\boldsymbol{x};R) \equiv \int \delta_0(\boldsymbol{x}')W(\boldsymbol{x}+\boldsymbol{x}';R)d^3\boldsymbol{x}'\,, \tag{4.8}$$

where $W(\boldsymbol{x};R)$ is a window function with characteristic radius of R. Each R corresponds to an M which, for a top-hat window function, evaluates to $M = 4\pi\bar{\rho}R^3/3$. Now, the PS formalism is based on two assumptions

- The fraction of mass residing in collapsed objects with mass M or greater, $F(>M)$, is the same as the probability that $\delta_{\rm s} > \delta_{\rm c}(t)$, in other words

$$F(>M) = \mathscr{P}(>\delta_{\rm c}(t)) = \int_{\delta_{\rm c}(t)}^{\infty} \mathscr{P}(\delta_{\rm s})d\delta_{\rm s}\,, \tag{4.9}$$

  where $\mathscr{P}$ represents probability.

- $\delta_{\rm s}$ follows a Gaussian distribution where the width is given by

$$S(R) \equiv \sigma^2(M) = \langle\delta_{\rm s}^2(\boldsymbol{x};R)\rangle = \frac{1}{2\pi^2}\int_0^{\infty} P(k)\tilde{W}^2(\boldsymbol{k}R)k^2dk\,, \tag{4.10}$$

  where $P(k)$ is the power spectrum and $\tilde{W}(\boldsymbol{k}R)$ is the Fourier transform of $W(\boldsymbol{x};R)$.

However, there is a general problem with the PS formalism. It can only count for overdense regions, while initially half of the matter resides in underdense regions. In reality, the matter in underdense regions can eventually become a part of a larger halo. This is known as the cloud-in-cloud problem. To resolve this issue, a fudge of 2 is inserted in Eq (4.9) i.e. $F(>M) = 2\mathscr{P}(>\delta_{\rm c}(t))$. Finally, the number density of halos with masses in the range $[M, M+dM]$ can be derived as

$$\begin{aligned} n(M,t)\,{\rm d}M &= \frac{\bar{\rho}}{M}\frac{\partial F(>M)}{\partial M}\,{\rm d}M = 2\frac{\bar{\rho}}{M}\frac{\partial\mathscr{P}[>\delta_{\rm c}(t)]}{\partial\sigma}\left|\frac{{\rm d}\sigma}{{\rm d}M}\right|{\rm d}M \\ &= \sqrt{\frac{2}{\pi}}\frac{\bar{\rho}}{M^2}\nu\exp\left(-\nu^2/2\right)\left|\frac{{\rm d}\ln\nu}{{\rm d}\ln M}\right|{\rm d}M\,, \end{aligned} \tag{4.11}$$

where in the last line ν is defined as $\nu(t,M) \equiv \delta_{\rm c}(t)/\sigma(M)$.

The extension to PS formalism, called excursion set theory (EST), relaxes the assumption of R being a constant. The goal is to see how $\delta_{\rm s}$ changes as we increase S (see Eq (4.10)). First, note that S is a decreasing function of R (or M), therefore, when $S = 0$, R is effectively infinite and encloses all the mass of the universe, hence $\delta_{\rm s} = 0$. To describe how $\delta_{\rm s}$ changes as S is increased from zero, a set of trajectories can be assumed. Then for each trajectory,

$\delta_{\rm s}$ behaves like a stochastic process as S is increased. Since $\delta_{\rm s}$ has a Gaussian distribution in terms of R, these trajectories are essentially some Markovian random walks. Then, $F(>M)$ can be estimated by counting the number of trajectories that do the first up-crossing from the $\delta_{\rm s} = \delta_{\rm c}(t)$ line at a particular $S(M)$. Mathematically, EST is similar to PS formalism, with the advantage of naturally calculating the fudge factor of 2.

A strength of EST is that it allows for statistically describing the merger history of halos. Let us assume a spherical region with mass M_0 (corresponding to $S_0 = \sigma^2(M_0)$) has the overdensity of $\delta_{\rm c}(t_0)$, hence, it formed a collapsed object at t_0. Now, the question is what fraction of M_0 was already collapsed in other objects at $t_1 < t_0$. This translates to the conditional mass function of $n(M_1, t_1|M_0, t_0)$ and can be easily calculated by the EST, only the origin of S-$\delta_{\rm s}$ plane should be moved by $\delta_{\rm c}(t_1) - \delta_{\rm c}(t_0)$ in the y direction and by $S_1 - S_0$ in the x direction. Then, it can be shown that the number of progenitors with masses in $[M_1, M_1 + dM_1]$ at t_1 that later merged to form an object of mass M_0 at t_0 is given by

$$n(M_1, t_1|M_0, t_0)dM_1 = \sqrt{\frac{2}{\pi}}\frac{M_0}{M_1^2}\nu_{10}\exp\left(-\nu_{10}^2/2\right)\left|\frac{{\rm d}\ln\nu_{10}}{{\rm d}\ln M}\right|{\rm d}M_1\,, \tag{4.12}$$

where $\nu_{10} = (\delta_{\rm c}(t_1) - \delta_{\rm c}(t_0))/\sqrt{S_1 - S_0}$.

The formalism for finding progenitors of halos allows for computing a parameter called *halo bias*, b_h. First, note that the spherical region with M_0 does not have to be a collapsed region. The formalism is still applicable if we think of M_0 as a region with size $R_0 = (3M_0/4\pi\bar{\rho}_0)^{1/3}$, and with linear overdensity δ_0. Consequently, Eq (4.12) can be interpreted as the number of halos existing in a region with linear overdensity $\delta_0 = \delta(t)/D(t)$. Finally, by defining the overdensity of halos in that region as

$$\delta_h(t, M) \equiv \frac{n_h(t, M)}{\bar{n}_h(t, M)} - 1\,, \tag{4.13}$$

where $n_h(t, M)$ is the number of halos with mass M, and $\bar{n}_h(t, M)$ is its spatial average, the EST formalism above leads to a linear relation between δ_h, and δ as

$$\delta_h(t, M) = b_h(t, M)\delta(t)\,, \tag{4.14}$$

where the bias factor is given by [5]

$$b_h(t, M) = 1 + \frac{\nu(t, M)^2 - 1}{\delta_{\rm c}(t)}\,. \tag{4.15}$$

This defines a characteristic mass scale where $\nu = 1$ or $\sigma(M^*) = \delta_{\rm c}(t)$. Therefore, dark matter halos with $M > M^*$ ($M < M^*$) are over-biased (under-biased) compared to the linear overdensity.

[5] See section 7.4 of [262] for a detailed derivation.

The same quantity can be defined for galaxies which is known as *galaxy bias* (see [266] for a detailed review), b_g (replace the lower index of h by g in Eq (4.14), and Eq (4.15)). It is generally accepted that galaxies reside in dark matter halos. Therefore, the galaxy bias can, in principle, be derived by a convolution of halo bias and $N_g(M)$ which is the number of galaxies per halo. However, obtaining $N_g(M)$ is complicated. The formation of galaxies involves various non-gravitational effects which makes it challenging to derive the galaxy bias from first principles. In addition, galaxies have different types, each with its own formation mechanism which results in a different bias. For instance, brighter galaxies generally tend to form in more massive halos which are more strongly clustered and more over-biased. Furthermore, each halo might be occupied by multiple galaxies. Nevertheless, on very large scales, where substructures of halo-galaxy could be ignored, the galaxy bias approaches the halo bias of the host halos in which galaxies reside.

In this chapter, we briefly described how the density field in the universe can be tracked from the post-inflation era to the present time. The linear evolution is encoded in the transfer function, and the nonlinear evolution is modeled by PS and EST formalisms. Then, the halo and galaxy biases are introduced as a proxy to relate the nonlinear evolution to the linearly evolved density field. Various modifications can be added to this picture, such as an ellipsoidal collapse model, a non-Gaussian initial density field, and a non-Markovian EST description. Also, the linear bias can be modified to include complex phenomena such as assembly bias, nonlinear bias, and stochastic bias (see [263, 266] for a detailed review). After introducing gravitational waves in Chapter 5, in Chapter 6 we define a new bias parameter for sources of gravitational waves.

Chapter 5

Gravitational waves

In this chapter, we review the fundamental physics of gravitational waves (GWs). We begin by formulating them within the framework of general relativity, following the formalism presented in [23, 272–274]. Next, we discuss the types of sources that can generate GWs in cosmology and how these sources themselves are formed. Finally, we explore observational signals and review the current and future status of GW detections.

5.1 Theory

As explained in Section 2.2, GWs can be thought of as the two dynamical DoFs of General Relativity (gravitons) as a spin-2 field theory. In some simpler cases, such as the weak-field limit, which is close to Minkowski spacetime, they can be described in terms of metric perturbations on a pre-established background, behaving and propagating as classical waves. In this section, we study GWs following the formalism presented in Section 2.2. In vacuum, the scalar and vector perturbations of the metric are not sourced, leaving the tensor perturbations as the only possible propagating DoF. For simplicity, in what follows, we analyze the propagation of GWs, far away from the sources that generated them, where the background can be estimated to be close to Minkowski. In other words, we set $a(t) = 1$ which means $t = \tau$. This updates Eq 2.39 as

$$h_{ij}(t, \boldsymbol{x}) = \int \frac{\mathrm{d}^3\boldsymbol{k}}{(2\pi)^3} \sum_{\lambda=+,\times} \epsilon_{ij}^{(\lambda)} h_{\boldsymbol{k},\lambda}(t) e^{i\boldsymbol{k}.\boldsymbol{x}} \,, \tag{5.1}$$

where the transverse-traceless conditions on $\epsilon_{ij}^{(\lambda)}$ is also modified as

$$k^i \epsilon_{ij}^{(\lambda)} = 0 \,, \qquad \eta^{ij} \epsilon_{ij}^{(\lambda)} = 0 \,, \qquad \eta^{im}\eta^{jn}\epsilon_{ij}^{(\lambda)}\epsilon_{mn}^{(\lambda')} = 2\delta^{\lambda\lambda'} \,, \tag{5.2}$$

where $\eta_{\mu\nu} \equiv \text{diag}(\eta_0, \eta_{ij}) = \text{diag}(-1, 1, 1, 1)$ is the Minkowski metric. The action given in Eq (2.27) is then modified as

$$S^{(2)}_{\text{tensor}} = \frac{M^2_{\text{Pl}}}{8} \int \text{d}t\text{d}^3\boldsymbol{x} \left(\dot{h}_{ij}\dot{h}^{ij} - \partial_l h_{ij} \partial^l h^{ij} \right) , \tag{5.3}$$

which gives the following EoM in vacuum

$$\ddot{h}_{ij} - \partial^l \partial_l h_{ij} = 0 \, . \tag{5.4}$$

It is common to write the EoM as

$$\Box h_{\mu\nu} = 0 \, , \tag{5.5}$$

where $\Box = \partial^\mu \partial_\mu$ is called the d'Alembertian operator and $h_{\mu\nu} \equiv$ block diag$(0, h_{ij})$. This, as expected, is a wave equation, therefore we update the ansatz of Eq (5.1) by performing another Fourier transform in the time domain

$$h_{ij}(t, \boldsymbol{x}) = \int \frac{\text{d}^4\boldsymbol{k}}{(2\pi)^4} \sum_{\lambda=+,\times} \epsilon^{(\lambda)}_{ij} h_\lambda(\boldsymbol{k}_\mu) e^{i\boldsymbol{k}_\mu \boldsymbol{x}^\mu} \, , \tag{5.6}$$

where $\boldsymbol{k}_\mu = (\omega, \boldsymbol{k})$. Plugging this into the EoM, we have

$$\boldsymbol{k}^\mu \boldsymbol{k}_\mu = 0 \tag{5.7}$$

which could be thought of GWs propagating at the speed of light in the vacuum. Without loss of generality, we can assume the wave is traveling in the x^3 positive direction, hence, $\boldsymbol{k}^\mu \equiv (\omega, 0, 0, \omega)$ where ω is the frequency of the propagating wave. Applying this to Eq (5.2) implies

$$\epsilon^{(+)}_{ij} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \qquad \epsilon^{(\times)}_{ij} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \, . \tag{5.8}$$

At this point, it shortens the notation if we combine $\epsilon^{(\lambda)}_{ij}$, and $h_\lambda(\boldsymbol{k}_\mu)$ to define

$$h_{\mu\nu}(t, \boldsymbol{x}) = \int \frac{\text{d}^4\boldsymbol{k}}{(2\pi)^4} e^{i\boldsymbol{k}_\mu \boldsymbol{x}^\mu} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & h_+(\omega) & h_\times(\omega) & 0 \\ 0 & h_\times(\omega) & -h_+(\omega) & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \, , \tag{5.9}$$

where $h_+(\omega)$, and $h_\times(\omega)$ are the two independent functions describing two waves with different polarizations.

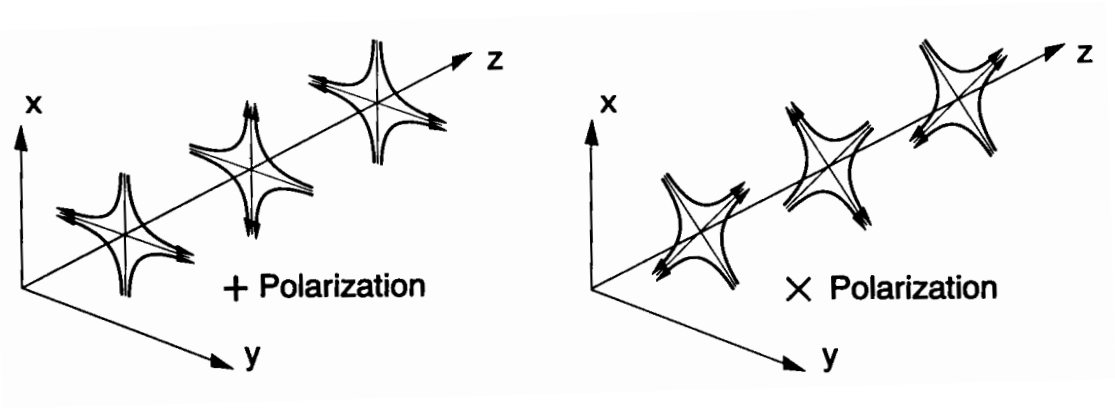


Figure 5.1: A schematic depiction of how test particles move under the passing of GWs with + (×) polarization shown in the left (right) panel. The wave is propagating in the positive z direction. The figure is adapted from [275].

To understand how passing GWs affect test particles, it is not sufficient to derive the geodesic of a single particle. In fact, the worldline of an individual particle can always be made to appear stationary under a suitable coordinate transformation. Therefore, to define a coordinate-independent measure of the GW effect, we should consider the relative motion of nearby particles. This is mathematically formulated in the geodesic congruence (see Section 2.4.1 for the definition) and the geodesic deviation equation. Here, we schematically present the result in Figure 5.1, which illustrates that test particles oscillate in the + (×) directions when a gravitational wave with only + (×) polarization passes by.

So far, we have studied the propagation of GWs in vacuum. However, a source is required for the generation of such waves. We must therefore revisit the calculation, now assuming $T_{\mu\nu} \neq 0$ (the energy-momentum tensor). The presence of $T_{\mu\nu}$ in the Einstein equation implies $h_{\mu\nu}$ cannot be traceless. As a result, the transverse-traceless gauge is no longer valid, and we instead impose the following gauge condition

$$\partial_\mu \bar{h}^{\mu\nu} = 0\,, \tag{5.10}$$

which is known as Lorentz gauge (similar to electromagnetism), and $\bar{h}_{\mu\nu}$, called the trace-reversed perturbation, is defined as

$$\bar{h}_{\mu\nu} \equiv h_{\mu\nu} - \frac{1}{2} h \eta_{\mu\nu}\,, \tag{5.11}$$

where h is the trace and is non-zero in the new gauge[1]. The Einstein equation then reads

$$\Box \bar{h}_{\mu\nu} = -16\pi G T_{\mu\nu}\,. \tag{5.12}$$

[1]Note that far away from any source where the traceless-transverse gauge is valid, we have $\bar{h}_{\mu\nu} = h_{\mu\nu}$.

To solve this, we should use a method analogous to electromagnetism, employing Green's function for delta-function-like sources (see [23, 54] for a detailed derivation). The final result for an isolated source located far from the observer at a distance r (i.e., r is much larger than the size of the source) is

$$\bar{h}_{ij}(t, \boldsymbol{x}) = \frac{2G}{r} \ddot{I}_{ij}(t - r)\,, \tag{5.13}$$

where $I_{ij}(t)$ is the quadruple moment tensor given by

$$I_{ij}(t) = \int x^i x^j T^{00}(t, \boldsymbol{x}) d^3\boldsymbol{x}\,. \tag{5.14}$$

Therefore, any source whose quadrupole moment has a second time derivative can generate GWs. An example is a binary system of stars, i.e. two stars that are orbiting around each other. For two stars, each with mass m, orbiting in a perfect circle with radius R, the GW is given by

$$\bar{h}_{ij}(t, \boldsymbol{x}) = \frac{8Gm\Omega^2 R^2}{r} \begin{pmatrix} -\cos(2\Omega(t-r)) & -\sin(2\Omega(t-r)) & 0 \\ -\sin(2\Omega(t-r)) & \cos(2\Omega(t-r)) & 0 \\ 0 & 0 & 0 \end{pmatrix}, \tag{5.15}$$

where $\Omega = \sqrt{Gm/4R^3}$ is the angular frequency of the binary system rotation, and we assumed the angular momentum is in the positive z direction. Comparing this with Eq (5.9) shows that this configuration sources only one plane wave with $\omega = 2\Omega$.

Waves are always expected to carry energy. However, it is not straightforward to define an energy-momentum tensor for the metric field. In fact, due to the equivalence principle, there is no local definition of energy in general relativity. As a result, the concept of energy carried by GWs presents both philosophical and mathematical challenges. Nevertheless, there are a few methods for a perturbative description of the energy in the weak field limit. This after lengthy calculations (see [23, 54, 276] for more details[2]) yields the following for the energy-momentum tensor of the GW given in Eq (5.9) corresponding to a single plane wave,

$$T^{\rm GW}_{\mu\nu} = \frac{1}{32\pi G}\omega^2(h_+^2 + h_\times^2) \begin{pmatrix} 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 1 \end{pmatrix}. \tag{5.16}$$

[2]Schematically, it can be shown that $T^{\rm GW}_{\mu\nu} \propto \langle \partial_\mu h_{\rho\sigma} \partial_\nu h^{\rho\sigma} \rangle$, where $\langle\rangle$ is an average over several wavelengths. For a single plane wave, this schematically yields $T^{\rm GW}_{\mu\nu} \propto \omega^2 (h_{\mu\nu})^2$.

This gives the following for the emission power[3] (brightness) of GWs given in Eq (5.15)

$$P = \frac{128}{5} G m^2 R^4 \Omega^6 = \frac{2}{5} \frac{G^4 m^5}{R^5} \,, \tag{5.17}$$

where the last equality is evaluated using $\Omega = \sqrt{Gm/4R^3}$.

The binary system studied here is simple with many simplifying assumptions; however, it provides a valuable insight. The power in GWs is proportional to the fifth power of m/R ratio, i.e. the compactness of the binary system. In other words, larger masses orbiting around each other at smaller distances can create stronger GWs. This is also evident from the amplitude of $\bar{h}_{ij}$ in Eq (5.15) which is proportional to $m\Omega^2 R^2/r \propto m^2/(rR)$.

5.2 Sources

In the previous subsection, we showed that any mass distribution which has a non-zero second time derivative of its quadrupole moment could generate GWs. The simplest of this configuration could be two masses orbiting around each other. Hence, in principle, even a bird flying around a tree is a possible source. Based on our current observation ability, which will be explained in Section 5.3, we list the following categories for possible GW sources.

- Transient signals, modeled: GWs from binary systems at their merger time, such as binray black hole mergers — known as compact binary coalescence (CBC) — are the best detectable candidates. In fact, as will be discussed in Section 5.3, these have been observed. Although Eq. (5.15) is an approximation that breaks down in the strong-field regime near the merger, such regimes are well understood using post-Newtonian calculations, numerical relativity, and black hole perturbation theory.

- Transient signals, challenging to model: such as the core collapse of stars, gamma ray bursts, and supernovas. Because these phenomena create highly dynamic environments, they certainly generate time-varying quadrupole moments, but they are also difficult to model. They also rarely happen and their signals are marginally within the sensitivity of current detectors, hence they are not entirely hopeless to be observable in the future.

[3] Roughly speaking, it can be shown that

$$P \propto G \left\langle \frac{d^3 J_{ij}}{dt^3} \frac{d^3 J^{ij}}{dt^3} \right\rangle \,,$$

where $\langle\rangle$ is an average over several wavelengths, and $J_{ij} \equiv I_{ij} - \frac{1}{3}\delta_{ij}\delta^{kl} I_{kl}$ is called the reduced quadrupole moment. This brings the Ω^6 factor here.

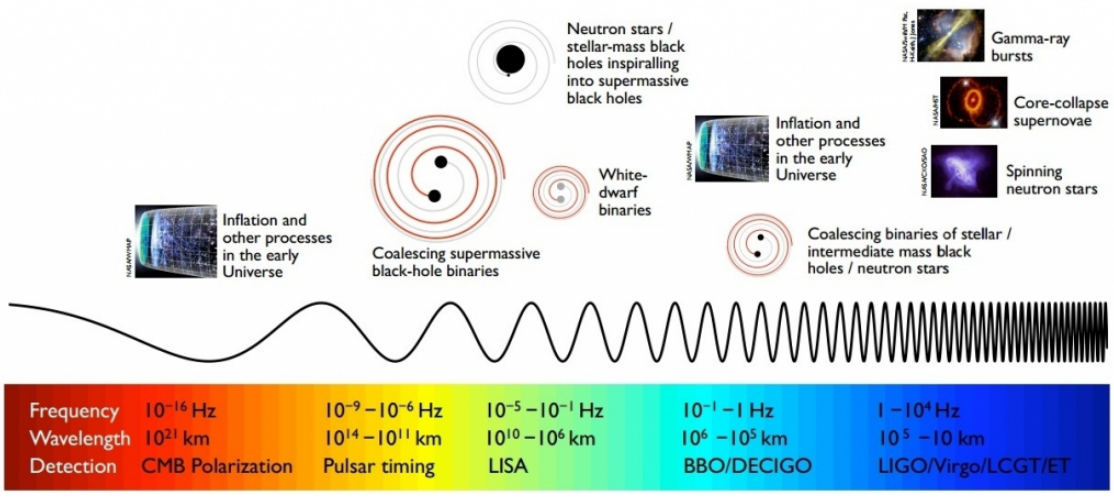


Figure 5.2: All possible GW sources along with their observing tools. The x-axis shows their estimated frequency. The figure is adapted from here.

- Long-duration signals, modeled: Examples are asymmetric neutron stars, such as a neutron star with a mountain [277], that are spinning fast. Their observing prospect is similar to that of supernovas. In addition, there is another exciting opportunity, the Extreme Mass Ratio Inspiral (EMRI) which refers to a system where a compact object (like a stellar-mass black hole or neutron star) spirals around a super massive black hole [278]. They are in principle observable by LISA (discussed in Section 5.3).

- Long-duration signals, challenging to model: We expect many different phenomena to generate GWs during cosmic history, such as inflationary GWs, early universe phase transitions, unresolved binary systems in the spiral phase, and high redshift mergers possibly from primordial black holes. Together, they create a GW background, called stochastic gravitational waves [279]. Since they are from a combination of many different sources and the phase information is lost, they are very challenging to model. Their signals remain below the sensitivity of current and possibly future detectors. However, an indirect method using pulsars to detect them has been developed by the NANOGrav team [280].

Figure 5.2 summarizes all possible sources along with their observing tools. The estimation of their frequency and the observational tools are discussed in Section 5.3. In addition, the possibility of detecting primordial tensor perturbations with CMB polarization was discussed in Section 2.3.1.

As mentioned before, mergers of compact objects generate observable GW signals. The energy carried by GWs is coming from the orbiting objects, hence, they should lose energy,

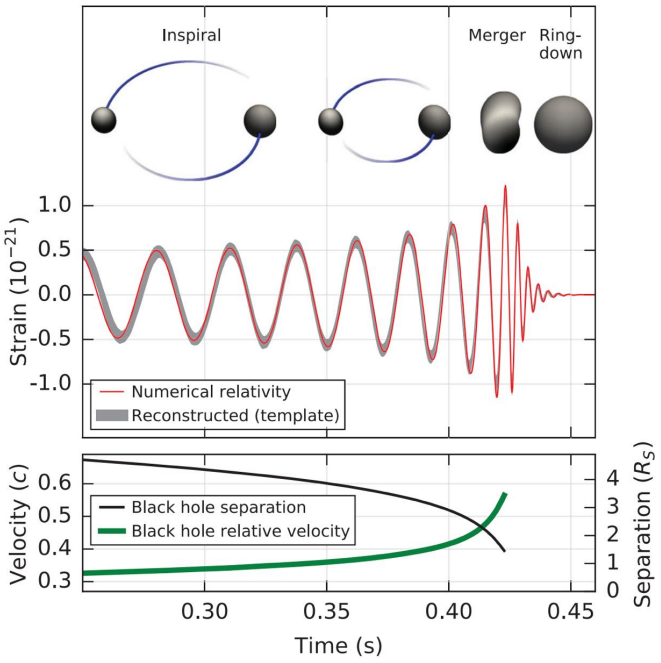


Figure 5.3: The reconstructed waveform for the first GW observation. Here we use it only for illustrative purposes, showing different phases of binary evolution alongside GW emission. Strain represents the amplitude of GWs. The figure is adapted from [281].

reducing their orbital radius. Therefore, as the binary system emits GW, the objects get closer to each other, emitting stronger GW until they finally merge with each other. At this point, the quadrupole moment dies off and no more GW is emitted. This is shown in Figure 5.3 alongside the change in orbital velocity, the separation between the objects, and amplitude(strain)-frequency of emitted GW. The amplitude-frequency evolution of a binary system is called a *waveform*. The evolution of the binary system is categorized into three phases; inspiral, merger, and ring-down[4]. This is also evident from Eq (5.15); as R decreases (at fixed m), the GW amplitude and its frequency increase. However, Eq (5.15) is just an approximation, and to model the true waveform one needs numerical relativity.

The astrophysical candidates for the binary systems from progenitor stellar populations can be categorized into three types: binary neutron stars (BNS), black hole-neutron star binaries (BHNS), and binary black holes (BBH)[5]. In principle, modeling the properties of these systems is a complex task. These properties include their masses, chemical composition (for neutron stars), spin, and orbital geometry. In general, these properties can exhibit different populations, which depend on the type of system, their formation redshift, and

[4]Currently, the ring-down phase is not directly observable, however, in principle, they can have interesting information about Hawking radiation, and quantum gravity [282].

[5]Binary white dwarfs are also possible; however, their signal is very weak.

their host galaxies. The formation of these binaries is a complicated process; they originate from a population of stars and follow different stellar evolution pathways to form the final stellar remnants [283]. These processes depend on the mass and metallicity of the progenitor stars, which, in turn, are functions of the mass, star formation rate, and metallicity of their host galaxies [283–287]. From stellar remnants to the final coalescence of the binary system, various astrophysical processes play a role, including stellar winds, mass transfer, supernova kicks, and common envelope physics. This complexity allows for multiple formation scenarios for these binary systems (see Section 3.3 in [283] and Section 3.1 in [284] for discussions on the formation channels of compact binary objects). However, this intricate relation can be advantageous; observing different populations of binary systems can provide insights into the unknown astrophysical phenomena governing their formation. We discuss their formation channels and explore the connection to the properties of their host galaxies in more detail in Chapters 6, and 7.

5.3 Observation

In order to specify sources with possible observable signals, we do a rough order of magnitude estimate of GW amplitude and its frequency. First, in Eq (5.15), we can assume $R \simeq R_{\rm s}$ where $R_{\rm s} = 2Gm/c^2$ is the Schwarzschild radius. We bring back c (the speed of light) to perform the estimation in SI units. This assumption is reasonable since the strongest GW signal originates from the merger moment when the separation becomes comparable to the size of the objects, which should be the Schwarzschild radius if they are sufficiently massive. Hence, we can approximate the amplitude of $\bar{h}_{ij}$ in Eq (5.15) (let us denote it by h) near the merging moment as $h \sim Gm/(rc^2)$, and its frequency as $\Omega \sim c^3/(Gm)$. Therefore, more massive binaries generate stronger GWs with smaller frequencies.

As discussed around Figure 5.1, a passing GW can change the distance between two test particles which are freely falling with respect to the background metric. If they are separated by the distance L, we have

$$\frac{\delta L}{L} \sim h\,, \tag{5.18}$$

where $\delta L/L$ is called *Strain*. Putting everything together, we have the following for typical solar mass objects at cosmological distances away from earth

$$\left(\frac{\delta L}{\rm cm}\right) \overset{R=R_{\rm s}}{\sim} 10^{-16} \left(\frac{m}{M_\odot}\right) \left(\frac{100{\rm Mpc}}{r}\right) \left(\frac{L}{\rm km}\right)\,, \tag{5.19}$$

which corresponds to $h \sim 10^{-21}$. For $f \equiv \Omega/(2\pi)$, we have

$$\left(\frac{f}{\mathrm{s}^{-1}}\right) \overset{R=R_{\mathrm{s}}}{\sim} 10^4 \left(\frac{M_\odot}{m}\right) \qquad \left(\frac{f}{\mathrm{s}^{-1}}\right) \overset{R=5R_{\mathrm{s}}}{\sim} 10^3 \left(\frac{M_\odot}{m}\right) , \tag{5.20}$$

where the second one is more valid for the inspiral phase (pre-merger phase).

5.3.1 Interferometers

The strain caused by GWs, shown in Eq (5.19), is substantially small. In comparison, the typical size of an atom nucleus is about 10^{-13}cm. We could increase the strain if we design a detector with a greater length, $L \simeq 10$ km. In addition, one more order of magnitude improvement is achievable if we aim for more massive binaries, i.e., $m \simeq 10M_\odot$. It turns out that light could be sensitive to such a tiny change in length using the clever interferometry design shown in Figure 5.4, which is the principal idea behind all GW detectors. A laser beam is split into two paths and travels between two mirrors more than 100 times. The mirrors are mounted to act as free-fall test particles. A passing GW alters the distance between the mirrors, changing the travel distance of light. As light bounces between mirrors multiple times, it accumulates a phase shift, which is observed through interference patterns in an interferometer.

Figure 5.4 shows the schematic illustration of the Advanced LIGO detector design alongside its sensitivity for their first observing run (see [288] for sensitivities of other observing runs). The sensitivity bounds in lower, middle, and higher frequencies are respectively set by seismic noise, thermal noise of the laser, and photon shot noise.

5.3.2 Summary of observations

Current observations The first observational evidence for GWs was detected in 1981 through the orbital energy loss of a binary pulsar system, consistent with gravitational radiation [289, 290]. The first direct detection, GW150914, occurred in 2015 by LIGO detectors [281]. They found two black holes with masses $29^{+4}_{-4}M_\odot$, and $35^{+5}_{-4}M_\odot$ merging at the luminosity distance of $410^{+160}_{-180}M_\odot$, forming a final black hole with mass of $62^{+4}_{-4}\ M_\odot$, and releasing the energy of $3.0^{+0.5}_{-0.5}\ M_\odot c^2$. The actual observed data consist of strain and frequency information as a function of time (see the bottom panel of Figure 1 in [281]). Since the GWs have to travel through the universe, the amplitude decays and the frequency gets redshifted. Hence, only considering the amplitude of GWs at their merger time, there is a degeneracy between the luminosity distance, redshift, and mass of the binary system. However, the degeneracy could be partially broken considering that the change in the frequency of the

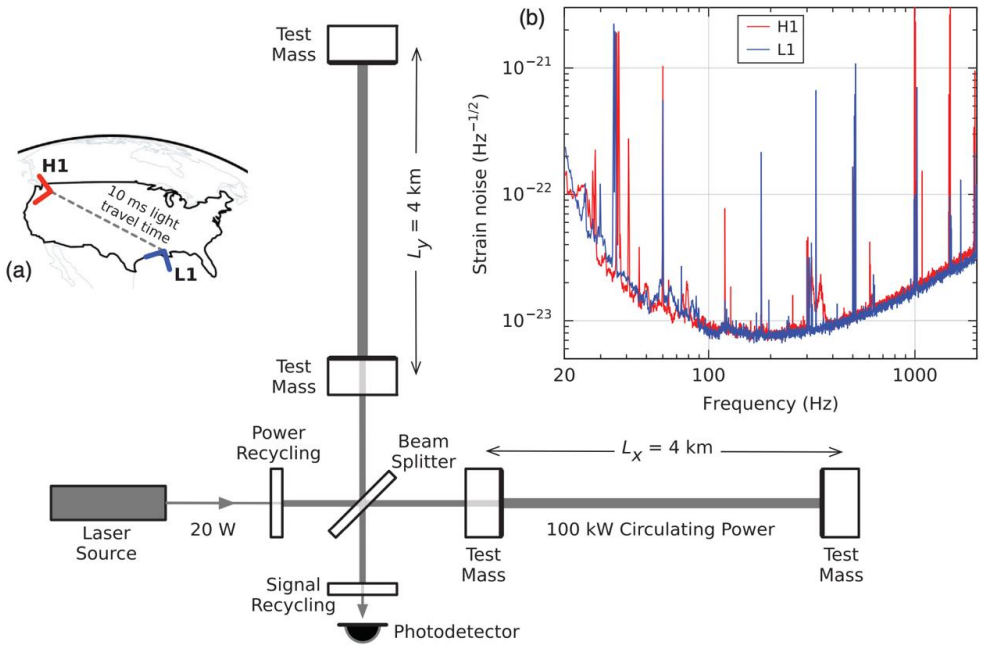


Figure 5.4: Schematic illustration of the Advanced LIGO detector design. The plot on the right shows the sensitivity of the first observing run in terms of strain vs frequency. See [288] for sensitivities of other observing runs. The figure is adapted from [281].

wave, $\dot{f}$, is only a function of intrinsic binary dynamics. Therefore, by measuring $\dot{f}$, they can constrain the following parameter, called the *Chirp Mass*, as [281]

$$\mathcal{M} \equiv \frac{(m_1 m_2)^{3/5}}{(m_1 + m_2)^{1/5}} = \frac{c^3}{G}\left(\frac{5}{96}\pi^{-8/3} f^{-11/3}\dot{f}\right)^{3/5}, \tag{5.21}$$

where m_1, and m_2 are the two masses in the binary system. In practice, modeling the emitted GW becomes more complex when considering spin, polarization, and orbital geometry details like eccentricity. Thus, detailed statistical inference (often using waveform templates from numerical relativity) is required to estimate model parameters. However, in any case, redshift is not directly measurable from GW events—only the luminosity distance can be determined. Distance is also degenerate with the inclination angle—the angle between the line of sight and the binary plane—introducing large uncertainties in distance measurements.

LIGO consists of two detectors located in Livingston and Hanford, while Virgo is based in Italy. Later, a third detector, KAGRA, located in Japan, joined the network. All have arm lengths of about 3-4 km[6]. As of the time of writing this thesis, the LIGO-Virgo-KAGRA

[6]There is another detector in Hanover, Germany, called GEO600 with shorter arm length (about 600 m) which mostly serves as a collaboration tool for the other detectors [291].

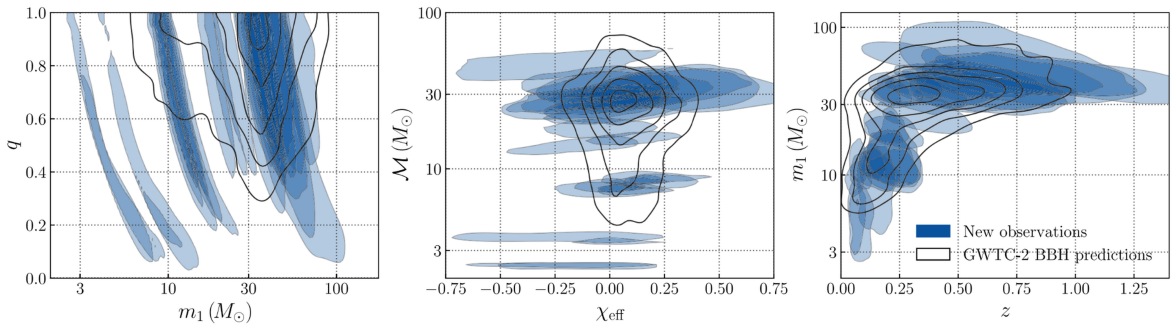


Figure 5.5: Summary of the most recent LIGO-Virgo-KAGRA results. $q = m_1/m_2$ is the mass ratio, and m_1 is the larger mass in the binary system. $\chi_{\rm eff}$ is the effective inspiral spin parameter. The redshift, z is measured assuming a particular cosmological model. The figure is adapted from [292].

network has detected 76 events[7] [292][8]. The list comprises 69 BBH events, 4 BHNS events, 2 binary BNS events, and one event that can be characterized as BBH or NSBH. Figure 5.5 summarizes their result.

Future observations The LIGO network is expected to have its third detector in India. The prospects of the full network are summarized in Figure 5.6. The improved detector sensitivities will allow the detection of sources located as far as $\sim$ 300Mpc. In addition, two more detectors are planned for the future, commonly referred to as the next generation of ground-based GW detectors. *Cosmic Explorer* is planned to begin construction in 2030 in the United States, though the exact location has yet to be determined [293]. With an arm length of around 40 km, Cosmic Explorer will increase sensitivity by an order of magnitude, particularly in higher frequency ranges. On the other hand, the *Einstein Telescope* is an underground GW detector, also set to begin construction around 2030 in Europe, with its exact location yet to be determined [294]. The Einstein Telescope will have a triangular design with an arm length of about 10 km, which improves the sky localization of events. Furthermore, since it will be built underground, it is more immune to seismic noise, allowing for better sensitivity at lower frequencies [295].

Perhaps the Laser Interferometer Space Antenna (LISA) is the most ambitious laboratory in human history. It is a space-based GW detector involving three spacecraft flying in a triangular formation, each separated by 2.5 million kilometers. It has successfully passed its

[7]On March 19, 2025, LIGO-Virgo-KAGRA announced the 200th GW detection of O4 (https://www.ligo.caltech.edu/WA/news/ligo20250320), however, the results have not been published yet.

[8]The paper is continuously updated with the latest results, so readers can refer to it for the most recent findings.

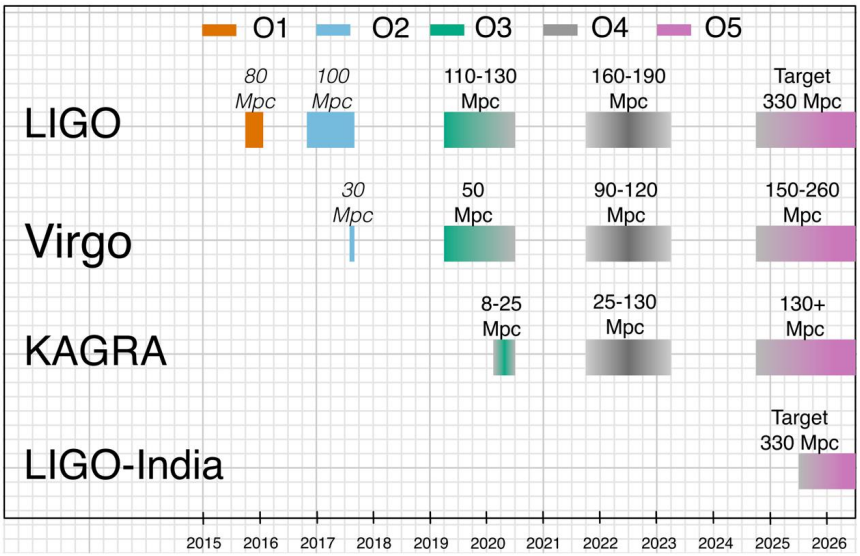


Figure 5.6: The prospects of full LIGO-Virgo-KAGRA network. The figure is adopted from [288]. The paper was updated with the last version in 2020. However, O4 and O5 were delayed, and as of today (March 2025), O4 is still in operation.

mission definition review, moving into the final design phase. The significantly increased arm length allows for probing a whole new range of frequencies, around μHz–mHz (as opposed to Hz–kHz for ground-based detectors). As Eq (5.19) suggests, this opens a new observational window to phenomena related to supermassive black holes ($m \sim 10^6 - 10^9 M_\odot$) [296, 297].

In this chapter, we showed the mathematical definition of GWs and the process of their generation. We briefly reviewed the current and future observations. GW physics is a new observational tool that can provide valuable insights into a wide range of phenomena, from quantum effects near black hole horizons to the influence of galaxy properties on black hole formation. In the next chapter, we explore what information can be extracted from the spatial distribution of gravitational wave sources.

Chapter 6

The Gravitational Wave Bias Parameter from Angular Power Spectra: Bridging Between Galaxies and Binary Black Holes

6.1 Introduction

Following the first direct observation of gravitational waves (GWs) [281] by the LIGO-Virgo-KAGRA (LVK) scientific collaboration [288, 298–300], a new observational window has opened up for studying the Universe [292, 301–303]. Since then, GWs have allowed us to study astrophysics, cosmology, and fundamental physics from a completely new perspective [304–309]. For instance, binary mergers of compact objects are shedding new light on the ongoing H_0 tension in cosmology [303, 310–320], and the first observations of the stochastic GW background from pulsar timing arrays are providing new insights in astrophysics and cosmology [280, 321–324].

There are many possible sources of GWs in the LVK frequency range, including different types of compact binary coalescing objects such as binary neutron stars, black hole neutron star binaries, and stellar binary black hole mergers (BBHs). Among these, BBHs are now mostly detected by the LVK network [307–309]. Understanding the properties of these binary systems requires comprehensive detailed models for their formation channels (see section 3.3 in [283] and section 3.1 in [284]). To establish such a model, one needs to start from a population of stars and follow different possibilities of stellar evolution to the final stellar remnants [283]. These processes depend on the mass and metallicity of the stars, which in

turn are functions of the mass, star formation rate, and metallicity of the host galaxies [283–287, 325–333]. From stellar remnants to the final coalescence of the binary system, a variety of phenomena play a role, including stellar winds, initial masses and orbital geometry of the binary system, mass transfer, supernova kicks, and common envelope physics [283, 284, 329]. BBHs could arise from stellar remnants that were born together in a binary system, or they could arise through hierarchical mergers of smaller black holes [334–336]. Depending on the choice of model parameters, different possibilities exist for the formation channel of the binary compact objects. This complexity also reflects itself in the number of free input parameters in simulation codes for stellar evolution [337], galaxy evolution [338, 339], and compact object binary formation [283, 284, 287, 329].

Therefore, the final properties of these GW sources are the convolution of several physical quantities that are redshift-dependent and host-dependent from formation time to merger time. In particular, for BBHs, it is expected that their merger rates are influenced by the stellar mass, stellar metallicity, and star formation rate of the host galaxy [331] and black hole (BH) masses may also have some dependencies on these properties [340–342]. These global properties of the host galaxies are related to the formation and evolution of the galaxies' underlying dark matter halo.

Galaxy clustering measurements have been a rich source of information about the galaxy-halo connection (see [343] for a review), from the most basic information about host halo mass to the more subtle effects such as assembly bias and velocity bias [344–355]. On large scales, clustering information is encoded in the bias parameter, which measures how much more galaxies are clustered compared to the underlying dark matter. In a similar way, the GW bias parameter relates the clustering of GW sources to the dark matter density, and naturally should depend on the formation history of the GW sources [317, 356–360]. The dependence of GW bias parameter on redshift and spatial scales can therefore bring useful insight into the connection between stellar GW sources and their host galaxies.

Modeling the GW bias parameter at large cosmological scales using the relationship to stellar properties at the sub-parsec scale can be done using cosmological simulations together with BBH population synthesis models, but is challenging due to the range of scales involved in the problem (from sub-pc to Gpc). It is also a difficult problem to construct realistic simulations that can capture the actual cosmic evolution of stars and galaxies [e.g. 361–363]. To avoid these difficulties, we use observed properties of galaxies and empirical galaxy characteristics to develop a phenomenological model for the GW-galaxy connection. In other words, we estimate the GW bias parameter by populating all-sky photometric galaxy surveys with synthetic GWs following a simple phenomenological relation for host-galaxy probability. We refer to this approach as the *phenomenological approach*, as the observed galaxy properties

are used to populate hosts of BBHs, instead of the *simulation-based approach*, where GWs are populated based on semi-analytic models and combining population-synthesis simulations with synthetic galaxy catalogs from cosmological simulations [331, 364, 365] or excursion set calculations [366].

This chapter thus investigates the potential of GW bias measurements to illuminate the GW-galaxy connection. Here, we primarily focus on the dependence of the GW bias parameter on the stellar mass of host galaxies. This is the first logical step to take since stellar mass of galaxies is the primary galaxy property linked to the halo mass (and consequently linked to the large scale bias). While this approach cannot fully resolve uncertainties or parameter degeneracies in linking the observed GW bias to the underlying physics of the GW-galaxy connection, improving our understanding of the GW-galaxy connection is a valuable aim. This relation has important implications for astrophysical formation mechanisms of GW sources and for further contextualizing their use as cosmological probes. We are also conducting follow-up research to incorporate additional galaxy characteristics that may influence GW bias behavior [3].

Many other works have considered the observability of the GW-galaxy connection by studying their correlation [367] and also cross-correlations to study the GW bias parameter [315, 317, 357–359, 364, 366, 368–372]. Prior knowledge about the GW bias parameter can also be useful for cosmological model inference. Moreover, measurements of the GW bias parameter can differentiate between an astrophysical and a primordial origin for BBHs [373–375]; in this work we will hereafter only consider astrophysical origins for BBHs.

The structure of the chapter is summarized as follows. In Section 6.2 we will review the basics of the bias parameter for tracers and introduce our phenomenological approach to generate mock catalogs of BBH mergers from galaxy catalogs. Then, in Section 6.3 we will discuss the galaxy catalogs that were used in our study. In Section 6.4, we will describe how the mock catalogs of BBH mergers were generated. We present our prescription for calculating the 2D angular power spectra of BBH mergers and galaxies that were used to compute the BBH GW bias parameter in Section 6.5. The main results of this study will be presented and discussed in Section 6.6. Finally, we make our concluding remarks in Section 6.7.

6.2 Modeling Gravitational Wave bias parameter from BBHs based on observed galaxies

6.2.1 GW bias parameter

In cosmology, the bias parameter quantifies how the overdensity field of a tracer population follows the underlying matter overdensity field. In this work, we explore the properties of the gravitational wave (GW) bias parameter, b_{GW}, defined similarly to the well-known galaxy bias. That is, given the matter overdensity as a function of redshift $\delta(x, z)$, and similarly the overdensity of gravitational wave sources $\delta_{GW}(x, z)$, we define the GW bias to parameterise the operational relation $\delta_{GW} \equiv b_{GW}[\delta]$ (which in the linear approximation reduces to a proportionality constant). Note that in this work, we are only exploring the GW bias parameter for stellar BBH sources of GWs. However, in principle, it can refer to any kind of GW source, including binary neutron stars, neutron star-BHs, supermassive BHs, and primordial BHs.

Since we will be working with a photometric survey and the angular power spectrum, we will focus on the inference of the bias parameter in spherical harmonic space. To be more specific, we assume that $\delta_{GW}(\ell, z)$ depends on $\delta(\ell, z)$ as

$$\delta_{GW}(\ell, z) = b_{GW}(\ell, z)\delta(\ell, z)\,. \tag{6.1}$$

We measure b_{GW} from the angular power spectrum of mock GW sources created from the observed galaxy catalog (described more fully in Section 6.4). This measurement thus requires a theoretical model for the angular power spectrum, $C_\ell^{\rm Mod}$, which we compute using the method described in [376] (in particular, see Eq. (44) in that paper),

$$C_\ell^{\rm Mod} = 4\pi \int d(\ln k)\ \mathcal{P}_\mathcal{R}(k)\,(\Delta_\ell(k))^2\ , \tag{6.2}$$

where $\mathcal{P}_\mathcal{R}(k)$ is the dimensionless power spectrum of the primordial curvature perturbation, and $\Delta_\ell(k)$ is the line-of-sight projection of the matter density, weighted by the galaxy redshift distribution dN/dz:

$$\Delta_\ell(k) = \int dz\, \frac{dN}{dz} T_\delta(k, z) j_\ell(k\chi(z))\,, \tag{6.3}$$

where χ is the comoving distance, $T_\delta(k, z)$ is the matter overdensity transfer function (which relates the linear matter power spectrum and the primordial curvature power spectrum $P_{\rm mm}(k) = (T_\delta(k, z))^2\mathcal{P}_\mathcal{R}(k))$ and j_ℓ is a spherical Bessel function of degree ℓ. We use the post-Limber calculation in the `CAMBSOURCES` module [376–378] to compute $C_\ell^{\rm Mod}$, which neglects contributions from the unequal-time nonlinear power spectrum and rescales the

linear transfer function to match the HALOFIT nonlinear matter power spectrum [379], while still performing the 1D redshift integrals over the spherical Bessels rather than replacing them with the Limber approximation.

Injecting dN/dz from the siren catalog and substituting the dark-matter power spectrum, the expected model angular power spectrum, $C_\ell^{\rm Mod}$, encodes the autocorrelation for a sample of the same number of sources if they were unbiased tracers of the dark matter distribution. We call this the "model angular power spectrum" as it is the projection of the theoretical 3D power spectrum onto the sky, using the observed redshift distribution. In this work, we are interested in inferring the GW bias at fixed cosmology, and so we do not vary the cosmological parameters. We use the most recent values of the cosmological parameters from the Planck Collaboration result [11] for the Flat ΛCDM (Lambda Cold Dark Matter) cosmological model, i.e. with $\Omega_c h^2 = 0.122$, $\Omega_b h^2 = 0.022$, $H_0 = 67.5$ km s^{-1} Mpc^{-1}, $n_s = 0.965$, $A_s = 2 \times 10^{-9}$ calculating the power spectrum with CAMB [376, 377].

We then proceed by fitting a constant b_{GW} within each redshift bin within a specified range of scales, chosen to ensure that we are within the range where linear bias is valid. We therefore fit the following model, $\hat{C}_\ell^{GWGW}$, to the observed power spectrum C_ℓ^{GWGW}:

$$\hat{C}_\ell^{GWGW} = b_{GW}^2 C_\ell^{\rm Mod}. \tag{6.4}$$

We also check the scale-dependence of the ratio of GW and matter power spectrum to ensure that linear bias is a valid description on the scales that we are interested in, by testing an empirical linear scale dependence in ℓ for the bias (see Section 6.6.4). The GW bias parameter, similar to the galaxy bias, may have other non-local corrections, but that is beyond the scope of this work.

6.2.2 Phenomenological approach: assigning mock BBHs to observed galaxies

A schematic overview of our analysis is illustrated in Figure 6.1. In order to estimate the GW bias parameter (the red box in Figure 6.1), we require the angular power spectrum for the merging BBHs and the model dark matter angular power spectrum. To obtain the former, we generate a mock siren catalog (the blue box in Figure 6.1) from existing observed galaxy catalogs. Catalogs and their properties will be discussed in more detail in Section 6.3. Generating the siren catalog requires three pieces of information (the green boxes in Figure 6.1), (I) *Number of mergers*: the expected number of sirens throughout cosmic history as a function of redshift, (II) *Galaxy catalog and GW host-galaxy probability*: which together give the angular distribution of sirens in the sky, and (III) *BH mass modeling*: which comes from

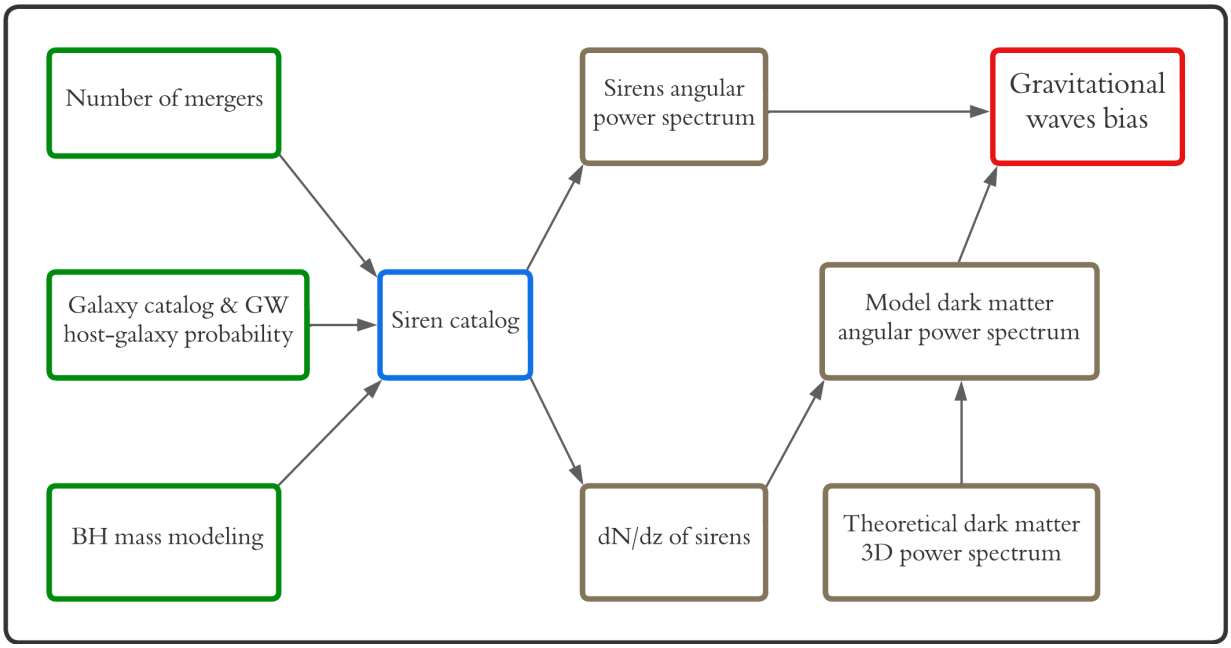


Figure 6.1: Flowchart showing our procedure for modeling the GW bias by populating galaxy catalogs with mock BBHs.

the astrophysical properties of sirens and their hosts. In principle, all of these ingredients can be correlated with each other and thus require careful treatment when generating a working mock catalog of BBHs. However, as will become evident in our exploration of GW bias, only the host-galaxy probability plays a significant role in affecting clustering. Lastly, the model angular power spectrum is calculated from the redshift distribution of the siren catalog and the matter power spectrum using the method explained in Section 6.2.1.

The number of BBHs at a given redshift is expected to primarily depend on the star formation rate [380]. By considering a relation between the star formation rate, the BH formation, and the BH coalescences, we can estimate the BH merger rate, which will be discussed in detail in Section 6.4.1. Once the BBH merger rate as a function of redshift is determined, the merger rate can then be integrated over the observing time and the size of a redshift bin to determine the total number of expected sirens in that redshift bin. Although the number of BBHs in a given redshift bin can be calculated using the BBH merger rate, to evaluate GW bias, we also need the angular distribution of BBHs in the sky. We adopt a *phenomenological survey-based* approach (or for short a *phenomenological* approach), populating host galaxies with BBHs based on a phenomenological relation for the host-galaxy probability. Specifically, we assume that the probability of a given galaxy hosting a GW siren is a broken power-law function of the stellar mass. This is discussed in

more detail in Section 6.4.2.

There are several advantages of the phenomenological approach. First, galaxy formation is a difficult problem for simulations [e.g. 361–363] and tying the BH properties directly to the observed stellar mass, star formation rate, and metallicity removes the need to marginalize over possible astrophysical models for sub-grid physics in galaxy simulations. On a related note, cosmological hydrodynamical simulations that can resolve properties of the gas in galaxies (and thus the star formation rate and metallicity) are limited to small volumes, usually $\sim 10^{-3}$ Gpc^3, whereas we use surveys that span much larger volumes, ~ 1 Gpc^3 for 2MPZ (2MASS Photometric Redshift catalog) and ~ 10 Gpc^3 for WISC (WISExSCOS Photometric Redshift catalog). Finally, our phenomenological model for the GW host-galaxy probability (i.e. the probability that a given galaxy hosts a BBH) allows flexibility in considering many different possibilities for the GW host-galaxy probability, rather than being tied to a single physical description of the galaxy-BBH connection. However, using simulations can be advantageous over using observations due to issues with the galaxy catalogs themselves, such as the completeness of a survey and observational errors (i.e., the fact that the galaxy survey does not contain all possible galaxies, but rather all galaxies above some fixed luminosity or mass threshold). In this work, we therefore always present our results in a way that minimizes the dependency on the uncertain details of the galaxy survey selection function.

The phenomenological approach requires an input galaxy catalog, either a photometric survey (requiring measurement of the angular power spectrum to infer the GW bias) or a spectroscopic survey (allowing us to use the 3D power spectrum instead). Although the 3D power spectrum using spectroscopic surveys can give a more precise measurement of clustering, and allows us to directly infer from the galaxy spectra the metallicity and the star formation rate, spectra are expensive to obtain and, therefore, spectroscopic catalogs are smaller and biased towards more massive and brighter galaxies. Furthermore, since most observations of GWs do not have a spectroscopic redshift measurement due to lack of an electromagnetic counterpart, the angular power spectrum is more closely tied to current observational prospects. While in this work we will use photometric galaxy surveys, we will explore the GW bias parameter using the 3D power spectrum and specifically using the SDSS-Main Galaxy Survey sample [381] in Chapter 7.

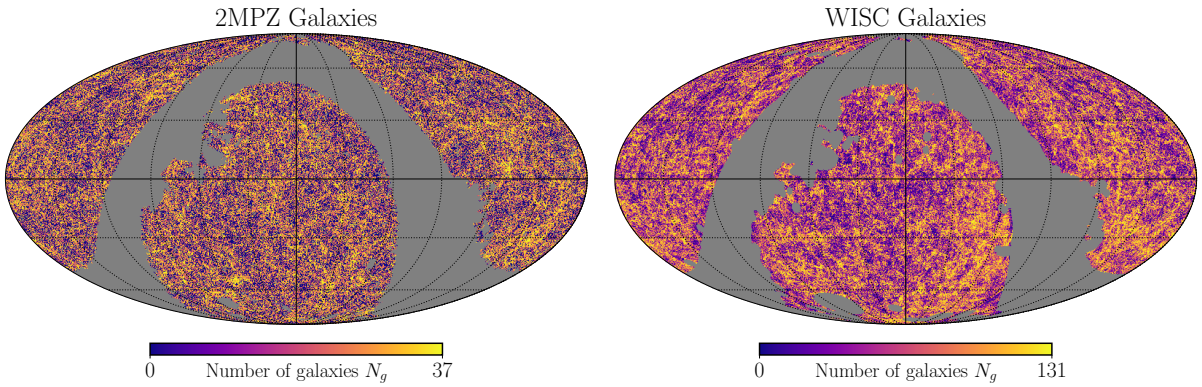


Figure 6.2: 2D sky maps of galaxies after masking, displayed in equatorial coordinates. The left panel shows the galaxies from 2MPZ, and the right panel shows the galaxies from WISC, both generated with $N_{\rm side} = 256$. Note that only WISC galaxies with $z \geq 0.1$ are shown since we did not use galaxies from WISC with $z < 0.1$ due to the higher stellar contamination rate. The gray regions are masked following the prescription in [319].

6.3 Summary of the galaxy catalogs used in the analysis

6.3.1 2MPZ and WISC catalogs

We populate mock sirens from two all-sky galaxy catalogs: 2MPZ and WISC (shown in Figure 6.2), the details of which are described below. These galaxy catalogs have a variety of advantages in creating mock GW sources. Most importantly, they cover a large fraction of the sky and have a simple selection function. Moreover, 2MPZ and WISC have previously been analyzed as part of GLADE+ [382], allowing for estimation of the stellar mass of each galaxy. However, only photometric redshifts are available for the majority of the galaxies in these surveys, restricting us to an analysis of the angular power spectrum. Also, without galaxy spectra, it is difficult to obtain accurate measurements of star formation rate and metallicity. As a result, in this work, we only consider GW host-galaxy probabilities that depend on the stellar mass of the host galaxies, which is more easily measured for these samples. In Chapter 7, we use the SDSS-MGS spectroscopic catalog instead, which provides a complementary view of GW clustering.

2MPZ catalog: The 2MASS Photometric Redshift Catalog (2MPZ[1], [383])) is an all-sky photometric redshift galaxy catalog. It was built by cross-matching the 2 Micron All-Sky Survey Extended Source Catalog (2MASS XSC, [384, 385]), the All-Sky data release of the Wide-field Infrared Survey Explorer (WISE, [386, 387]), and SuperCOSMOS [388–390].

[1] http://ssa.roe.ac.uk/TWOMPZ.html

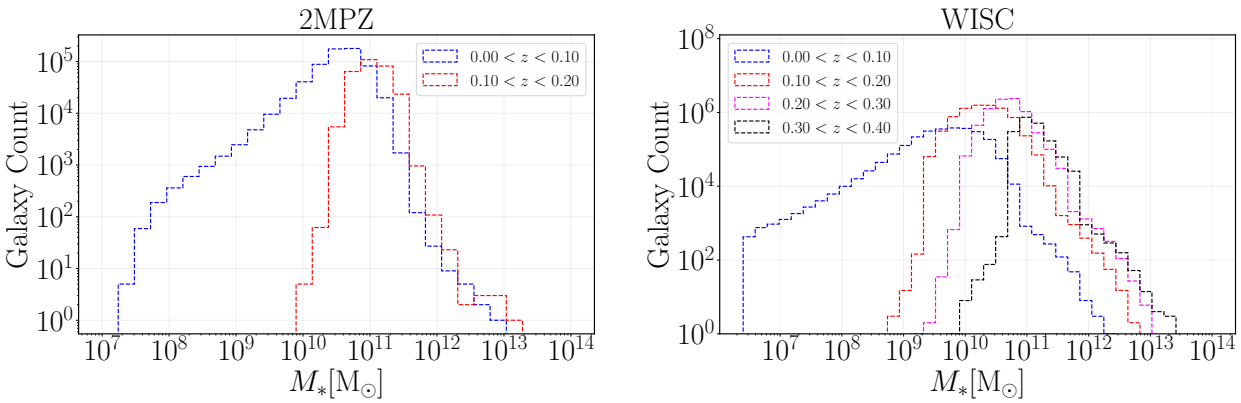


Figure 6.3: Stellar mass distribution of the 2MPZ and WISC galaxy catalogs. This does not include masked galaxies as we do not use them in our simulation.

Then, the photometric redshifts were computed using the `ANNz` artificial neural network package [391] trained on several overlapping redshift surveys, such as SDSS [392], 6dFGS [393], and 2dFGRS [394]. The resulting catalog is comprised of almost 1 million galaxies, with a median $z = 0.07$.

2MPZ includes all resolved sources in 2MASS XSC and is complete up to the magnitude band of $K_s = 13.9$ [395]. This translates to a redshift-dependent stellar mass cut, i.e. low-mass galaxies are only detected in the lowest redshift bins, as shown in Figure 6.3. The stellar mass limit for 2MPZ galaxies (*i.e.* the minimum stellar mass of a galaxy in a given redshift bin) can be seen in the right panel of Figure 6.4. This completeness limit is an unavoidable feature of using a galaxy catalog and its impact on our results will be discussed later. However, the simple selection of galaxies in the 2MPZ catalog guarantees that above the limiting magnitude of the survey, the completeness is high and it contains all types of galaxies[2]. This is critical for allowing us to produce a realistic mock catalog of GWs. Using a galaxy sample tailored for a particular galaxy population (for example, SDSS luminous red galaxies) is undesirable for creating a realistic mock GW catalog, since GWs may occur in any galaxy.

WISC catalog: We also use the WISExSCOS Photometric Redshift catalog (WISExSCOSPZ, or WISC for short[3], [397]) which was constructed by cross-matching the ALLWISE [398] data release of WISE and SuperCOSMOS Sky Survey (SCOS, [388–390]) all-sky samples. `ANNz` was once again used to derive the photometric redshifts of these samples, but with GAMA-II as the training data [399, 400]. This catalog is much larger and deeper than 2MPZ, containing about 18.5 million galaxies with a median of $z = 0.2$.

[2]Indeed, the completeness for 2MASS galaxies for $K_s < 13.9$ is about 98.6% [396].

[3]http://ssa.roe.ac.uk/WISExSCOS.html

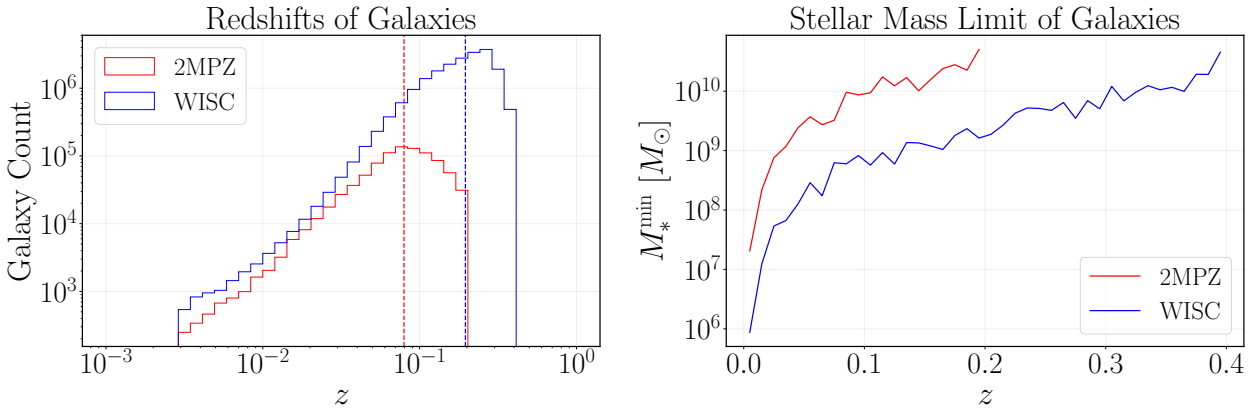


Figure 6.4: *Left panel:* Distribution of heliocentric redshifts for 2MPZ (red) and WISC (blue) in the range $z \in [0, 0.2]$ for 2MPZ and $z \in [0, 0.4]$ for WISC. The dashed vertical lines indicate median redshifts of these catalogs in these ranges, with $\bar{z}_{\rm 2MPZ} \approx 0.0795$ and $\bar{z}_{\rm WISC} \approx 0.196$. *Right panel:* Stellar mass limit for 2MPZ (red) and WISC (blue). The stellar mass limit is the lowest stellar mass of galaxies within redshift bins of $\Delta z = 0.01$. Note that 2MPZ only goes up to $z = 0.2$ since we do not consider 2MPZ galaxies with $z > 0.2$ for this work.

WISC has a more complicated selection than 2MPZ, with selection limits in both the mid-infrared $W1$ band and the optical B and R bands, as well as a restriction to use only extended sources in optical SuperCOSMOS imaging [397]. They additionally apply a Galactic-latitude dependent infrared color cut to remove stars, which are difficult to remove from the $2''$ optical imaging alone[4]. These restrictions lead to $\sim 90\%$ completeness (i.e. the fraction of galaxies included in the catalog within the observational limit) in WISC at high Galactic latitudes ($|b| > 30°$), with losses both from misclassification of compact galaxies as stars, and losses due to the color cut. However, despite these complications, WISC is advantageous as it is much larger than 2MPZ, leading to a lower stellar mass completeness limit (see the right panel of Figure 6.4).

As noted previously, Figure 6.2 illustrates the two galaxy catalogs, where the left panel presents the 2MPZ samples and the right panel shows the WISC samples. To generate the maps, we have used the Hierarchial Equal Area isoLatitude Pixelization (`HEALPix`[5], [401]) library, employing the Python package `HEALPy` [6] [402]. The gray-shaded regions in the figure represent masked areas, which were obtained using masks from [319], which followed the procedures in [403] to mask 2MPZ and [404] to mask WISC. Furthermore, for 2MPZ, the authors of [319] manually excluded the region around LMC and SMC, as well as areas

[4]WISE has a $6''$ PSF and is therefore unsuitable for star-galaxy separation.
[5]https://healpix.sourceforge.io
[6]https://healpy.readthedocs.io/en/latest/

with $< 85\%$ completeness in 2MPZ (determined by comparing the counts of 2MPZ sources and 2MASS XSC sources with $K_s < 13.9$). For WISC, the regions with high extinction ($E(B-V) > 0.1$) and high stellar density were masked [319]. These masks leave $\sim 68.5\%$ ($\sim 68.1\%$) of the sky for 2MPZ (WISC). For further information on the generation of masks for 2MPZ and WISC, we refer interested readers to [319].

6.3.2 Galaxy redshifts

For 2MPZ, we used spectroscopic redshifts when available and photometric redshifts when not. Around 33% of the 2MPZ sources have spectroscopic redshifts, nearly all from the overlapping areas of SDSS, 6dF, and 2dF, which provide redshifts for many 2MPZ sources in their coverage areas. WISC provides only photometric redshifts; we use the photometric redshifts corrected for a hemispherical offset in the photometry. To convert the heliocentric redshifts to the frame of the cosmic microwave background (CMB), we use the standard formalism to convert redshifts between different frames, explained in Appendix B. We have only corrected for our relative motion to the CMB and have not corrected for the peculiar velocities of the galaxies themselves. The peculiar velocity corrections are small compared to the coarse redshift bins that we use ($\Delta z = 0.1$) and hence we do not include them. Finally, galaxies with heliocentric redshifts $z < 0.003$ were excluded from our catalog due to the risk of stellar contamination [405]. However, for WISC as will be described later, the findings suggested that contamination could be more significant, affecting our measurements to $z \lesssim 0.1$. The redshift distribution is shown in the left panel of Figure 6.4.

6.3.3 Determining the stellar masses of galaxies

To estimate the stellar mass of the galaxies, we first classified galaxies as active or passive following the methodology described in [382, 405], with the former having active star formation and the latter without. To determine whether a galaxy was passive or active, we employed the $W2$ and $W3$ magnitudes, so that $W2 - W3 \leq 1.5$ (> 1.5) indicated active (passive) star formation. However, since 2MPZ and WISC did not explicitly provide $W3$ magnitudes in the catalogs, we first performed a one-to-one cross-matching search between the galaxies in these catalogs and those in ALLWISE [398] to extract the $W3$ magnitudes. Then, we corrected the magnitudes for dust extinction using the dust extinction coefficients for WISE filters of $A/E(B-V) = 0.184$, 0.113, and 0.0241 for the $W1$, $W2$, and $W3$ bands, respectively [406] and $E(B-V)$ for each galaxy sample. Galaxies lacking a corresponding magnitude $W2$ or $W3$ from our cross-matching procedure were considered to be active, following the methodology in GLADE+ [382].

After categorizing the galaxies, we computed the luminosity in the $W1$ band, using the K correction given in [405]

$$K = -7.1 \log_{10}(1+z), \tag{6.5}$$

which arises from the fact that observed light from sources has different observed and emitted frequencies. Note that the redshift z in Eq. (6.5) is the heliocentric redshift of the galaxy. We then employed the mass-to-light in the $W1$-band ratio for passive galaxies (those with insignificant star formation) presented in [405]

$$M_*/L_{W1} = 0.65 \pm 0.07, \tag{6.6}$$

where M_* is the stellar mass of the galaxy and L_{W1} is its luminosity in the $W1$-band. According to [405], this straightforward method can effectively approximate the masses of galaxies with redshifts $z \leq 0.15$, and its results are comparable to those obtained using more sophisticated techniques. We applied this method to the entire catalog. We similarly estimated the stellar mass for active galaxies (those with significant star formation) using the mass-to-light ratio in the $W1$-band ratio provided for active galaxies in the same study

$$\log_{10}(M_*/L_{W1}) = -0.4 \pm 0.2. \tag{6.7}$$

The resulting distribution of stellar masses in various redshift bins can be seen in Figure 6.3.

6.4 Populating galaxies with BBHs

Provided with the galaxy catalogs, we now take the next step of making synthetic mock siren catalogs as outlined in Section 6.2.2, and schematically in Figure 6.1. The following three subsections, respectively, explain the building blocks related to generating the binary black hole catalog (green boxes) in Figure 6.1.

6.4.1 The merger rate and $N(z)$ of BBHs

The merger rate of BBHs [283, 285–287, 325–329, 367] is a complex process that is highly dependent on various astrophysical properties of the host galaxy of the BBH (which in turn determine the properties of the star population) and the delay time distribution of the BBH merger events. The delay time of binary mergers, $t_d = t_m - t_f$, is the time elapsed between the formation time of the progenitor stars t_f in a binary system and the time of the merger event t_m. Numerous theoretical studies have investigated the behavior of the delay-time distribution [325, 329–332, 342, 367, 407–409], and its functional form, along with the

necessary parameters to describe it, remain a topic of ongoing discussion. In this work, we consider a simple power-law delay-time distribution with a minimum cutoff expressed as

$$P_D(t_d) \propto \Theta(t_d - t_{d,\min}) t_d^{-\kappa}. \tag{6.8}$$

For formation histories from BBHs with initial uniform distribution in the log-space distribution of the binaries' separation, $\kappa \sim 1$ is expected [325, 329, 332]. Assuming the BBH formation rate scales with the star formation rate and incorporating the delay time distribution, the merger-rate density of BBHs at a specific redshift can be written as

$$R_{\rm GW}(z_m) = \mathcal{A}_0 \int_{z_m}^{\infty} P_D(t_d) \frac{dt_f}{dz_f} R_{\rm SFR}(z_f)\, dz_f, \tag{6.9}$$

where z_m and z_f denote the merger and formation redshifts, respectively, and $R_{\rm SFR}(z_f)$ represents the redshift-dependent SFR density. Here, $\mathcal{A}_0$ denotes a normalization factor such that the local merger rate of gravitational wave sources at redshift $z_m = 0$ matches the observed value $\mathcal{R}_{\rm BBH} = 23.9^{+14.3}_{-8.6}$ Gpc^{-3} yr^{-1} [410]. In other words, $\mathcal{A}_0$ is given by

$$\mathcal{A}_0 = \frac{\mathcal{R}_{\rm BBH}}{\int_0^{\infty} P_D(t_d) \frac{dt_f}{dz_f} R_{\rm SFR}(z_f)\, dz_f}. \tag{6.10}$$

To estimate the quantity $R_{\rm SFR}(z_f)$ in Eq. (6.9) and Eq. (6.10) we take an analytical best-fit formula motivated by observational data [380] (the Madau-Dickinson SFR model), given by

$$R_{\rm SFR}(z) = 0.015 \frac{(1+z)^{2.7}}{1 + [(1+z)/2.9]^{5.6}} M_\odot \text{ yr}^{-1} \text{ Mpc}^{-3}. \tag{6.11}$$

Finally, dt_f/dz_f in Eq. (6.9) is the standard change-of-coordinates Jacobian between time coordinate and redshift,

$$\frac{dt_f}{dz_f} = \frac{1}{H(z_f)} \frac{1}{(1+z_f)}, \tag{6.12}$$

The SFR density $R_{\rm SFR}(z)$ can vary with stellar metallicity, which in turn can impact the BBH population [411, 412]. However, we have not included this effect in this analysis and instead consider a universal Madau-Dickinson SFR model [380]. A possible way to explore that will be to study the correlation between the evolution of different BBHs mergers with the evolution of emission-line-galaxies [367].

We plotted the evolution of the BBH merger rate given in Eq. (6.9) using the delay time model in Eq. (6.8) for different values of the minimum delay time and values of κ in the left panel of Figure 6.5. As expected, a higher value of $t_{d,\min}$ results in a peak at lower redshifts, since on average merger events take longer to occur. Meanwhile, a higher value of κ increases the steepness of the delay time distribution, resulting in a sharper fall-off from the peak.

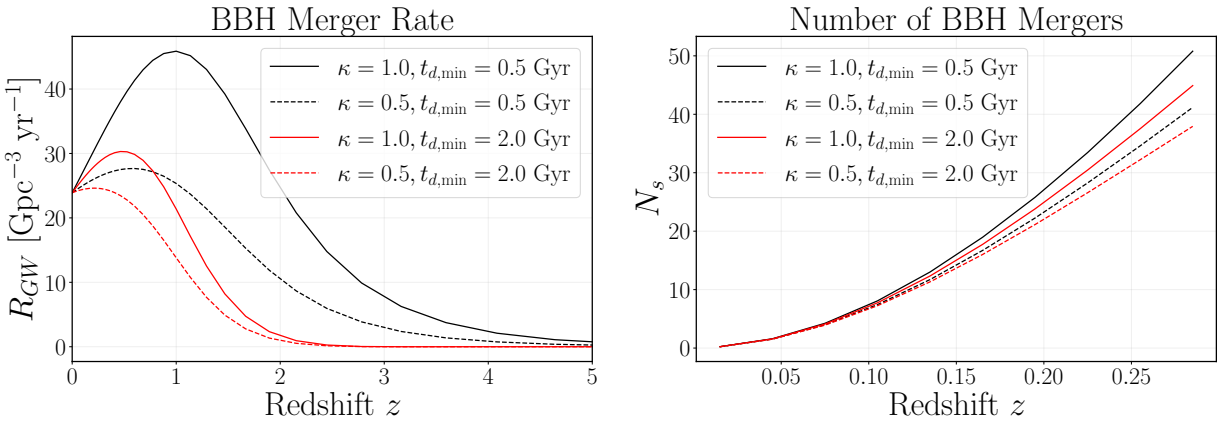


Figure 6.5: *Left panel*: BBH merger rate $R_{\rm GW}(z)$ for two different minimum delay times $t_{d,\rm min} = 0.5$ Gyr (*black*) and $t_{d,\rm min} = 2$ Gyr (*red*). The solid curves correspond to the merger rates computed with the power law index $\kappa = 1.0$ while the dashed curves correspond to $\kappa = 0.5$. As discussed in the text, these curves have been normalized so that they coincide at redshift $z = 0$ with a value of $\mathcal{R}_{\rm BBH} = 23.9$ Gpc^{-3} yr^{-1} [410]. *Right panel*: Number of BBH mergers between redshifts $z = 0$ and 0.3, computed using redshift bins of $\Delta z = 0.03$ and $T_{\rm obs} = 1$ yr. The colors and line styles of the curves are the same as in the left panel.

Given the merger rate of BBHs, one can calculate the number of BBH mergers in a redshift bin by integrating the merger rate over the width of the bin. We write the number of observable mergers in a given redshift bin as

$$N_s(z_a) = \int_{z_a-\Delta z/2}^{z_a+\Delta z/2} \frac{dV}{dz}\frac{R_{GW}(z)}{1+z} T_{\rm obs}\, dz, \tag{6.13}$$

where z_a is the midpoint of each redshift bin labeled with index $a = 1, \cdots, N_{\rm bins}$ (the total number of bins), and Δz is the width of each redshift bin. In this expression, $T_{\rm obs}$ is a constant corresponding to the observing time of the GW detector. The number of mergers in redshift bins of size $\Delta z = 0.03$ from $z = 0$ to $z = 0.3$ is displayed in the right panel of Figure 6.5 for various choices of the parameters κ and $t_{d,\rm min}$ in the delay time distribution, assuming $T_{\rm obs} = 1$ yr. In this work, we do not examine the impact of variations of the parameters κ and $t_{\rm d,min}$ defined in Eq. (6.8) on merger rate, setting their fiducial values to $t_{\rm d,min} = 0.5$ Gyr, $\kappa = 1.0$, as the number of mergers does not directly change the clustering properties and these parameters are expected to have negligible effects on our results for GW bias. In other words, in our result the delay time parameters only influence the number of BH mergers but do not affect the GW-galaxy connection. However, as we discuss in the next subsection, the delay time distribution is expected to indirectly impact the host-galaxy probability. While our fiducial host-galaxy probability model assumes shorter minimum delay times ($t_{d,\rm min} < 0.5$ Gyr), we will also present variations in the host-galaxy probability appropriate for longer

time delays (though parameterized through the shape of the host-galaxy probability function rather than explicitly via the delay time).

6.4.2 Host-galaxy probability for BBHs

In the previous section, we provided the description for obtaining the number of BBH mergers at a given redshift. Here, we introduce our approach to modeling the probability distribution of host galaxies for BBH mergers, referred to as the "GW host-galaxy probability"[7] (in agreement with the terminology used in [413]).

The host-galaxy probability can depend on different properties of the galaxy, such as stellar mass, metallicity, and SFR, and as shown in previous studies [331, 414–416], these properties are not independent of each other. However, the 2MPZ and WISC surveys contain only mass information; therefore, for the selection of host galaxies from the catalog, we only apply a conditional probability based on stellar mass, $P(GW|M_*)$. In Chapter 7, we consider a more generic host-galaxy probability that also includes the dependencies on SFR and metallicity.

From a practical perspective, in order to populate BBHs into the catalogs, we need to formulate $P(GW|M_*)$ that best encapsulates the physics from the formation to the merger of BBHs, but all of it solely manifested in terms of the stellar mass of the host galaxy at the time of the merger event. The stellar mass at merger time is influenced by other astrophysical properties at earlier times, such as the initial stellar mass, star formation history, and merger history of the galaxies. From a physics standpoint, one expects that the BBH merger rates are primarily linked to the star formation rate of the host galaxy during the period when the progenitor stars formed. The integral of the star formation rate then gives the stellar mass (up to a correction for mass loss as stars evolve). This leads to the expectation that the merger rate per galaxy will be positively correlated with the stellar mass at later times, which is also consistent with the results of population synthesis simulations [287, 331, 413].

The stellar mass–BBH merger rate relationship may not be completely monotonic, as it is also influenced by changes in the galaxy properties during the time between binary formation and merger. Specifically, high mass galaxies form stars rapidly at early times before their star formation is quenched [e.g. 417], leading to the well-known bimodality between star-forming and quiescent observed at late times (see Figure 16 in [415], showing that the division is around $M_* \sim 10^{11} M_\odot$). Hence the relationship between BBH host-galaxy probability and late-time stellar mass depends on the two timescales in the problem, the delay

[7] We use the following terms interchangeably throughout this chapter: "host-galaxy probability", "host-galaxy probability function", and "GW host-galaxy probability".

time and the quenching time. If the delay time is shorter than the quenching timescale, one expects a suppression in the number of BBHs at high stellar mass, because BBHs that formed at early times when the SFR was high would have already merged. Inferring from the empirical relation between SFR and stellar mass [415], we expect that the suppression in the BBH host-galaxy probability function would also occur around $10^{11} M_{\odot}$. Current observational constraints on the delay time distribution are weak and seem to indicate that the minimum delay time is less than 5 Gyrs [342]. On the other hand, some population synthesis simulations [e.g. 330, 331, 418] prefer longer delay times for $z < 1$, leading to a host-galaxy probability function with no suppression and even some amplification at higher stellar masses at $z \sim 0.1$ [e.g. Figure 1 in 331].

Consequently, considering all these factors, we utilize a broken power law to represent the GW host-galaxy probability function, where the break in the power law is designed to reflect the effect of suppressed SFR history in high-mass galaxies. In this framework, one can also easily study the possibility of removing the suppression at higher masses if considering longer delay times [e.g. 327, 330, 331, 414], by pushing the break scale to higher values (recovering a monotonic host-galaxy probability function in the limit of an infinite pivot mass). We parameterize the host-galaxy probability function described above[8] in the following way

$$P_{\rm host}({\rm GW}|{\rm M}_*) = \begin{cases} A_{M_*} 10^{(\log_{10}(M_*) - \log_{10}(M_{\mathcal{K}}))/\delta_l} & \text{for } 7 \leq \log_{10}(M_*) \leq \log_{10}(M_{\mathcal{K}}), \\ A_{M_*} 10^{(\log_{10}(M_{\mathcal{K}}) - \log_{10}(M_*))/\delta_h} & \text{for } \log_{10}(M_*) \geq \log_{10}(M_{\mathcal{K}}), \end{cases} \tag{6.14}$$

where $1/\delta_l$ and $1/\delta_h$ control the power of the growing and decaying regime of the host-galaxy probability function, $M_{\mathcal{K}}$ is the pivot mass that divides the broken power laws, and A_{M_*} is the normalization factor. For our fiducial model (Figure 6.6), we set the values of the power indices as $\delta_l = 4$ assuming a mild growth (positive slope of ~ 0.25), $\delta_h = 0.5$, a steep fall off in the suppression region (negative slope of ~ -2), and $M_{\mathcal{K}} = 10^{11} M_{\odot}$ for the pivot mass. As described earlier, this fiducial model is motivated by a scenario where the delay time is shorter than the quenching time, and thus binaries in high-mass galaxies have already merged.

In our full analysis on the GW bias discussed in Section 6.6, we will vary these parameters (δ_l, δ_h and $M_{\mathcal{K}}$), exploring scenarios with milder suppression $\delta_h \geq 1$, as well as those closer to the scenario obtained in [331] with faster growth and no suppression, by setting $\delta_l \sim 1$ and taking the limit of $M_{\mathcal{K}} \sim 10^{12} {\rm M}_{\odot}$. Overall, we investigate a wide range of δ_l, δ_h and $M_{\mathcal{K}}$ to explore the impact of these parameters and understand limiting cases. We investigate

[8] Practically, we use `numpy.random.choice` package for selecting the hosts with probabilities coming from the host-galaxy probability introduced here.

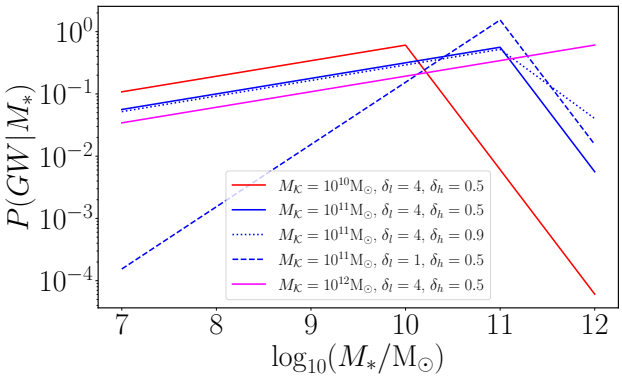


Figure 6.6: Stellar mass-dependent BBH host-galaxy probability, Eq. (6.14), for different choice of model parameters. The probability of a galaxy hosting binary BH peaks at the pivot mass $M_\mathcal{K}$ and declines afterwards, following the power laws relations above and below the pivot mass. The host-galaxy probabilities are normalized to all integrate to unity. The solid blue line corresponds to our fiducial model.

$\delta_l \in [0.5, 10]$, $\delta_h \in [0.1, 5]$, and $M_\mathcal{K} \in [10^9\mathrm{M}_\odot, 10^{12}\mathrm{M}_\odot]$. These parameters are roughly chosen for consistency with host-galaxy probabilities presented in past works based on simulations [287, 327, 331]. To compare with [331], their parameter $\alpha_1 \sim 0.8$ at $z \leq 1$ translates into $1/\delta_l$ for us, yielding $\delta_l \sim 1.2$. Similarly [287] estimate the parameter b in their models based on MZR and FMR relations to be ~ 0.47 and ~ 0.58 (see Table 2) which translates to $\delta_l \sim 2.1$ and $\delta_l \sim 1.72$. Last, the host-galaxy probability of [327] can be inferred by comparing the stellar mass distributions of BBHs from their Figure 2 (derived from populating the Millenium-II simulation with semi-analytic models for galaxies and BHs) to the stellar mass function of [419] for these simulations. For BBHs with delay times of 2 Gyr or less, this corresponds to $\delta_l \sim 1$ and $\delta_h \sim 0.5$–2; for very long delay times of 12.7 Gyr, the [327] host-galaxy probability has $\delta_l \sim 0.5$ with no break at high mass.

6.4.3 BH mass distribution

In the previous subsection, we outlined our method for selecting host galaxies for BBH mergers. The next step involves sampling the masses of the BBHs. We follow a phenomenological model fitted to the BH mass distribution from the GWTC-3 LVK catalog [342]. This model assumes that BHs follow the same trend as stellar populations described by the Kroupa initial mass function (Kroupa IMF) [420], i.e. a power-law distribution for BH masses subject to additional minimum and maximum mass cut-offs. The lower limit is imposed since we know that below some mass scale, stellar BHs do not form. In addition to the minimum mass cutoff, there is an upper bound for the allowed range of BH mass due to pair insta-

bility supernovae, called $m_{\rm PISN}$ [340–342]. This instability is expected for stars surpassing a specific mass limit, where the core generates electron-positron pairs, leading to shock waves and consequent instability, establishing a maximum permissible mass for the resulting BH. This distribution can be expressed as

$$P(m) \propto \begin{cases} 0 & m < m_{\rm min}, \\ m^{-\alpha} & m_{\rm min} \leq m \leq m_{\rm PISN} \\ 0 & m > m_{\rm PISN}\,, \end{cases} \tag{6.15}$$

where m is the BH mass and we set $m_{\rm min} = 5M_\odot$ [421], and $m_{\rm PISN} = 45M_\odot$ [340]. α is a free parameter describing the power-law index of the mass distribution which is currently weakly constrained from the observations of GW sources [292, 302, 342]. We set $\alpha = 2.3$, the same as the power-law index of the Kroupa IMF [420] in the analysis. Though the parameters that control the BBHs mass distribution are impacted by the stellar metallicity of the parent stars [283], it will not impact the large scale ($\gtrsim 1\ h^{-1}$ Mpc scale) clustering of the GW sources and hence the GW bias parameter is not influenced by this choice of mass modeling. In other words, with this model, the GW bias, by construction, will be independent of the BH masses.

6.5 Calculating the angular power spectra for sirens and galaxies

6.5.1 Power spectrum measurement

Given the setup of the mock GW catalog in the previous sections, and as mentioned in Section 6.2, we wish to compare the 2-D angular power spectra of sirens and galaxies and measure the bias by comparing to the angular matter power spectrum (Eq. 6.4). To do so, we will generate overdensity maps δ_X in redshift bins for both galaxies and sirens and compute the angular power spectrum (pseudo)-C_ℓ using `anafast` [422]. Given a catalog of either the galaxies or mock sirens, we produced a density map of $\{\delta^a_{X,i}\}$, by computing the local overdensity in each pixel with label i and within the redshift bin labeled a such that

$$\delta^a_{X,i} = \frac{N^a_{X,i}}{\bar{N}^a_X} - 1, \tag{6.16}$$

where $X = g\,(GW)$ for galaxies (sirens). $N^a_{X,i}$ is the number of the tracer X in the pixel i and the redshift bin a. $\bar{N}^a_X$ is the average number of galaxies or sirens in the redshift bin computed from $\bar{N}^a_X = N^a_X/n_{\rm pix}$, where N^a_X and $n_{\rm pix}$ are respectively the total number of

objects and the number of pixels in that redshift bin. We then measure C_ℓ as the spherical harmonic transform of the density field, using `anafast`. We divide by the pixel window function[9] for the value of $N_{\rm side}$ that we use when creating the overdensity maps, $N_{\rm side} = 512$. Since galaxies or sirens are only obtained in locations where the mask is nonzero, we need to correct the power spectrum for mask convolution. On scales smaller than the characteristic scale of the mask, this is equivalent to dividing the power spectrum by the sky fraction of the mask [423], $f_{\rm sky}$. We adopt this approximation since we always work on sufficiently small scales where it is valid ($\ell > 20$). Finally, the shot noise is subtracted from the power spectrum; the shot noise itself is computed as the inverse of the number density observed given by $n_X^a = \frac{N_X^a}{4\pi f_{\rm sky}}$.

We smooth the power spectrum by bin-averaging with width $\Delta\ell = 10$. For estimating the bias uncertainty, we use the error for a Poisson-sampled Gaussian random field [424]

$$\sigma_{C_\ell^{XX}} = \sqrt{\frac{2(\overline{C_\ell^{XX} + 1/n_X})^2}{(2l+1)\Delta l f_{\rm sky}}}, \tag{6.17}$$

where $\overline{C_\ell^{XX} + 1/n_X}$ means first the shot noise is added back and then we calculate the average. ℓ and $\Delta\ell$ are the center and width of the bin. Also note that we have dropped the index a in Eq. (6.17) to make the equation less crowded, but the calculations are done for each redshift bin separately. An example of the galaxy and GW power spectra is shown in Appendix C in Figure C.1, along with their shot noise, error bars, and $\Delta\ell = 10$ binned power spectra.

Note, here we provide computable theoretical models of GW bias by establishing a mapping between GW sources and galaxies. Therefore, since our primary aim is to explore the phenomenological model describing GW bias dependence on galaxy stellar mass, we consider only Poisson noise from a limited number of sources. While this work focuses on theoretical modeling, any work aiming for a realistic forecast of GW bias measurement from observations will require independent frameworks that incorporate GW selection effects and instrumental noise to infer the GW bias. In particular, for the redshift range considered here, future GW observations are expected to have a selection function close to 1 (see e.g. Figure 6 in [425]), suggesting a nearly complete GW catalog that mitigates the impact of detector-related systematics on inferences of the GW-host galaxy probability.

We measure the angular power spectrum using relatively broad redshift bins. For the 2MPZ catalog, we opt for two equal bins within the range $z \in [0, 0.2]$, while for WISC, we choose three equal bins within the range $z \in [0.1, 0.4]$. The range $z \in [0, 0.1]$ for WISC is excluded due to the anticipated stellar contamination in the dataset [397, 404].

[9] https://healpy.readthedocs.io/en/latest/generated/healpy.sphtfunc.pixwin.html

The number of mergers in a redshift bin can be calculated using Eq. (6.13), with z_a representing the center of the bin. Having chosen $\kappa = 1.0$, and $t_{d,\min} = 0.5$ Gyr, determining the total number of mergers requires specifying the observational time, $T_{\rm obs}$ in Eq. (6.13). As we aim to estimate the theoretical GW bias parameter that links galaxies and BBHs, it is crucial to have a substantial number of binaries to minimize the error bars, thereby allowing us to effectively study the behavior of the theoretical GW bias parameter. A realistic observational time, spanning only a few years, is not a viable choice for our purposes of modeling the bias, as it would yield an insufficient number of sources in the siren catalog, resulting in noise-dominated bias calculations. Therefore, in our computations, we assign a considerably higher value to $T_{\rm obs}$.

However, there is still an unavoidable practical statistical limit for $T_{\rm obs}$. This upper bound is dictated by the fact that there are finite numbers of galaxies at different stellar masses in the catalogs. When we select hosts for sirens from the galaxy catalog, we do so without replacement, to avoid any galaxy being selected in the catalog multiple times. If $T_{\rm obs}$ is so high that all galaxies within a certain mass range are selected as hosts, then hosts will subsequently be selected from other mass ranges, misrepresenting the properties of the hosts away from the desired mass distribution set by the selection criteria. We could instead choose to select host galaxies with replacement, but this is undesirable, as it can distort the shot noise and small-scale clustering properties, due to some galaxies being included in the sample multiple times. Hence, $T_{\rm obs}$ must be within a sweet spot of being large enough to make the clustering measurement shot noise-dominated, but not so large as to select all galaxies with a particular stellar mass.

To quantify this, we applied the host-galaxy probability function to the survey several times while each time increasing the observational time, i.e. selecting more and more galaxies. The maximum observational time was chosen so that the mean stellar mass for the GW hosts changed by $< 5\%$. Our results are insensitive to the exact value of the observational time chosen, near this threshold. This allows us to choose $T_{\rm obs}$ for analyzing the GW bias parameter with minimal noise, as permitted by our galaxy catalogs. Following this process, the fractions $\{0.0091, 0.13\}$ and $\{0.0047, 0.012, 0.090\}$ of galaxies were selected as hosts in the corresponding redshift bins (in ascending order) of the two equal bin widths in the range $z \in [0, 0.2]$ for 2MPZ and the three equal bin widths in the range $z \in [0.1, 0.4]$ for WISC, respectively[10].

Finally, since we anticipate the bias to be approximately scale independent on sufficiently large scales [426], we will fit a constant bias within a specific range of ℓ in $[\ell_{\rm min}, \ell_{\rm max}]$. The

[10] In terms of the number of sources, this translate to $\{5664, 37123\}$ and $\{36903, 92203, 163336\}$, respectively for the redshift bins of 2MPZ and WISC.

value of $\ell_{\rm max}$ is determined based on the maximum value of the 3D comoving wavenumber $k_{\rm max} = 0.4\ h\ {\rm Mpc}^{-1}$, translating into $\ell_{\rm max} \approx \chi(z_a) k_{\rm max}$, with $\chi(z_a)$ being the comoving distance to the middle of the redshift bin z_a, calculated using our reference cosmology [11]. For $z_a = \{0.05, 0.15, 0.25, 0.35\}$, this yields $\ell_{\rm max} = \{60, 180, 290, 400\}$. As we will explain in Section 6.6.4, we also verified the robustness of our bias calculation against this assumption by varying $k_{\rm max}$ to higher values and checking a simple model for scale-dependent bias. In Appendix D we also discuss our strategy for choosing the best values for $\ell_{\rm min}$, avoiding large scales where contamination in the catalogs leads to spurious additional clustering. We start with $\ell_{\rm min} = 20$, and if the linear bias model does not fit the data well, we increase $\ell_{\rm min}$ until we find a good fit. We use $\ell_{\rm min} = 20$ for 2MPZ and $\ell_{\rm min} = [30, 50, 60]$ for the three WISC bins. The higher values of $\ell_{\rm min}$ for WISC are due to stellar contamination in the WISC sample, which adds additional power at large scales.

6.5.2 Model angular power spectrum

The estimation of the GW bias requires calculation of the model angular power spectrum $C_\ell^{\rm Mod}$, in Eq. (6.2). This process is described in Section 6.2.1 and requires accurate knowledge of the redshift distribution of galaxies and sirens. We therefore take into account the redshift uncertainties in the photometric surveys. We discuss the modeling of the photometric redshift uncertainty and how it is propagated to the inference of the redshift distribution dN/dz in the following subsection.

Photometric redshift modeling

To estimate the actual redshift distribution of the tracers, given their measured photometric redshift distributions, we convolve the photometric redshift distribution with a Gaussian or modified Lorentzian function denoted $P_{\rm conv}(\delta z)$ as per [427, 428]. This can be represented by the equation

$$\frac{dN}{dz_{\rm conv}}(z_{\rm conv}) = \int \frac{dN}{dz_{\rm photo}} P_{\rm conv}(\delta z)\, dz_{\rm photo}\,, \tag{6.18}$$

where, $z_{\rm photo}$ represents the photometric redshift, $z_{\rm conv}$ represents the convolved redshift,[11] and $\delta z \equiv z_{\rm conv} - z_{\rm photo}$. For 2MPZ, as shown in [428], $\frac{dN}{dz_{\rm conv}}$ can be estimated from the photometric redshift distribution by convolving with a redshift-dependent Gaussian. The

[11] In the literature, $z_{\rm spec}$ typically refers to the convolved redshift, which is meant to reflect the true spectroscopic redshift distribution. However, in our case, some 2MPZ galaxies have both photometric and spectroscopic redshifts, whereas others have only photometric redshifts. We therefore save $z_{\rm spec}$ for the spectroscopic redshift of galaxies with both photometric and spectroscopic redshifts.

method given in [428] uses 2MPZ galaxies with both photometric and spectroscopic redshifts to determine the scatter of the observed spectroscopic redshift, $\sigma(z)$, in particularly narrow photometric redshift bins and finds

$$\sigma(z) = 0.034 \tanh\left(-20.78z^2 + 7.76z + 0.05\right). \tag{6.19}$$

This quantity is later used to estimate the spectroscopic redshift distribution of all the galaxies in the survey. In this analysis, Ref. [428] find that a Gaussian PDF centred at that particular photometric redshift provides a good fit for $\frac{dN}{dz_{\rm conv}}$ compared to the spectroscopic redshift $\frac{dN}{dz_{\rm spec}}$ (see Figure 6 in [428]). That is, the photometric redshifts have some scatter but are overall unbiased relative to the spectroscopic redshifts. Although this is justified for the entire 2MPZ catalog, in our work we measure the clustering of mock siren catalogs, which, depending on the GW host-galaxy probability, have different stellar mass distributions from 2MPZ as a whole. Given that the photometric redshift scatter can depend on the stellar mass (fainter galaxies with lower mass will likely have worse photometric redshifts), then their photometric redshift distribution may be biased depending on the mass selection. Therefore, assuming unbiased photometric redshifts and using Eq. (6.19) for mock siren catalogs with different GW host-galaxy probabilities may be incorrect.

As an example, we extract the subset of the 2MPZ catalog that includes both photometric and spectroscopic redshifts and divide it into two equal redshift bins within the range $z_{\rm spec} \in [0, 0.2]$. For each bin, we apply the host-galaxy probability function at the lowest and highest values of $M_{\mathcal{K}}$, which respectively prioritize low-mass and high-mass galaxies (keeping δ_l and δ_h fixed at their fiducial values). We then plot the resulting distributions of photometric and spectroscopic redshifts. The results, shown in Figure 6.7, indicate that the choice of $M_{\mathcal{K}}$ results in a different photometric redshift distribution in comparison to the spectroscopic redshift distribution, particularly in the first bin with $0 < z_{\rm spec} < 0.1$. Although the photometric and spectroscopic redshift distributions are more similar for $M_{\mathcal{K}} = 10^{12}\ M_\odot$, the photometric redshift distribution is shifted to higher values than the spectroscopic redshift distribution for $M_{\mathcal{K}} = 10^{9}\ M_\odot$.

To address this potential issue, we allow the redshift distribution for sirens to differ compared to the entire galaxy population in 2MPZ and to vary depending on the host-galaxy probabilities used. To be specific, we repeat the analysis given in [428] for siren hosts obtained in different runs of the host-galaxy probability, that is, different values of $M_{\mathcal{K}}$, δ_l and δ_h. We fit Gaussians with two free parameters, the scatter $\sigma(z)$ and an offset parameter $z_{\rm shift} \equiv z_{\rm photo} - \overline{z_{\rm spec}}$, in narrow photometric redshift bins of width 0.02. In other words, we allow the centre of the photometric redshift distribution to deviate from the mean spectroscopic redshift within each narrow bin by $z_{\rm shift}$. Within each bin in photometric

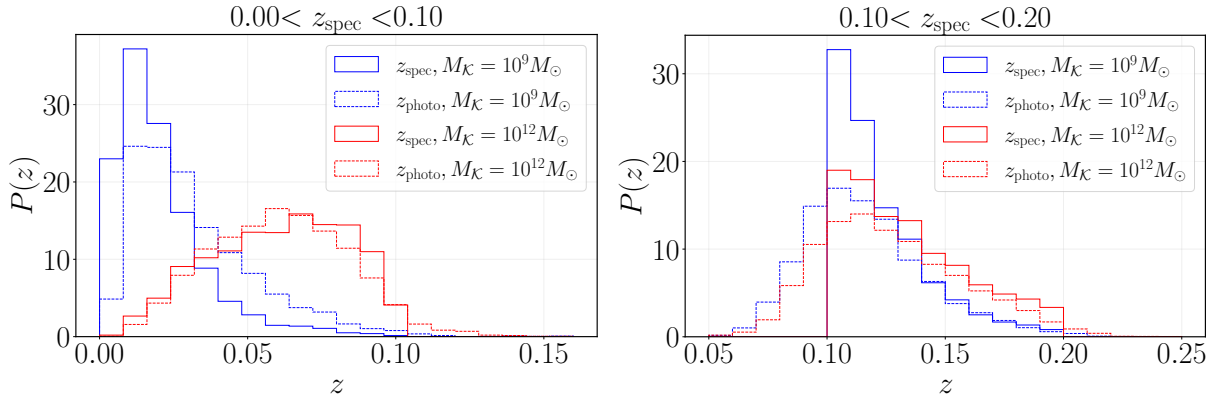


Figure 6.7: The distribution of photometric and spectroscopic redshifts for the subset of the 2MPZ catalog that includes both photometric and spectroscopic redshifts. This subset then is divided into two equal redshift bins within the range $z_{\rm spec} \in [0, 0.2]$. For each bin, we apply the host-galaxy probability function at the lowest and highest values of $M_{\mathcal{K}}$, which respectively prioritize low-mass and high-mass galaxies. Here, the values for δ_l and δ_h are set at 4 and 0.5, respectively. This plot shows that the scatter between photometric and spectroscopic redshifts is similar for both values of $M_{\mathcal{K}}$, but the low-$M_{\mathcal{K}}$ sirens show a substantial bias in their photometric redshifts, which overestimate the spectroscopic redshifts.

redshift, indexed with i, we fit the function,

$$P_{{\rm conv},i}(\delta z) \propto \exp\left(-\frac{1}{2}\left(\frac{z_{\rm conv} - \overline{z_{{\rm spec},i}}}{\sigma_i}\right)^2\right)$$
$$\propto \exp\left(-\frac{1}{2}\left(\frac{\delta z + \overline{z_{{\rm shift},i}}}{\sigma_i}\right)^2\right). \tag{6.20}$$

We find that there is significant deviation from a zero value of $z_{\rm shift}$ for different values of $M_{\mathcal{K}}$, δ_l and δ_h. For example, Figure 6.8 illustrates how $z_{\rm shift}$ changes with respect to $M_{\mathcal{K}}$ at $\delta_l = 4$ and $\delta_h = 0.5$, where the dashed lines represent the actual data. Note that after $z \simeq 0.11$, there are relatively few hosts in the siren catalog and therefore the errors on $z_{\rm shift}$ are large. We tested a few methods to determine $z_{\rm shift}$ in this region: (I) a constant or (II) a linear extrapolation. The difference between the two has a minor impact on the final bias calculation. For fiducial values of δ_l, and δ_h, we found the constant $z_{\rm shift}$ generally matches the data better so we applied that extrapolation. The red dashed line in Figure 6.8 shows $z_{\rm shift}$ for the entire 2MPZ catalog. It follows zero closely, which matches the previous assumption for the entire catalog and the findings of [428]. We can see that, as $M_{\mathcal{K}}$ decreases, the offset becomes larger, indicating that the photometric redshift measurement is worse for lower mass galaxies. While constant extrapolation works well for fiducial values of δ_l, and

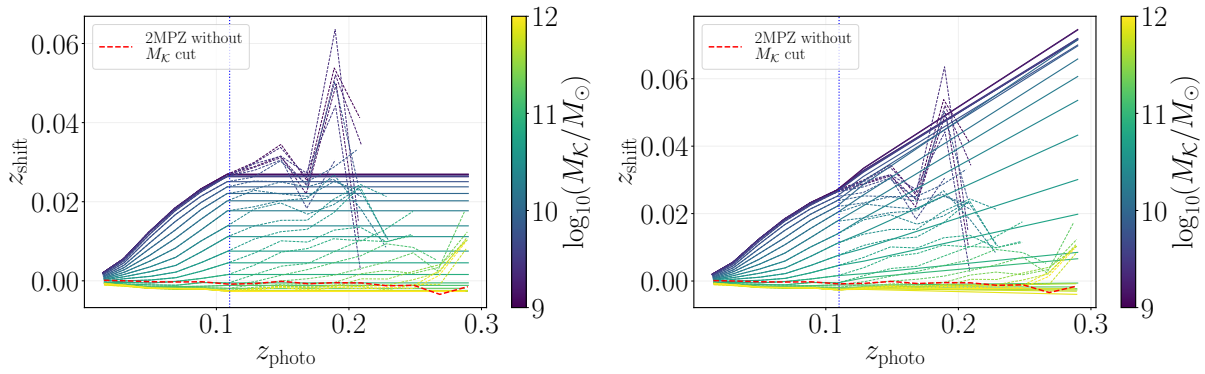


Figure 6.8: The offset between the spectroscopic and photometric redshift distribution, i.e. $z_{\rm shift} = z_{\rm photo} - \overline{z_{\rm spec}}$, for sirens of various $M_\mathcal{K}$, setting $\delta_l = 4$ and $\delta_h = 0.5$), in narrow photometric redshift bins (~ 0.02), as generated from the 2MPZ catalog. This quantity is used in Eq. (6.20). The dashed lines represent the actual data. From $z \simeq 0.11$ onward, there are not enough hosts (in the siren catalog), the data fluctuate, and we approximate it with (I) a constant (left panel) or (II) a linear extrapolation (right panel), shown by solid lines. The red dashed line ($z_{\rm shift} \approx 0$) corresponds to the result for the entire 2MPZ catalog.

δ_h, for other configurations the fit may or may not remain valid. In principle, it is possible to determine the optimal fit for each individual case, but the sparse number of galaxies at $z > 0.11$ led us to simplify this by directly using the actual data with an upper limit imposed on $z_{\rm shift}$. That is, we generate the corresponding dashed lines of Figure 6.8 for other configurations of δ_l or δ_h, but to avoid producing spurious data, we impose the condition of $|z_{\rm shift}| < 0.03$.

Next, we investigate the effect of the host-galaxy probability on the scatter. We find that the $\sigma(z)$ formulation of Eq. (6.19) is still a good fit, as illustrated in Figure E.1 for different choices of the host-galaxy probability. This is further discussed in Appendix E and the impact of choosing different convolution schemes on GW bias is presented in Figure E.2.

To summarize our results for 2MPZ, the left panel of Figure 6.9 shows the original distribution of photometric redshifts in 2MPZ alongside the convolved redshift distribution derived using this approach.

Meanwhile, for WISC, we convolve the photometric redshift distribution with a generalized Lorentzian [427],

$$P_{\rm conv}(\delta z) \propto \left(1 + \frac{\delta z^2}{2a(z)s(z)^2}\right)^{-a(z)}, \tag{6.21}$$

where $a(z) = -4z + 3$ and $s(z) = 0.04z + 0.02$. Figure 6.9 (right) shows the photometric redshift convolution for the galaxy catalogs in their corresponding redshift bins. We cannot conduct the same detailed tests for WISC as for 2MPZ, since we do not easily have access

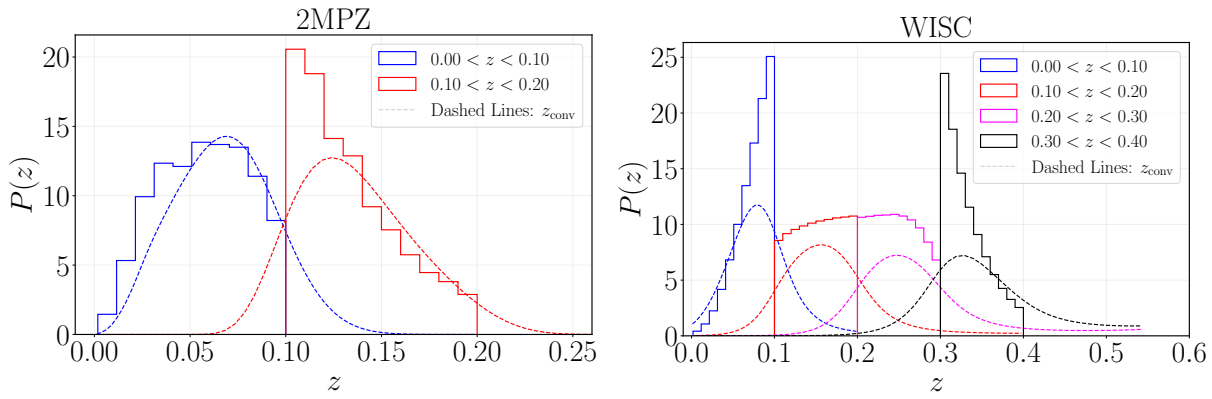


Figure 6.9: Comparison between the photometric redshift distribution (solid) and the convolved redshift distribution (dashed), for the entire galaxy samples of 2MPZ and WISC. The latter is taken into account in the calculation of the model angular power spectrum. This figure shows only galaxies that do not have spectroscopic redshifts.

to the corresponding spectroscopic redshift catalogs. Moreover, due to the higher mean redshifts of the WISC bin, the impact of changes in the redshift distribution is expected to be smaller. Hence, we consider only the simple case where the redshift convolution is the same for all the host-galaxy probabilities considered.

6.6 Results for gravitational wave (GW) bias and discussion

To confirm the validity of our analysis pipeline, we first measured the GW bias for a uniform host-galaxy probability (see Appendix C). Once we showed that the GW bias agrees with the galaxy bias in that limit as expected, we used the GW host-galaxy probability and investigated how the GW bias depends on the parameters of the host-galaxy probability function, redshift, and angular scale.

In this section, we present our results for the distinct siren catalogs associated with 2MPZ and WISC, analyzed in their respective redshift bins as outlined in Section 6.5. There is one common redshift bin between the two galaxy catalogs used here $z \in [0.1, 0.2]$. Within that shared redshift bin, the results for the GW bias agree on the overall trend for varying different parameters; however, the overall amplitude is different. This is due to different completeness levels in the two catalogs: WISC is more complete for lower masses compared to 2MPZ, and consequently yields a lower GW bias, consistent with the lower galaxy bias in WISC.

In the following, we discuss this and provide more details on how the GW bias is influenced by various parameters, namely (I) variation in relation to the parameter $M_{\mathcal{K}}$, (II) dependence on the parameters δ_l and δ_h, (III) changes in redshift, and (IV) variation with respect to the choice of $\ell_{\rm max}$.

6.6.1 Variation in pivot mass $M_{\mathcal{K}}$

We start with investigating how the GW bias changes in response to the host-galaxy probability of Eq. (6.14) for different values of $M_{\mathcal{K}}$. The plots in Figure 6.10 show b_{GW} as a function of $\log(M_{\mathcal{K}})$, while δ_l and δ_h are kept fixed at their fiducial values. We observe similar trends in both catalogs and across all redshift bins, that is, monotonically increasing, flatter towards the endpoints and a steeper slope at some midpoint, with some exceptions in the first redshift bin for 2MPZ, which will be discussed separately. To capture the observed behavior of the GW bias versus $\log(M_{\mathcal{K}})$, we fit a logistic function to the data points in Figure 6.10 as follows

$$b_{GW} = \frac{A}{1+\exp\{(-R(\log_{10}(M_{\mathcal{K}}/\mathrm{M_{mid}}))\}} + b_{\rm min}\,, \tag{6.22}$$

where A, R, $M_{\rm mid}$, and $b_{\rm min}$ are fitting constants whose values are shown in Table 6.1. $b_{\rm min}$ represents the asymptotic minimum value of b_{GW} as $M_{\mathcal{K}}$ approaches its lower end. $M_{\rm mid}$ represents the point where the slope is steepest, and $AR/4$ is the value of the derivative at that point. $A + b_{\rm min}$ represents the maximum value of the GW bias, and thus A represents how much the GW bias changes in the range of $M_{\mathcal{K}} \ll M_{\rm mid}$ to $M_{\mathcal{K}} \gg M_{\rm mid}$. While in Figure 6.10, we also display a fitted function for the first redshift bin of 2MPZ (blue curve in the left panel of Figure 6.10), in this case the best-fit function does not flatten at endpoints of $M_{\mathcal{K}}$, and the fit is not accurate in the tails, since the data points probe only the slope-changing part of the logistic function. Therefore, we do not report the fitted parameters for that bin in Table 6.1.

To better understand the trends in Figure 6.10 we examine the variation of the GW host stellar mass distribution with respect to $M_{\mathcal{K}}$ and compare it to the stellar mass distribution of galaxies in 2MPZ and WISC. Figure 6.11 shows the mean stellar mass of all the galaxies, $\overline{M_g}$, as well as the host galaxies, $\overline{M_{\rm hosts}}$, at different values of $M_{\mathcal{K}}$. This plot arises by the combination of two effects, (i) GW host-galaxy probability[12] and (ii) stellar mass distribution of the galaxies in the catalog. When $M_{\mathcal{K}} > \overline{M_g}$, the increasing part of the host-galaxy probability function becomes more relevant causing the mean stellar mass of hosts, $\overline{M_{\rm hosts}}$,

[12] As mentioned in Section 6.4.2, our fiducial model (illustrated with a solid blue line in Figure 6.6) exhibits two phases: a gentle increase followed by a steep decline just past $M_{\mathcal{K}}$.

Catalog	z	A	R	$b_{\rm min}$	$\log_{10}(M_{\rm mid}/M_\odot)$	$\log_{10}(\overline{M_g}/M_\odot)$
2MPZ	$0.1 < z < 0.2$	0.18 ± 0.01	11.0 ± 4.4	1.36 ± 0.01	11.0 ± 0.04	11.1
WISC	$0.1 < z < 0.2$	0.14 ± 0.03	10.1 ± 11.8	0.95 ± 0.02	10.2 ± 0.1	10.4
WISC	$0.2 < z < 0.3$	0.14 ± 0.02	8.4 ± 6.3	1.10 ± 0.01	10.7 ± 0.1	10.8
WISC	$0.3 < z < 0.4$	0.08 ± 0.02	11.2 ± 11.4	1.32 ± 0.01	11.2 ± 0.1	11.1

Table 6.1: Parameters of Eq. (6.22) fitted to the data points in Figure 6.10. A, R, $M_{\rm mid}$, and $b_{\rm min}$ are the fitting coefficients and constants. $b_{\rm min}$ represents the asymptotic minimum value of b_{GW} as $M_{\mathcal{K}}$ approaches its lower end. A represents how much GW bias is changed in the whole range of $M_{\mathcal{K}}$, in other words, $A + b_{\rm min}$ represents the maximum value of GW bias. $M_{\rm mid}$ represents the point where the slope is steepest, and $AR/4$ is the derivative at that point. The last column corresponds to the horizontal dashed lines of Figure 6.11, i.e. the mean mass of all the galaxies in the bin.

to be around or slightly higher than the mean stellar mass of galaxies, $\overline{M_g}$ (Figure 6.11). As $M_{\mathcal{K}}$ increases further above $\overline{M_g}$, $\overline{M_{\rm hosts}}$ rises marginally since most of the galaxies correspond to masses less than $M_{\mathcal{K}}$ and the increase rate (for $\overline{M_{\rm hosts}}$ vs. $M_{\mathcal{K}}$) is set by the shallow slope of the growing regime of the host-galaxy probability function. As a result of the increase in $\overline{M_{\rm hosts}}$ with higher $M_{\mathcal{K}}$, since galaxy bias is related to stellar mass [352, 429, 430], there is a gradual rise in the GW bias, making $b_{GW} \gtrsim b_g$ (Figure 6.10). Conversely, when $M_{\mathcal{K}} < \overline{M_g}$, the decreasing part of the host-galaxy probability function is more relevant, and the sharp drop-off causes the host-galaxy probability to strongly favor low-mass galaxies. Consequently, as $M_{\mathcal{K}}$ decreases, $\overline{M_{\rm hosts}}$ also decreases rapidly (see Figure 6.11). As a result, b_{GW} decreases below b_g (see Figure 6.10). Naturally, the transition between the two phases from the $M_{\mathcal{K}} < \overline{M_g}$ regime to the $M_{\mathcal{K}} > \overline{M_g}$ regime occurs near $M_{\mathcal{K}} \sim \overline{M_g}$. In fact, the values of $M_{\rm mid}$, where the slope is steepest for each bin, align well with the values of $\overline{M_g}$ in the corresponding bin (see Table 6.1). When $M_{\mathcal{K}}$ drops into the very low-mass regime, where the catalog is incomplete and there are very few galaxies, further lowering $M_{\mathcal{K}}$ makes no difference, and $\overline{M_{\rm hosts}}$ (and therefore b_{GW}) saturates.

The behavior of the GW bias in the tails of $M_{\mathcal{K}}$ range can also be understood in terms of the Schechter function [431, 432] with a cut-off at low masses due to the incompleteness of the catalogs. If $M_{\mathcal{K}}$ is greater than the characteristic stellar mass $\hat{M}_\star$ of the Schechter function then further increasing $M_{\mathcal{K}}$ prefers more sirens in high-mass galaxies. But since there are exponentially few galaxies with stellar mass higher than $\hat{M}_\star$, this will have minimal impact on the mean stellar mass of the siren catalog and thus the GW bias. Hence both the GW bias and the mean stellar mass flatten out for $M_{\mathcal{K}}$ above the characteristic mass of the Schechter function $\hat{M}_\star$. Likewise, if $M_{\mathcal{K}}$ drops below the stellar mass where the catalogs become incomplete, since there are so few galaxies at these low masses, further decreasing $M_{\mathcal{K}}$ will have little impact on the mean stellar mass and thus the GW bias approaches a

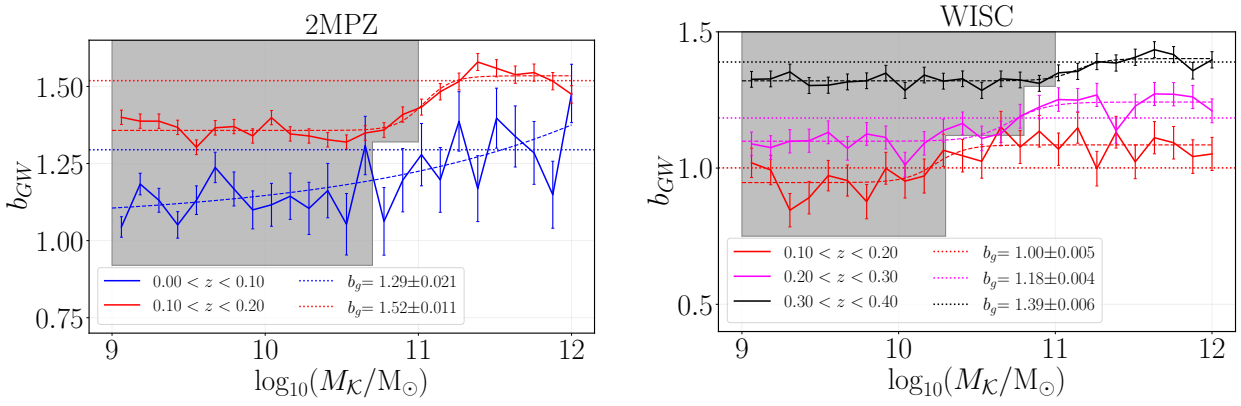


Figure 6.10: Bias for varying the pivot mass $M_{\mathcal{K}}$ in the GW host-galaxy probability for 2MPZ and WISC catalogs. In this case the low and high-mass slopes are fixed by setting $\delta_l = 4$, and $\delta_h = 0.5$. The gray areas show roughly the range that galaxy catalogs are incomplete. The dotted lines represent the galaxy bias. The dashed lines show the fitted logistic function introduced in Eq. (6.22). The fitted parameters are shown in Table 6.1. Even though we show a fitted function for the first redshift bin in 2MPZ (blue curve in the left panel), the fit is not an accurate description for the entire range of $M_{\mathcal{K}}$ and the data points probe only the slope-changing part of the logistic function. The bars indicate 1σ error.

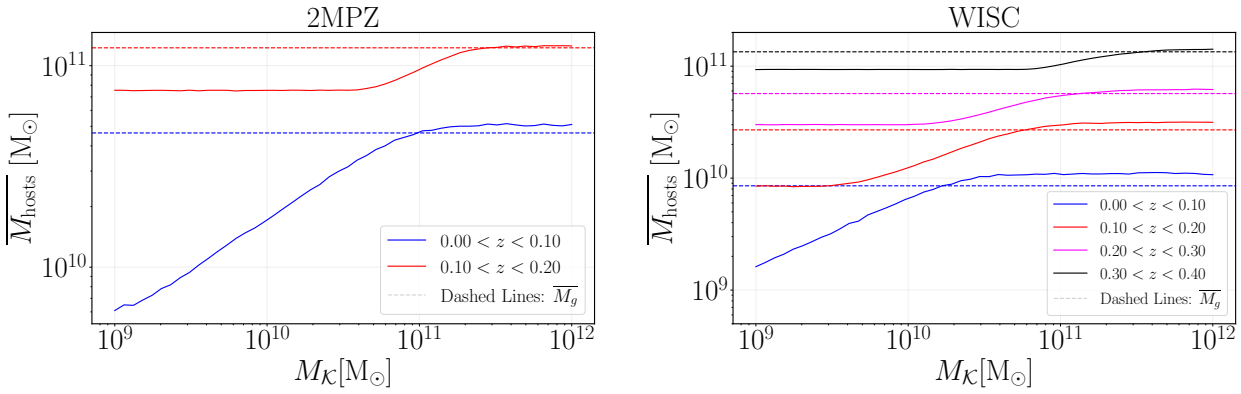


Figure 6.11: Mean stellar mass of host galaxies for different $M_{\mathcal{K}}$ values. In this case $\delta_l = 4$, and $\delta_h = 0.5$. The dashed lines show the mean mass of all galaxies (without any $M_{\mathcal{K}}$ cut) within that redshift bin. Even though we are not modeling the GW bias in WISC for $0 < z < 0.1$ (the blue curve in the right panel), we show the host mass distribution to demonstrate similar behavior for 2MPZ and WISC in this particular redshift bin.

constant value. Due to the above reason, we find that the asymptotic value of the GW bias at low $M_{\mathcal{K}}$ is highly impacted by the completeness of the galaxy catalogs. In Figure 6.10, we roughly indicate the incomplete regime of the galaxy catalogs in different redshift bins with the shaded gray area, taking the completeness threshold to start around the peak of the mass distribution in Figure 6.3, motivated by the fact that the Schechter function is a monotonically decreasing function of mass.

Another important feature noticeable from Figure 6.10 is the difference in the variation of the GW bias with $M_{\mathcal{K}}$ between the first redshift bin (blue curve in the left panel of Figure 6.10), in comparison to the other redshift bins. Firstly, note that it has relatively larger error bars, since there are fewer mergers in that redshift bin. Secondly, the stellar mass distribution of the galaxies within this low redshift bin is more complete at the lower-end of the stellar mass in comparison to a higher redshift bin. We will discuss in more detail the variation with redshift in Sec. 6.6.3.

6.6.2 Variation in slopes δ_l and δ_h

In the last sub-section, the fiducial model is chosen with $\delta_l = 4$ and $\delta_h = 0.5$. Here we consider the effect of varying δ_l and δ_h on the GW bias while keeping $M_{\mathcal{K}} = 10^{11}\ M_{\odot}$ (shown in Figure 6.12). For this particular choice of $M_{\mathcal{K}} = 10^{11}\ M_{\odot}$, the GW bias is weakly dependent on both δ_l and δ_h. The reason is that $M_{\mathcal{K}} = 10^{11}\ M_{\odot}$ is close to the peak of the catalog mass function (Figure 6.3) for all the redshift bins. Therefore, changing the slopes of the host-galaxy probability function only alters the selection probability of galaxies at the low- and high-mass ends, which have competing effects on the GW bias. This only leads to a noticeable impact on GW bias in cases where there are a significant number of galaxies in the low or high-mass tails, e.g. WISC at $0.1 < z < 0.2$ (for δ_l, from the long tail of galaxies to low mass in this bin) or $0.3 < z < 0.4$ (for δ_h, from the long tail of galaxies to high mass in this bin). However, this may not hold for a different choice of $M_{\mathcal{K}}$, especially if the delay times are long so the pivot mass is at $M_{\mathcal{K}} > 10^{11}\ M_{\odot}$ or there is no suppression in the host-galaxy probability function at all.

Therefore, we performed a more complete variation of the parameters of host-galaxy probability function, allowing all three parameters to vary simultaneously. We display examples of GW bias as a function of $M_{\mathcal{K}}$ with different choices of δ_l and δ_h in Figure 6.13 for both 2MPZ and WISC. These examples range from a linear increasing slope $\delta_l = 1$ showing a significant effect on GW bias at high $M_{\mathcal{K}}$ values (blue lines) to an almost flat growth phase with $\delta_l = 10$ (magenta), where $b_{GW} \sim b_g$. We also considered one case where the increasing phase is steeper than linear growth ($\delta_l = 0.5$) to investigate how the values of $\delta_l < 1$ can enhance the GW bias. Similarly, the high stellar mass slope in these examples ranges from

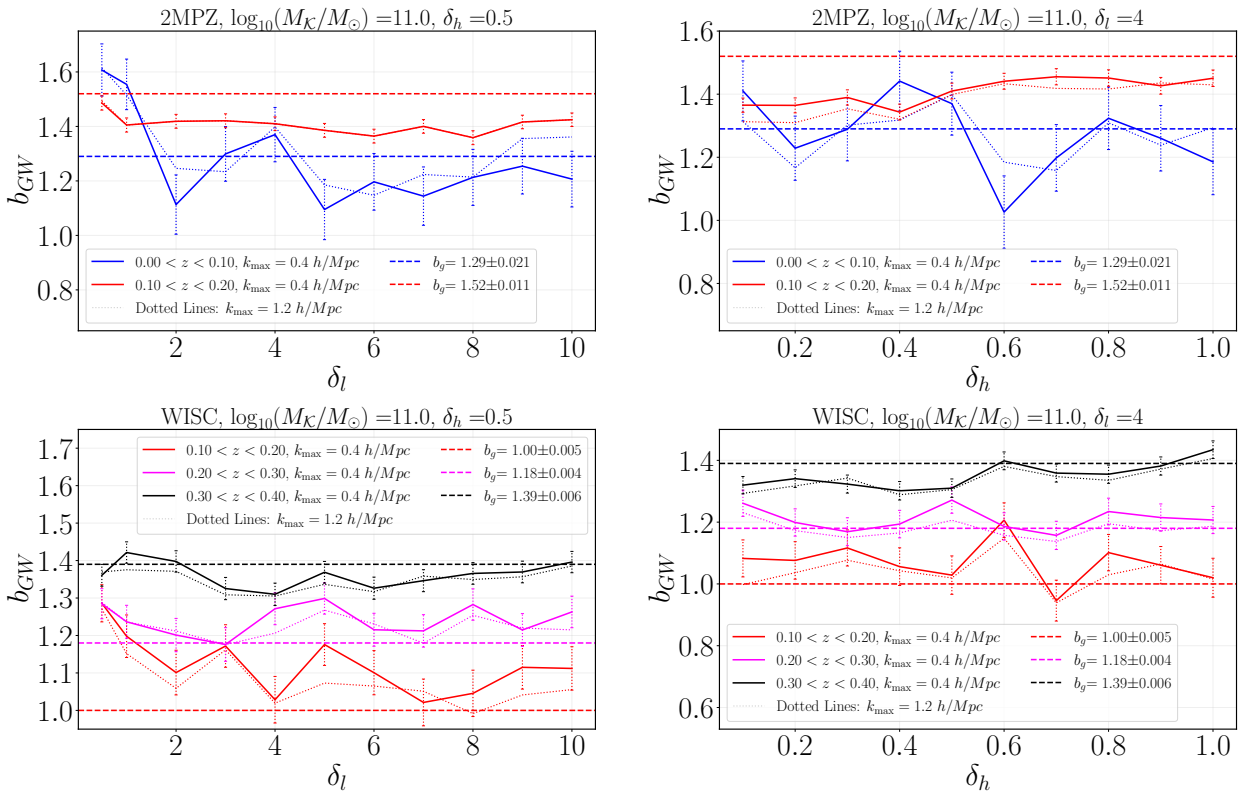


Figure 6.12: Bias for varying δ_l (left panel) and δ_h (right panel) in the GW host-galaxy probability function setting $M_{\mathcal{K}} = 10^{11}\ M_{\odot}$. GW bias is weakly dependent on both δ_l and δ_h for this particular choice of $M_{\mathcal{K}}$, which is close to the mean stellar mass of the galaxies in different bins. Nevertheless, the left (right) panels illustrate that in lower (higher) redshift bins, where $\overline{M_g}$ is below (above) $10^{11}\ M_{\odot}$, there is a noticeable increase (decrease) in GW bias as we increase the growth (suppression) rate by adopting $\delta_l \lesssim 1$ ($\delta_h \lesssim 1$). The dashed lines represent the galaxy bias. The dotted lines represent the same results for a different choice of $k_{\max}$ which confirms the scale independence to higher ℓ values. The bars are 1σ error. For the simultaneous varying all three parameters of the host-galaxy probability function δ_l, δ_h, and $M_{\mathcal{K}}$, see Appendix F.

a moderate suppression at $\delta_h = 5$ to a strong suppression at $\delta_h = 0.1$. Specifically, for cases with close to linear rise in the host-galaxy probability function with stellar mass, consistent with the simulation results in Ref. [331] ($\delta_l \sim 1$), refer to the red curves in those figures, with the solid (dotted) red curves corresponding to an insignificant (significant) suppression at high mass. In all bins except for the first redshift bin of 2MPZ where the GW bias fluctuates more, as $M_{\mathcal{K}} \to 10^{12}\ M_{\odot}$ and passes the Schechter upper limit $\hat{M}_{\star}$, the decreasing part of the host-galaxy probability function becomes irrelevant and the solid red line and the dotted line converge (as they should) to the same value (within the error bar). We also show a 3D plot of the GW bias vs. δ_l, δ_h, and $M_{\mathcal{K}}$ in Appendix F to illustrate the variation across the range of these parameters.

In summary, depending on the value of $M_{\mathcal{K}}$, changing the slopes of the host-galaxy probability function can impact the GW bias. In particular, for $\delta_l \sim 1$ and $\delta_l \sim 0.5$, there can be, respectively, up to $\sim 10\%$ and $\sim 30\%$ enhancement on GW bias relative to galaxy bias at high $M_{\mathcal{K}}$ (e.g. $0.1 < z < 0.2$ for both WISC and 2MPZ in Figure 6.13. On the other hand, at low $M_{\mathcal{K}}$, the bias is suppressed by $\sim$10% for small values of δ_h, corresponding to a steeply falling slope of the host-galaxy probability function at low mass. We find that WISC at $0.2 < z < 0.3$ is an exception to this trend, driven by anomalously high bias in the galaxy catalog at $M_* < 10^{10} M_{\odot}$. This increase is unphysical and may come from contamination in the galaxy catalog or problems with the photometric redshift (both of which occur more often for faint, low-mass galaxies).

6.6.3 Variation with redshift

As is evident in Figure 6.3 and Figure 6.10, the completeness is different for the different surveys and in different redshift bins. This leads to the trend of increasing b_{GW} with increasing redshift in Figure 6.10. This is because the galaxy catalogs become increasingly incomplete at higher redshifts, favoring higher-mass galaxies and thus more biased galaxies (greater values of b_g). Likewise, at $0.1 < z < 0.2$, where we can compare the results between 2MPZ and WISC, we find lower bias values for galaxies and sirens in WISC, which has a lower completeness limit ($\log_{10} M_* \sim 10$ vs. ~ 10.5 for 2MPZ). As a result, the most relevant measurement within each redshift bin is not the absolute value of b_{GW} but rather the comparison between b_{GW} and the galaxy bias b_g. That is, we cannot reliably compare the GW clustering to matter clustering ($b = 1$) because our mock GWs occupy a biased subset of galaxies, and thus dark matter halos. Therefore, it is more meaningful to always check whether the host-galaxy probability causes the GWs to be more or less biased with respect to the underlying galaxy bias inferred from a catalog.

The redshift dependence of the GW bias for our fiducial values of $\delta_l = 4$ and $\delta_h = 0.5$,

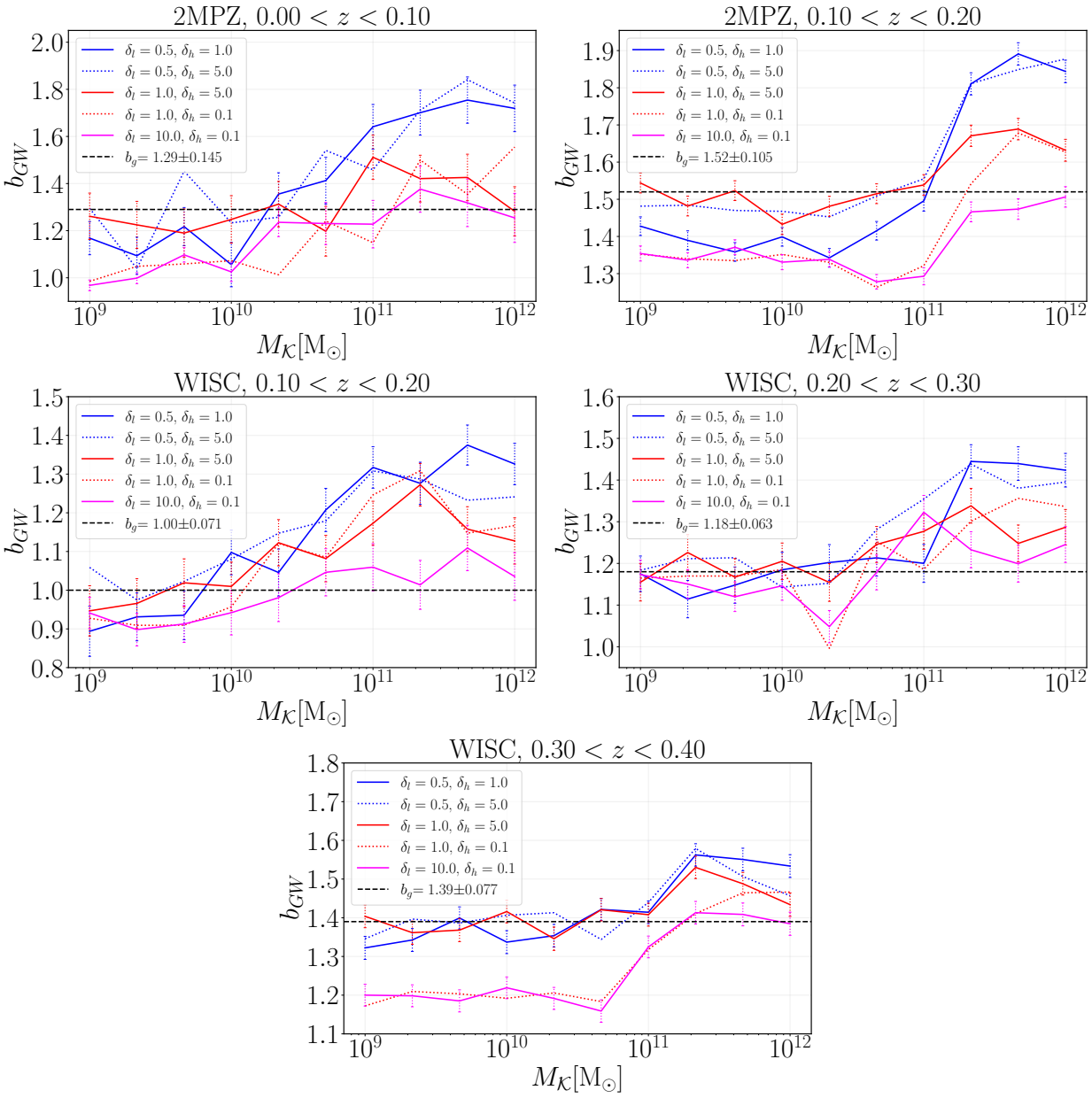


Figure 6.13: Comparisons of GW bias relative to $M_{\mathcal{K}}$ for 2MPZ and WISC under different δ_l and δ_h choices. These choices include very strong $\delta_l = 0.5$ with a very clear impact at high values of $M_{\mathcal{K}}$ to almost flat growth $\delta_l = 10$ resulting in $b_{GW} \sim b_g$ in that limit. It also includes a mild suppression $\delta_h = 5$ and a strong suppression $\delta_h = 0.1$. The solid red curve in the limit of $M_{\mathcal{K}} \to 10^{12}\ M_\odot$ depicts scenarios that align with the simulations of [331], that is, a rise with $\delta_l \sim 1$ followed by minimal suppression. We also observe greater fluctuations in the first redshift bin of 2MPZ. The bars indicate 1σ error. To avoid cluttering the figure, we removed the error bars from the dotted lines. However, the errors are approximately the same as those for the corresponding solid lines.

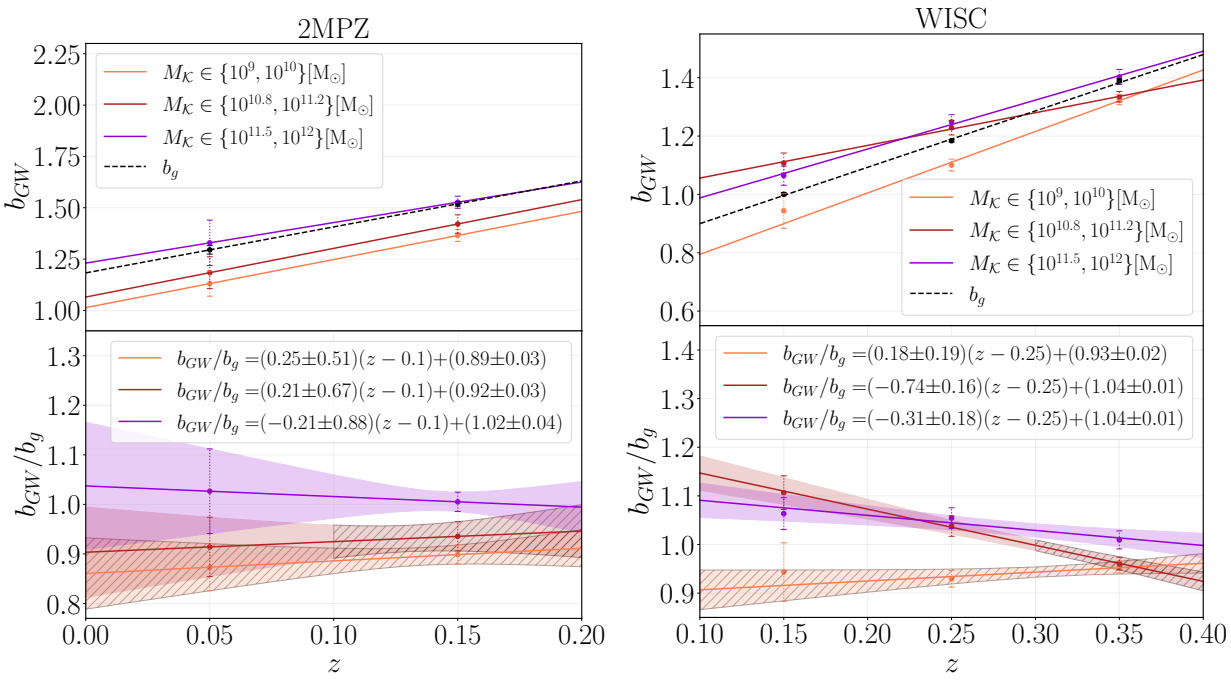


Figure 6.14: *Upper panels:* The redshift dependence of GW bias for three broad ranges in the pivot mass $M_\mathcal{K}$; (I) average of 8 data points for different $M_\mathcal{K}$ values with $M_\mathcal{K} \in [10^9, 10^{10}]M_\odot$, (II) the fiducial model represented by average of 4 data points for different $M_\mathcal{K}$ values with $M_\mathcal{K} \in [10^{10.8}, 10^{11.2}]M_\odot$, and (III) average of 5 data points for different $M_\mathcal{K}$ values with $M_\mathcal{K} \in [10^{11.5}, 10^{12}]M_\odot$. The galaxy bias is also plotted in the black dashed line. *Lower panels:* the same information plotted as the ratio of GW bias to galaxy bias. The hatched bands indicate the data points affected by the galaxy catalog's incompleteness. The bars are 1σ error.

is plotted in Figure 6.14. We compare three broad ranges in pivot mass $M_{\mathcal{K}}$, aggregating the results within each range of $M_{\mathcal{K}}$ to reduce errors. We consider the average bias across (I) 8 different values of $M_{\mathcal{K}}$ with $M_{\mathcal{K}} \in [10^9, 10^{10}]M_\odot$, (II) 4 different $M_{\mathcal{K}}$ values with $M_{\mathcal{K}} \in [10^{10.8}, 10^{11.2}]M_\odot$, and (III) 5 different values of $M_{\mathcal{K}}$ with $M_{\mathcal{K}} \in [10^{11.5}, 10^{12}]M_\odot$. The error bars are calculated by evaluating the standard deviation of the b_{GW} data points within the corresponding $M_{\mathcal{K}}$ range. This allows us to study the redshift dependence of the GW bias for host-galaxy probabilities favoring low-mass or high-mass galaxies.

The redshift dependence of GW bias inherits the trend in the redshift dependence of galaxy bias, i.e. GW bias increases with redshift, and this fact stands for all choices of host-galaxy probabilities. As discussed above, this is mostly an observational effect, since as redshift is increased, the surveys are biased to observe massive galaxies more than low mass galaxies due to the flux limit. The two bottom panels of Figure 6.14 show the ratio of GW bias over the galaxy bias (b_{GW}/b_g) for both 2MPZ and WISC along with 1σ errorbar on the fit. The plot indicates that the variation in the ratio b_{GW}/b_g with redshift is less than 15% up to $z = 0.4$ considered in this analysis for different choices of $M_{\mathcal{K}}$. The ratio shows a value less than one primarily for those scenarios that are impacted by the galaxy catalog incompleteness (lower $M_{\mathcal{K}}$ values), shown by striped lines in Figure 6.14. This mild redshift evolution for different scenarios of GW source and host galaxy connection is an interesting signature that can be explored in the future with more observations.

6.6.4 Variation in $\ell_{\rm max}$

We investigated extending our calculations to smaller scales (that is, larger $k_{\rm max}$ and $\ell_{\rm max}$) for both catalogs. In Figure 6.15, we show the GW bias versus $M_{\mathcal{K}}$ for different choices of $k_{\rm max}$. This shows that our results are robust to changing $k_{\rm max}$: using smaller scales does not noticeably change the trends in bias vs. $M_{\mathcal{K}}$, and Figure 6.15 also shows that the linear bias provides a good fit even with higher $k_{\rm max}$. We also explored the possibility of $b_{GW}(\ell)$ with a linear scale dependence. A linear function in multipoles (ℓ) did not fit much better than the constant bias. We noticed a slight preference for a decreasing bias at high ℓ; however, it was not statistically significant. We conclude that our results are robust to the choice of the minimum scale, $k_{\rm max}$.

6.6.5 Comparison with previous studies

Our results can be compared to other results in the literature studying the GW bias in simulations [364–366, 371]. These works have found a strong evolution of b_{GW} with redshift

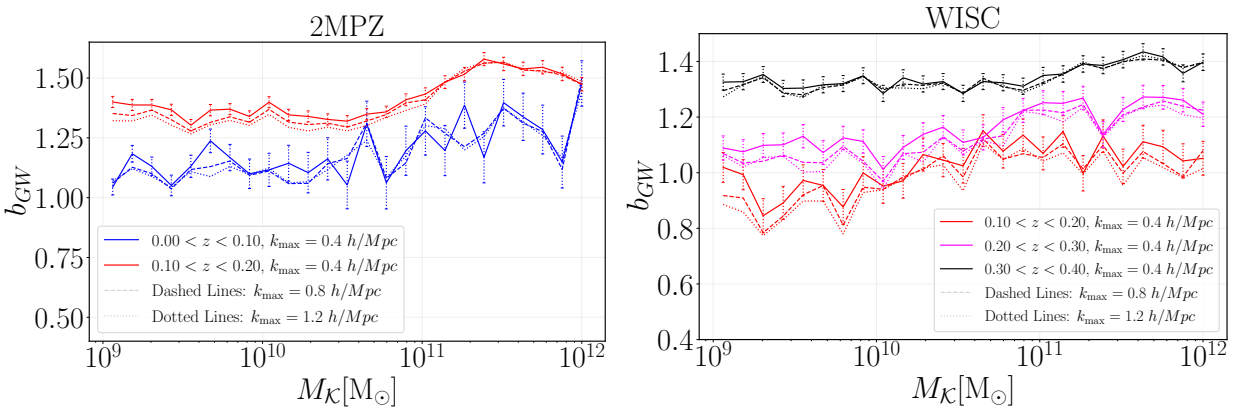


Figure 6.15: Same as Fig. 6.10 but for different values of $k_{\max}$ (different line styles), showing that our results are robust to the scales used. In this case $\delta_l = 4$, and $\delta_h = 0.5$. The bars are 1σ error.

but have not considered how b_{GW} depends on the GW host-galaxy probability and the physics relating BBHs to their host galaxies. We reiterate that our study is most effective in examining the sensitivity of the ratio between b_{GW} and b_g to the parameters of the host-galaxy probability function, given the incompleteness of the galaxy catalogs. We find b_{GW} is within 15% of b_g for our fiducial model at $0 < z < 0.4$ (see red data points in lower panel of Figure 6.14), in good agreement with [365] who find $b_{GW} = 1$ at low redshift. In contrast, [364] finds a higher value of $b_{GW} \sim 2$ at low redshift using a galaxy catalog from the EAGLE simulation and using a HOD-based approach. They also find a higher galaxy bias of $b_g \sim 1.5$, implying b_{GW} is $\sim 30\%$ higher than b_g. In scenarios with longer delay times and a moderately rising host-galaxy probability phase ($\delta_l \sim 1$), we find higher GW bias values. However, in our host-galaxy probability models, achieving a 30% enhancement of GW bias would require a rapid growth rate, around $1/\delta_l \gtrsim 2$ (refer to the blue lines in Figure 6.13). The comparison of our results and those from simulations shows how this "phenomenological" approach can complement the "simulation-based" approach of directly populating GWs into cosmological simulations. However, completeness of galaxy catalogs and photometric redshift uncertainties are potential hurdles in the "phenomenological" approach.

6.7 Conclusion

In this work, we have studied the dependence of the GW bias parameter for BBHs on the GW host-galaxy probability, using angular power spectrum measurements of mock sirens populating galaxy catalogs. We generated mock catalogs of BBHs using physically motivated relations between GWs and the properties of the host galaxies, populating the 2MPZ

and WISE all-sky galaxy catalogs. In particular, we examined an astrophysically inspired phenomenological model for the GW host-galaxy probability, characterized by a broken power law in stellar mass that rises at low mass and falls at high mass. The presence of a turnover point in the host-galaxy probability function $M_{\mathcal{K}}$ is motivated by the expected transition scale between star-forming and quiescent galaxies, which is predicted to be around $M_\star \sim 10^{11}\ M_\odot$. Essentially, our reasoning is that the BBH merger rate is expected to follow the SFR with a certain non-zero delay time. If the delay time is short compared to the quenching timescale of the galaxies, high-mass galaxies whose star formation ceased in the past should have a smaller probability of hosting a GW merger, leading to a drop in the host-galaxy probability at high mass. On the other hand, if the delay time is long, then we would expect the host-galaxy probability to trace the stellar mass and increase at high mass. We explored these physical differences by varying the pivot mass and slopes in our model.

We observed that increasing $M_{\mathcal{K}}$ and $1/\delta_l$ increases the GW bias. For the fiducial values of $\delta_l = 4$ (mild increase rate) and $\delta_h = 0.5$ (strong suppression rate), we observed that increasing $M_{\mathcal{K}}$ increases the bias, but only modestly. The GW bias at $M_{\mathcal{K}} = 10^{12}\ M_\odot$ exceeds the galaxy bias by at most $\sim 5\%$ and exceeds the galaxy bias at $M_{\mathcal{K}} = 10^{11}\ M_\odot$ by $\sim 10\%$ (see Figure 6.10). Assuming that the host-galaxy probability increases at a higher rate, as predicted by certain population synthesis models ($\delta_l \sim 1$) [331], our findings indicate that the GW bias can surpass the galaxy bias by approximately 10% when $M_{\mathcal{K}} = 10^{12}\ M_\odot$, as illustrated in Figures 6.13. It is important to note that the model we have described for the GW host-galaxy probability is quite simple, largely due to limitations imposed by the photometric galaxy catalogs, which lack direct measurements of the star formation rate and galaxy metallicity. We also explore more intricate host-galaxy probabilities that take these factors into account in Chapter 7.

We also note that since the hosts of the BBHs were taken as a sub-sample of the 2MPZ and WISC galaxy catalogs, the incompleteness of the galaxy catalogs themselves is also reflected in the mock BBH catalog. Hence, our calculation of the GW bias parameter has some systemic error depending on the incompleteness level of the input galaxy catalog. We address this issue by mainly comparing the GW bias to the galaxy bias of the primary galaxy catalog (which is similarly affected by incompleteness). However, for the host-galaxy probabilities with $M_{\mathcal{K}}$ greater than the catalog completeness limit, the bias values are less influenced by incompleteness so we believe the result is more robust in this regime. Our "phenomenological" approach, populating real galaxy catalogs, is complementary to "simulation-based" approaches, allowing more flexibility for the host-galaxy probability; probing a larger volume than hydrodynamical simulations; and avoiding dependence on uncertain galaxy formation physics. However, the phenomenological approach generally suffers from the incompleteness of the galaxy catalogs, observational errors, and less depth in redshift space, out to $z \sim 0.3$

compared to simulations spanning $0 < z < 6$.

To summarize, this method is the first of its kind, using a phenomenological approach with a physics driven host-galaxy probability to connect observed low-redshift galaxy catalogs to GW sources in terms of the GW bias parameter. This provides a framework that can be easily tested with future observations. The inference of the GW bias parameter provides a complementary probe to explore the astrophysical population of BBHs. We acknowledge that it is likely that with a single large-scale clustering measurement, uncertainties and parameter degeneracies would make it difficult to infer more than one or two effective parameters describing the GW-host galaxy probability (i.e. break mass and slope). However, this would still provide valuable insight into the GW-galaxy connection, and this information could be enhanced by smaller-scale clustering measurements or higher-point measurements of the small-scale GW-galaxy cross-correlation. We leave these avenues to future work.

Insight into the GW bias parameter would be valuable for inference of cosmological parameters, such as the Hubble constant and dark energy equation of state using BBH-galaxy cross-correlation [317, 359]. In these studies, it is crucial to include the variation of GW bias and marginalize over its uncertainties; hence a better understanding of the GW bias parameter could strengthen the analyses. Similarly to the GW bias parameter for BBHs, it is possible to extend this technique for calculating the GW bias parameter for other GW sources such as binary neutron stars and neutron star black hole systems. The future holds even greater potential for this approach to modeling the GW bias, as galaxy catalogs from DESI [433], Euclid [434], Vera Rubin Observatory [435], and Roman Space Telescope [436] will illuminate the GW-galaxy connection up to high redshifts.

Chapter 7

Gravitational Wave Bias from 3D Power Spectrum: Tracing Binary Black Hole Mergers Through Galaxy Surveys

7.1 Introduction

As proposed in Chapter 6, the gravitational wave (GW) bias parameter measures the relationship between the clustering of GW sources and the underlying matter field. This relationship can be used to study the formation history and environmental dependence of these sources[315, 317, 357–359, 364, 366–372]. Our goal is to explore how the GW bias parameter is influenced by the astrophysical properties of the galaxies hosting the merger events. In our previous study [2], we investigated the theoretical GW bias parameter by populating simulated GW sources in photometric all-galaxy surveys (2MPZ and WISC). We used a phenomenological model for the GW host-galaxy probability function consisting of a broken power law in stellar mass. We then measured the bias of these synthetic GW sources using the angular power spectrum and studied the impact of variations in the host-galaxy probability function. In this work, we extend that analysis by populating the simulated GW sources into a spectroscopic galaxy survey, which allows for more complex host-galaxy probability functions depending on galaxy properties other than stellar mass. Moreover, spectroscopic galaxy surveys offer more accurate redshift measurements and better ability to measure galaxy clustering via the three-dimensional power spectrum. While this simulated three-dimensional GW catalog is further from what will be available observationally (where the majority of GW sources will not have redshift information), it allows a more detailed theoretical probe of the relationship between the GW sources and the matter field.

The structure of this chapter is outlined as follows. In Section 7.2, we review the theoretical GW bias parameter and also introduce the phenomenological approach in this study, where GW sources are populated within observed galaxies using a model for the host-galaxy probability depending on stellar mass, SFR and metallicity. Section 7.3 presents the spectroscopic galaxy survey utilized in this analysis, the Sloan Digital Sky Survey (SDSS DR7) Main Galaxy Sample. In Section 7.4, we describe in detail the host-galaxy probability function. Section 7.5 describes the theoretical framework for calculating and modelling the 3D galaxy power spectrum of the mock GW sources. In Section 7.6, we measure the best-fit value of the GW bias parameter and show how it is sensitive to variations in the host-galaxy probability function. Finally, in Section 7.7, we summarize our findings and offer concluding remarks.

7.2 Modeling gravitational wave bias parameter

7.2.1 Gravitational wave bias parameter

The bias parameter measures to what extent a discrete tracer follows the underlying matter over-density field. For instance, the galaxy bias quantifies the over-density of galaxies, δ_g with respect to the over-density of the underlying distribution of matter, δ. The over-density for galaxies can be generally expressed as [266]

$$\delta_g(x,\tau) = \sum_W b_W(\tau) W(x,\tau) \tag{7.1}$$

where W is a set of local and non-local operators on the matter over-density, with each having a corresponding bias parameter b_W – for example, powers of the density and tidal field, time derivatives of the tidal field and higher-derivative terms. The simplest model for galaxy bias, which is a good approximation on large scales, can be described through a linear and local relationship as

$$\delta_g = b_g \delta_m. \tag{7.2}$$

where δ_g and δ_m is the galaxy and matter over-density. Note that Since b_g is a constant, this relation remains local in both Fourier space and real space. However, in general and particularly on small scales, this relation can be non-local as well. In this chapter, we explore the leading order, local, and linear bias. In other words, we define the bias parameter through the relationship between the power spectrum of the galaxies $P_g(k,z)$ and the matter power spectrum $P_m(k,z)$ as

$$P_g(k,z) = b_g^2 P_m(k,z). \tag{7.3}$$

Similarly, we define the GW bias parameter to describe how effectively the spatial distribution of the GW sources[1] follows the underlying matter distribution. GW bias parameter can be defined as

$$\delta_{GW} \equiv b_{GW}\delta_m \,, \tag{7.4}$$

where δ_{GW} is the over-density field of gravitational wave sources, and b_{GW} is the GW bias parameter. For the power spectrums we have

$$P_{GW}(k,z) = b_{GW}^2 P_m(k,z), \tag{7.5}$$

where $P_{GW}(k,z)$ is the corresponding power spectrum of the GW sources. To determine the GW bias parameter, we fit a phenomenological model to the redshift-space power spectrum of the mock sirens, consisting of a linear galaxy bias, linear redshift space distortions, and a Finger-of-God velocity dispersion parameter. More details are given in Section 7.5.2.

7.2.2 A Phenomenological framework for measuring GW bias parameters using mock GW source catalogs generated from galaxy surveys

In this study, we adopt a *phenomenological* approach to compute the GW bias parameter, as illustrated in the flowchart in Figure 7.1. To calculate the GW bias parameter, two essential components are required: the 3D power spectrum of BBH mergers[2] and the theoretical 3D power spectrum of dark matter, which can be conveniently computed using CAMB (Code for Anisotropies in the Microwave Background, with details provided in Section 7.5). To calculate the power spectrum of the BBH mergers, we begin by creating a mock catalog of BBHs by selecting their host galaxies from the spectroscopic galaxy survey SDSS DR7. The initial step involves determining the expected number of mergers occurring at a particular redshift, which depends on the astrophysical properties of both the formation channel of BBHs as well as the environment of the host galaxy. For instance, the merger rate depends on redshift, the observing time of a detector, SFR- density within the host galaxy and the time delay between the formation of a BH and the subsequent coalescence of two BHs (explained in section 7.4.1).

In the second step, we use two distinct host-galaxy probabilities to compute the probability of a galaxy hosting a BBH merger based on its astrophysical properties. First, following

[1]GW sources include binary neutron stars, binary black holes, neutron star- black hole binary and primordial black holes.

[2]Based on current observations, the majority of GW merger events are expected to originate from BBHs, which form the primary focus of this study. However, it is worth noting that our framework can be readily extended to incorporate binary neutron star (BNS) and neutron star-black hole (NSBH) systems as well.

[2], we use a host-galaxy probability function that only depends on galaxy stellar mass to allow us to directly compare our results to those of [2]. Second, taking advantage of the additional information in the spectroscopic galaxy catalog, we consider a host-galaxy probability function that depends on SFR and metallicity as well as stellar mass. This is a more realistic host-galaxy probability function because the occurrence of GW sources likely depends on more than just the host galaxy's stellar mass. Using these probabilities, we then select certain galaxies as mock sirens (see Section 7.4.2 for more details). In the third step, we assign masses to the BHs using the method explained in Appendix J.1 – allowing us to study the dependence of GW bias parameter on the BH masses. Finally, we compute the best-fit GW bias parameter using galaxy and siren power spectra and study the effect of redshift, parameters of the host-galaxy probability function, and the constituent BH masses on the GW bias parameter.

One key advantage of the phenomenological approach is its ability to tie black holes directly to observed galaxy characteristics, such as stellar mass, SFR, and metallicity, thereby avoiding uncertainties introduced by the sub-grid astrophysical models used in galaxy simulations. Spectroscopic surveys like SDSS DR7 cover large volumes, allowing for robust statistical analyses of galaxy clustering across extensive datasets. In contrast, cosmological hydrodynamical simulations [361–363], while offering detailed insights into gas properties within galaxies and being more complete than observational surveys, are limited to much smaller volumes, which makes large-scale studies of galaxy clustering and its connection to gravitational waves more challenging. Spectroscopic surveys such as SDSS DR7 also enable more precise clustering measurements and allow for the direct determination of galaxy properties like metallicity and SFR. Although these surveys are small-scale and biased toward brighter galaxies, the constraints due to selection biases and incompleteness in this work are carefully addressed to minimize the corresponding effects. Having considered these factors, it is worthwhile to note that incorporating simulations instead of observational data in future research could provide complementary insights.

7.3 Galaxy catalog

The Sloan Digital Sky Survey (SDSS DR7) includes the full SDSS-I and SDSS-II data sets. In these surveys, CCD imaging [437, 438] was obtained in five bands (u, g, r, i, z) across 11663 deg^2, with a uniform photometric calibration described in Ref. [439]. From this data set, a subset of galaxies within an area of $9,380$ deg^2 [440] underwent spectroscopic follow-up as part of the Main Galaxy Sample (MGS,[441]). The MGS selection criteria approximately encompass all galaxies with an extinction-corrected r-band Petrosian magnitude (r_{pet}) below

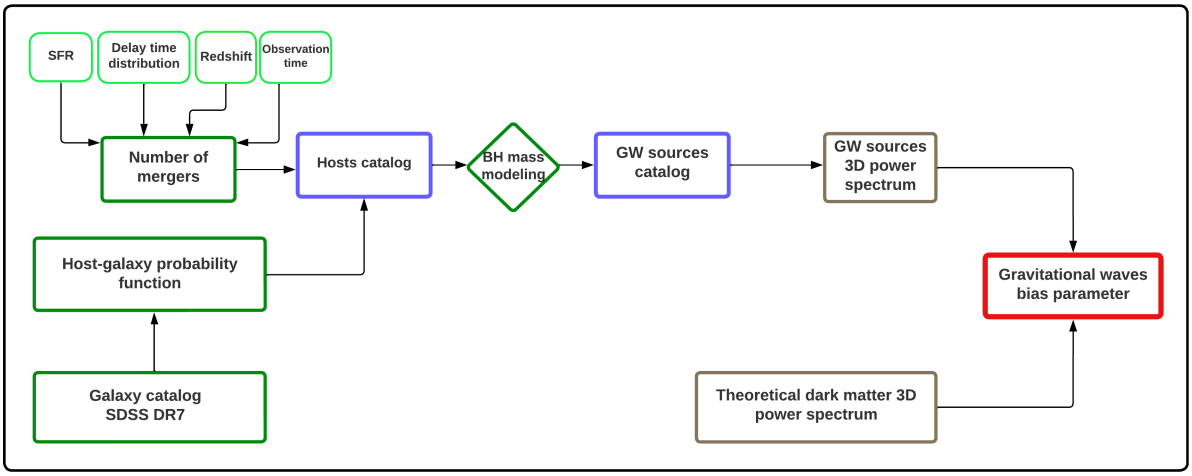


Figure 7.1: The flowchart that shows the framework for the GW bias parameter calculation.

17.77 [442, 443]. On the other hand, VAGC was created from the SDSS DR7 MGS redshifts [444]. NYU-VAGC "safe0" galaxy sample specifically includes galaxies with a range of $14.5 < r_{pet} < 17.6$ which cover a total of 7356 deg^2. By considering $r_{pet} = 14.5$ as the lower limit, it is guaranteed that galaxies with trustworthy SDSS photometry are included, while the upper limit enables us to make a homogeneous selection across the entire area of 7356 deg^2 [444]. However, at the bright end, saturation in SDSS images can lead to inaccurate galaxy photometry, potentially affecting measurements for the most luminous galaxies.

7.3.1 Sampling from SDSS DR7

Our galaxy sample has redshift and completeness cuts similar to the SDSS-MGS BAO sample of Ross et al. [381] but not identical; they restrict the sample to high-mass, passive galaxies by applying additional color cuts, whereas we want as complete a galaxy sample as possible. For our analysis, we focus exclusively on the continuous region within the North Galactic Cap to achieve a more homogeneous density map. Additionally, we limit our analysis to areas where the completeness exceeds 0.9. Completeness is defined as the fraction of observed galaxies relative to the expected number, accounting for observational effects such as survey geometry and galaxy selection criteria. In this context, completeness is calculated without considering galaxies missed due to fiber collisions. This approach ensures that the sample remains robust and representative of the true galaxy distribution, consistent with the methodology outlined in the SDSS-MGS BAO sample analysis by Ross et al. [381]. We also apply a redshift cut of $0.07 \leq z \leq 0.2$. As a sanity check, we also considered all cuts[3] applied in Ref. [381] and

[3]Details can be found in Appendix H.

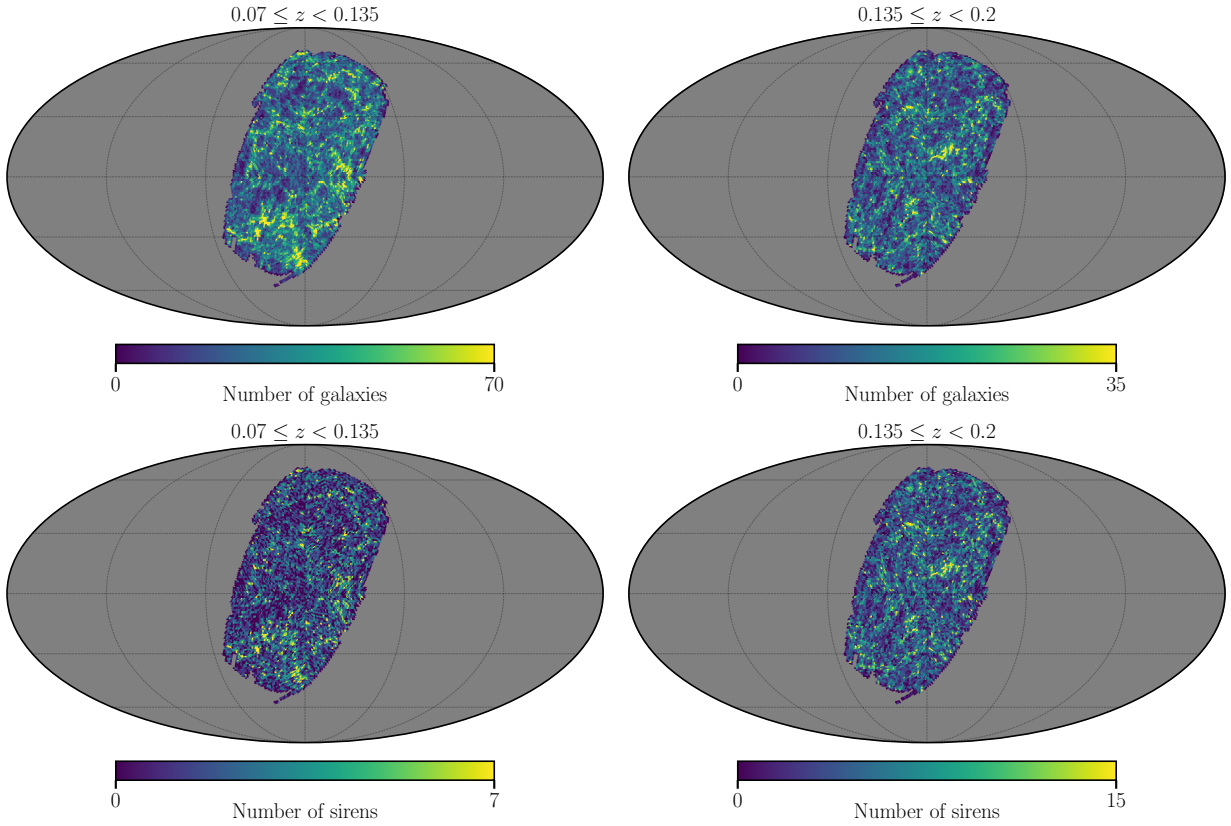


Figure 7.2: Upper plots show the counts map of the galaxy catalogs (obtained from safe0 file in SDSS DR7) in two different redshift bins $0.07 < z \leq 0.135$ and $0.135 \leq z < 0.2$. Lower plots show the map of randomly selected sirens from the galaxy catalog in the two same redshift bins (obtained from safe0 file in SDSS DR7). The color bars show the number of sources in each pixel. All maps are generated with Nside=64.

reproduced the galaxy power spectra to compare with their power spectrum, as shown in Figure H.1. After applying the aforementioned cuts to the actual data and random data catalogs, the count maps of both galaxies and sirens (the process of generating the mock siren catalog is explained in Section 7.4) in the two redshift bins $0.07 \leq z < 0.135$ and $0.135 \leq z < 0.2$ are shown in Figure 7.2. The number of galaxies in the first and second redshift bins is 237084 and 98734, respectively.

7.3.2 Physical properties of galaxies in SDSS DR7

As previously mentioned in Section 7.2.2, we select host galaxies based on their astrophysical properties, including SFR, stellar mass, and metallicity. In what follows, we briefly describe the procedure of obtaining the value of these parameters for galaxies in SDSS DR7 catalog.

- **Stellar mass in SDSS DR7:** In SDSS DR7, the stellar mass estimation involves the utilization of both observational data and theoretical models. Specifically, by acquiring the broadband photometry of the galaxies in the SDSS and 2MASS bands, ugrizJHK [4], one can measure the observed spectral energy density (SED) – which encompasses stellar ages, initial mass functions, and metallicities – and compare it with model-generated SED, which were obtained using stellar population synthesis models. By doing so, one can fine-tune various model parameters, such as the galaxies' age and metallicity, to identify the best-fit model. Finally, the stellar mass was measured using the best-fit SED model [445, 446]. Note that, we consider stellar masses in units of $10^{10} \times M_{\odot}/h^2$, considering a value of $h = 0.67$ which is favored from the ΛCDM fitting model to the cosmic microwave background (CMB) observations [11].

- **SFR in SDSS DR7:** The SFR is a crucial parameter that enables us to study star formation processes, galaxy evolution, and the cosmological history of star formation in a galaxy. It signifies the rate at which a galaxy generates new stars, usually measured in units of solar masses per year ($M_{\odot}$/yr). To calculate the SFR in the SDSS DR7 catalog, it is essential to start by defining two measured quantities provided in the catalog:

  - **B1000** measures the specific star formation rate (sSFR) over the 1 Gyr preceding the observed epoch (i.e., in the galaxy's rest-frame), providing valuable insights into the more recent star formation history and the evolutionary status of galaxies.
  - **INTSFH** is a parameter commonly used to quantify the cumulative star formation history of a galaxy. It represents the integrated star formation from the galaxy's formation up to the time it is observed (i.e., the observed redshift), rather than the present-day universe. Using these two parameters, we can calculate the SFR for each galaxy as [445, 446]

$$SFR = \frac{\text{B1000} \times \text{INTSFH}}{h^2 \times 10^9}. \tag{7.6}$$

- **Metallicity in the SDSS DR7 :** Metallicity measures the abundance of elements heavier than hydrogen and helium in a star or galaxy. It is intricately linked to the life cycles of stars, including their formation and evolution. In the SDSS DR7 catalog, the **METS** parameter represents the average metallicity of a group of objects. This measurement is obtained by comparing the observed spectrum of an object with a set of template spectra that have well-characterized properties. The metallicity is then estimated based on the best-fitting template spectrum.

[4]For more information see: http://sdss.physics.nyu.edu/vagc/kcorrect.html

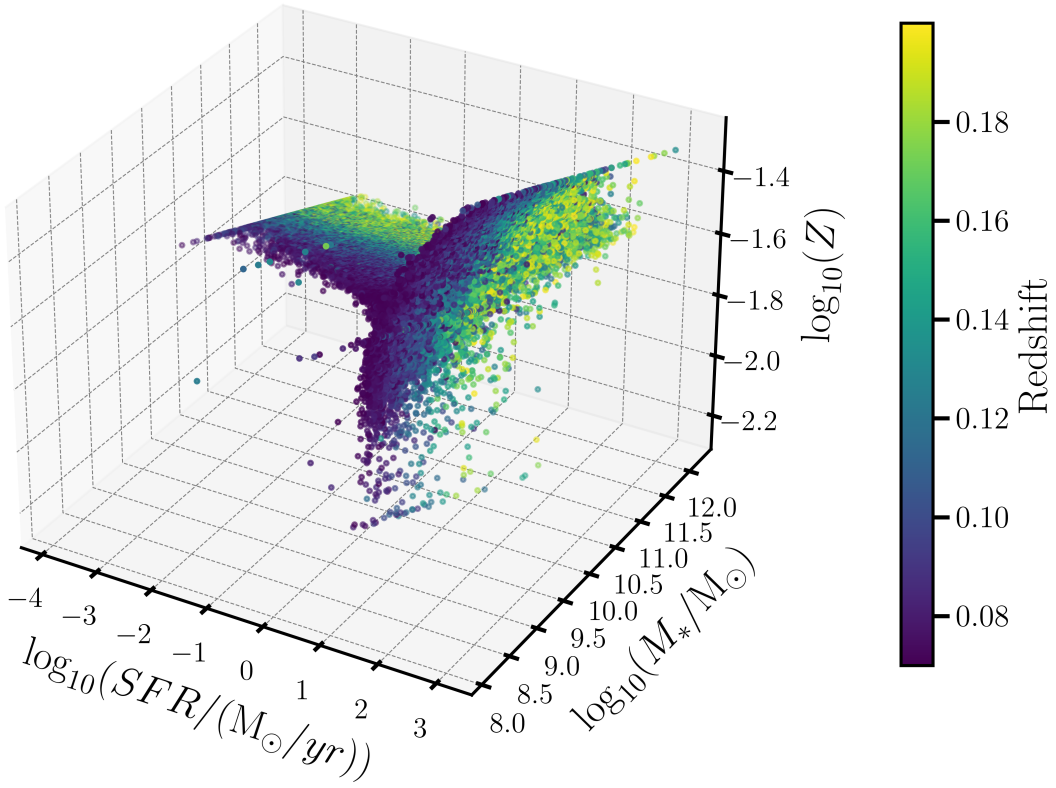


Figure 7.3: The distribution of stellar mass, SFR, and metallicity of the galaxies in the redshift range of $0.07 \leq z < 0.2$.

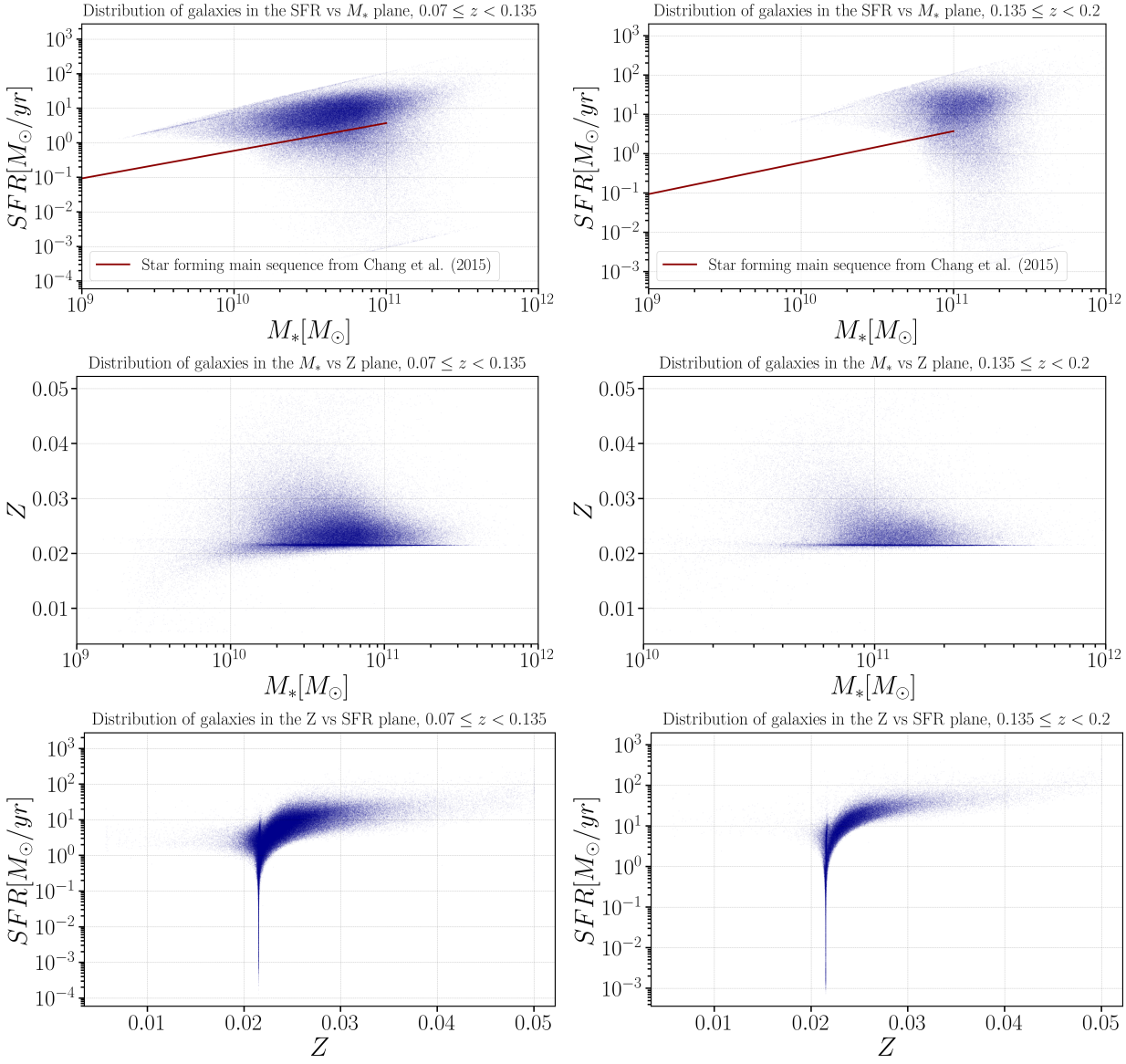


Figure 7.4: The 2D projected distribution of the SFR, stellar mass and metallicity of all galaxies in the SDSS DR7 catalogs within redshift of $0.07 \leq z < 0.135$ and $0.135 \leq z < 0.2$. The red lines in the upper panels represent the star-forming main sequence (SFMS), illustrating the relation between stellar mass and star formation rate for star-forming galaxies [447].

Figure 7.3 illustrates the distribution of SDSS DR7 galaxies as functions of stellar mass, SFR, and metallicity within the redshift range $0.07 \leq z < 0.2$. Figure 7.4 shows the 2D mass, SFR, and metallicity distribution of galaxies. In addition, Figure G.1 in Appendix G shows the 1D distribution of galaxies as a function of mass, SFR, and metallicity. As shown in all panels of Figure 7.4, especially in the upper plots, the galaxy sample is incomplete for stellar masses below $10^{10} M_{\odot}$ across both redshift ranges. According to [448], the minimum stellar mass required for a galaxy to be reliably detected and included in the sample at redshift $z < 0.06$ is $1.3 \times 10^{10} M_{\odot}$. Using the scaling relationship, the limiting stellar masses at $z = 0.07, 0.135, 0.2$ are 1.7×10^{10}, 6.3×10^{10}, 1.3×10^{11} $M_{\odot}$, respectively. Note that in this study, we use observed galaxy catalogs to construct the siren catalogs, which introduces limitations arising from the depth and completeness of the observational data. The upper plots also indicate that the sample is complete within SFR $\in [10^{-1}, 10^{1}]$. It is worth noting that our GW bias measurement results will be influenced by incompletenesses in stellar mass and SFR, which are discussed in greater detail in Sections 7.6.1 and 7.6.2. In the middle and bottom plots, we observed a sharp spike at a metallicity value of approximately $Z = 0.215$. This feature can be traced back to the underlying template-fitting methodology employed by k-correction [446]. The software fits a linear combination of five fixed galaxy templates to each galaxy's photometry. For galaxies that fall on the metallicity spike, we found that the fit is dominated almost entirely by the fourth template — with the corresponding coefficient being significantly nonzero while the others are nearly zero. According to Figure 4 in the kcorrect paper, the metallicity associated with this fourth template at the present-day lookback time is approximately $Z = 0.215$. Thus, the spike in the histogram does not indicate a failure of the fit, but rather reflects the discrete nature of the template set: in these cases, the fit selects a single dominant template whose metallicity is directly imprinted in the result.

7.4 Generating mock GW source catalogs from galaxy catalog

7.4.1 Binary merger rates and merger counts for binary black holes (BBHs)

The number of BBH mergers at a specific redshift determines the sample size of host galaxies, and, given that it is considerably smaller than the total number of galaxies, it does not strongly correlate with clustering properties to affect the GW bias parameter, even with extended observation periods. Nonetheless, it plays a crucial role in minimizing errors. Here, we outline the calculation method. However, in producing our siren catalog, discussed

in detail later, we do not change the parameters controlling the merger rate, as they do not impact our computations. Instead, we focus on optimizing the "observing times" to effectively model GW bias behavior. The BBH merger rate density, $R_{\rm GW}$ [283, 285–287, 325–329, 408] controlled by cosmic SFR density, $R_{\rm SFR}$, at the formation redshift, z_f, and the delay time probability distribution, $P(t_d)$, where t_d characterizes the delay time between the formation of stars, t_f, and the merger of the compact objects, t_m. This can be modeled by writing the merger rate density of GW sources at a given merger redshift of z_m, as

$$R_{\rm GW}(z_m) = \mathcal{A}_0 \int_{z_m}^{\infty} P_D(t_d) \frac{dt_f}{dz_f} R_{\rm SFR}(z_f)\ dz_f. \tag{7.7}$$

The redshift dependence of $R_{\rm SFR}(z_f)$ can be estimated analytically using the best-fit formula to observation [380] as

$$R_{\rm SFR}(z) = \psi(z) = 0.015 \frac{(1+z)^{2.7}}{1+[(1+z)/2.9]^{5.6}} M_\odot\ {\rm yr}^{-1}\ {\rm Mpc}^{-3}. \tag{7.8}$$

$\mathcal{A}_0$ in Eq. (7.7) denotes the normalization factor so that the local merger rate of gravitational wave sources at the present day redshift $R_{\rm GW}(z_m)|_{z_m=0}$ matches the observed value of $\mathcal{R}_{\rm BBH} = 23.9^{+14.3}_{-8.6}\ {\rm Gpc}^{-3}\ {\rm yr}^{-1}$ [410]. This implies that

$$\mathcal{A}_0 = \frac{\mathcal{R}_{\rm BBH}}{\int_0^{\infty} P_D(t_d) \frac{dt_f}{dz_f} R_{\rm SFR}(z_f)\ dz_f}.$$

The other parameter needed to evaluate the merger rate in Eq. (7.7) is the delay time probability $P_D(t_d)$. While there are some theoretical studies on this, the exact behavior of the delay time distribution is not well established [325, 329, 330, 332, 407, 408, 449]. In this work we model the delay time distribution as an inverse power-law function in t_d with a minimum threshold for delay time, $t_{d,\rm min}$,

$$P_D(t_d) \propto \Theta(t_d - t_{d,\rm min}) t_d^{-\kappa}. \tag{7.9}$$

Here κ is a non-negative free parameter and $\Theta(t)$ is the Heaviside function such that for $t_d < t_{d,\rm min}$, the probability distribution is zero. Lastly, the factor dt_f/dz in Eq.(7.7) is the Jacobian for change of variable from time coordinate time to redshift which can be calculated as

$$\frac{dt_f}{dz_f} = \frac{1}{H_0} \frac{1}{(1+z_f)E(z_f)}, \tag{7.10}$$

where

$$E(z) = \sqrt{\Omega_m(1+z)^3 + \Omega_r(1+z)^4 + \Omega_\Lambda + \Omega_K(1+z)^2}, \tag{7.11}$$

and we consider a cosmology with vanishing curvature and radiation density parameters (i.e. $\Omega_K,\ \Omega_r \simeq 0$). In addition, Ω_m, Ω_Λ and H_0 are taken from the data release of the Planck Collaboration for the Flat ΛCDM (Lambda Cold Dark Matter) [11].

Once the merger rate density is obtained, one can integrate it over the volume to compute the total number of mergers in a given redshift bin of width Δz and midpoint z_a,

$$N_{GW}(z_a) = T_{\rm obs} \int_{z_a-\Delta z/2}^{z_a+\Delta z/2} \frac{dV}{dz} \frac{R_{GW}(z)}{1+z}\, dz. \tag{7.12}$$

Here, the indices of the midpoint redshifts are labeled as $a \in \{1, \cdots, N_{\rm bins}\}$ where $N_{\rm bins}$ is the number of bins in the redshift range of the galaxy survey. $T_{\rm obs}$ denotes the observation time and depends on the running time of the GW detectors. As previously mentioned, the number of mergers does not alter our result regarding the GW bias. More specifically in our calculations, the delay time parameters solely affect the number of BBH mergers without impacting the GW-galaxy relationship. Consequently, we do not assess the effect of variation of the parameters κ and $t_{\rm d,min}$ defined in Eq. (6.8) on the merger rate and we will fix $\kappa = 1$ and $t_{d,min} = 500$ Myr, for all tests in this chapter. However, we will elaborate on this in the following subsection that the distribution of delay times is expected to have an indirect effect on the probability of host galaxies and hence GW bias. Although our merger rate model assumes shorter delay times that peaked around 500 Myr, we will also explore variations in host-galaxy probability that account for longer delay times by modifying the shape of the probability function, rather than explicitly using delay times.

7.4.2 Selecting host galaxies for BBH mergers using astrophysical host-galaxy probability functions

In Section 7.4.1, we computed the expected number of BBH mergers at each specific redshift. In what follows, we develop a general statistical framework to calculate the probability for galaxies to potentially host BBH mergers based on their astrophysical properties. We refer to these probability functions as *GW host-galaxy probability function*, or for short *host-galaxy probability function*. Note that, the formulation of an analytical expression for each individual host-galaxy probability function that accurately characterizes the distribution of BBH merger rates is an ongoing area of research. There are various studies about the relation between BBH merger rates and the astrophysical properties of their host galaxies [287, 327, 330, 331, 365, 414, 449, 450].

In general, the full joint host-galaxy probability function can be written as a joint probability distribution composed of individual host-galaxy probabilities. In this study, we consider three important astrophysical properties governing the host-galaxy probabilities: the

stellar mass M_*, the SFR, and the metallicity Z. For simplicity we assume the joint probability distribution $P_{\rm host}({\rm GW}|M, {\rm SFR}, Z)$ which we refer to as M_*-Z-SFR host-galaxy probability function, is a product of the host-galaxy probabilities corresponding to each of these properties individually[5],

$$P_{\rm host}({\rm GW}|M_*, {\rm SFR}, Z) = P_{\rm host}({\rm GW}|M_*) P_{\rm host}({\rm GW}|{\rm SFR}) P_{\rm host}({\rm GW}|Z). \tag{7.13}$$

We then model $P_{\rm host}({\rm GW}|M_*)$, $P_{\rm host}({\rm GW}|{\rm SFR})$, and $P_{\rm host}({\rm GW}|Z)$ as the individual host-galaxy probabilities associated with stellar mass, SFR, and metallicity, respectively. Of course, as these properties are not uncorrelated, it is expected that our result will have some degeneracy directions or planes in the parameter space. In the following, we formulate the stellar mass, SFR and metallicity host-galaxy probabilities to populate the BBHs through galaxies guided by the physics principals governing BBH formation and merger events.

M_*-Z-SFR host-galaxy probability function

M_* host-galaxy probability function, $P_{\rm host}({\rm GW}|M_*)$: The stellar mass of a galaxy at any given time is determined by its 'prior' stellar mass and star formation history. At the beginning of a galaxy's life, due to the hierarchical clustering and more frequent merging with other galaxies, the SFR is often higher, leading to galaxies accumulating more stellar mass by forming new stars. Therefore, since BBH merger rate is also expected to be correlated with the host galaxy's SFR during formation, it should also be positively correlated with stellar mass. In fact a strong positive correlation has been shown in population synthesis simulations such as [331]. Note that this relation can be affected by other factors like delay times between BH formation and BBH mergers. BBH mergers that occur at $z = [0 - 0.2]$ might have formed at early times (higher redshift for longer delay times) or late times (lower redshift for shorter delay times). If delay times on average are shorter than quenching timescales, BBH numbers at high stellar mass should be suppressed, since BBHs formed early would have already merged. This suppression is expected around $10^{11} M_\odot$ based on the empirical SFR-mass relation. The relationship suggests that for massive galaxies, after experiencing rapid star formation at high redshifts, they are quenched later in time [e.g., 417], leading to a bimodal stellar mass–SFR relation characterized by low-mass, high-SFR star-forming galaxies and high-mass, low-SFR quiescent galaxies [415]. However, if the average delay times are long, then there should not be a suppression of merger rates in massive galaxies as most are merging after the quenching.

Considering these factors, we use a broken power-law for the M_* host-galaxy probability function, reflecting the potential impact of suppressed SFR history in high-mass galaxies.

[5] In practice, we use `numpy.random.choice` for selecting host galaxies based on host-galaxy probabilities.

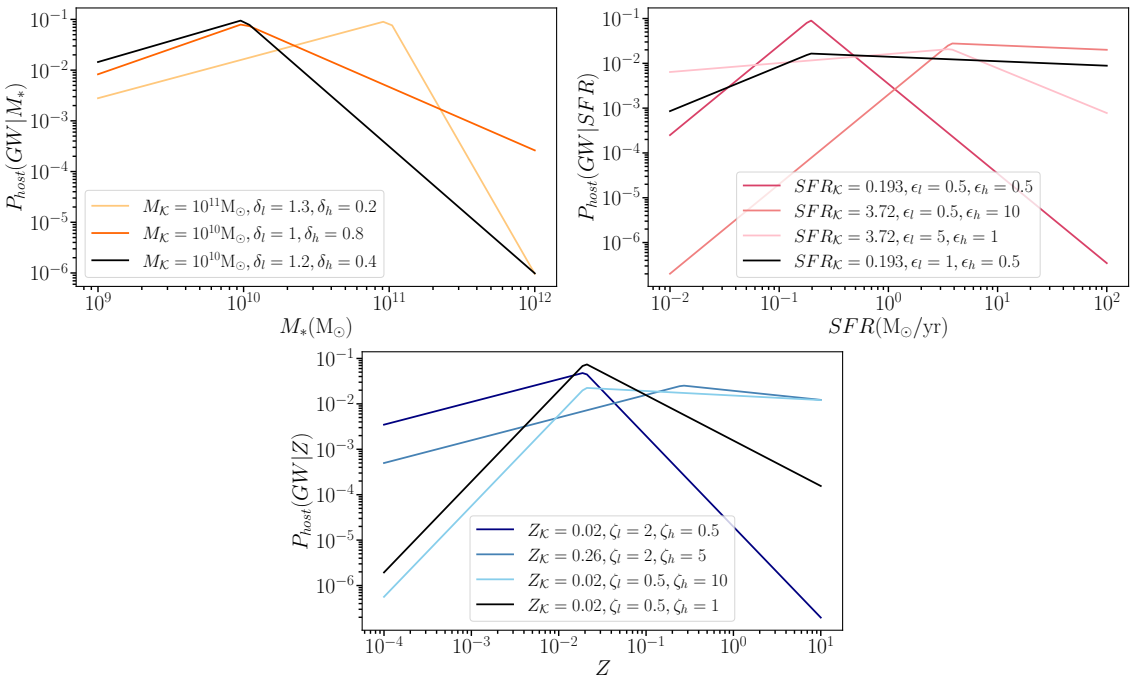


Figure 7.5: M_*, SFR and Z host-galaxy probability functions.

Note that, in the case of considering longer delay times, the absence of a suppression is automatically implemented by variables controlling the peak and slopes of the power-law to make the host-galaxy probability function a monotonically increasing function of stellar mass [327, 330, 331, 414]. The M_* host-galaxy probability function is parameterized as follows

$$P_{\rm host}(\mathrm{GW}|M_*) = \begin{cases} A_M 10^{(\log(M_*)-\log(M_{\mathcal{K}}))/\delta_l}, & \text{for } 7 \leq \log(M_*) \leq \log(M_{\mathcal{K}}), \\ A_M 10^{(\log(M_{\mathcal{K}})-\log(M_*))/\delta_h}, & \text{for } \log(M_*) \geq \log(M_{\mathcal{K}}), \end{cases} \tag{7.14}$$

where M_* is the stellar mass of the host galaxy obtained from the catalog and $M_{\mathcal{K}}$ sets the turn-around point. The slopes of the low and high M_* parts for the broken power-law are adjusted by δ_l and δ_h and A_M is the normalization factor. In this study, we consider many different parameters choices in the host-galaxy probability function i.e. selecting $M_{\mathcal{K}} \in [10^9, 10^{12}]\ M_\odot$, the set of values of $\delta_l \in \{0.5, 1, 1.5, 2, 10\}$ and $\delta_h \in \{0.5, 1.5, 2.5, 3.5, 4.5\}$ (See top left panel of Fig. 7.5 for few examples).

SFR host-galaxy probability function, $P_{\rm host}(\mathrm{GW}|\mathrm{SFR})$: As previously mentioned, there is good evidence indicating a significant correlation between the probability of a galaxy undergoing a merger and its stellar mass since it is a strong indicator of the galaxy's star formation history. However, as the relationship is not one-to-one, further insight can be obtained by introducing an additional probabilistic dependence to the SFR at the merger

event. The merger rate of BBHs, may be tied to the SFR of their host galaxies, with different behavior at low and high SFRs. In early-type, low-SFR galaxies, the merger rate is influenced by the cumulative star formation history rather than ongoing star formation. The compact binaries formed during those early, intense star-forming periods (regardless of the value ongoing SFR) may or may not be merging depending on the delay times. As a result, the merger rate in low-SFR galaxies correlates more strongly with stellar mass than SFR, since their merger history reflects past stellar populations. In late-type, high-SFR galaxies, active star formation leads to the formation of massive stars, which evolve into compact binaries. These galaxies are efficient at forming BBHs, but due to the delay times, BBHs formed in these galaxies remain unmerged. At intermediate SFRs, the merger rate is shaped by both ongoing star formation and the legacy of past star formation, underscoring the complex relationship between SFR, stellar mass, and merger rates. The transition from low- to high-SFR reflects a shift in merger rate dominance, from early-type galaxies to late-type galaxies. This transition is a consequence of the interplay between cosmic star formation, galaxy assembly, and the delay time distribution of BBHs mergers [449]. We model this interplay in the SFR host-galaxy probability function as

$$P_{\rm host}({\rm GW}|{\rm SFR}) = \begin{cases} A_{\rm SFR}\, 10^{(\log({\rm SFR})-\log({\rm SFR}_{\mathcal{K}}))/\epsilon_l}, & \text{for } \log({\rm SFR}) \leq \log({\rm SFR}_{\mathcal{K}}), \\ A_{\rm SFR}\, 10^{(\log({\rm SFR}_{\mathcal{K}})-\log({\rm SFR}))/\epsilon_h}, & \text{for } \log({\rm SFR}) > \log({\rm SFR}_{\mathcal{K}}), \end{cases} \tag{7.15}$$

where $\mathrm{SFR}_{\mathcal{K}}$, and SFR are respectively the pivot SFR scale, and the measured SFR for each galaxy is extracted from the catalog. ϵ_l and ϵ_h control the growing and decaying regime of the probability and $A_{\rm SFR}$ is also the normalization factor. In this study, we consider $\mathrm{SFR}_{\mathcal{K}} \in [10^{-2}, 10]\ M_\odot/\mathrm{yr}$, the set of values for $\epsilon_l \in \{0.5, 1, 2, 5, 10\}$ and $\epsilon_h \in \{1, \infty\}$, where $\epsilon_h = \infty$ means the the decaying slope is zero.

Z host-galaxy probability function, $P_{\rm host}({\rm GW}|Z)$: Metallicity may also influence the formation and consequently the merger rates of BBHs by affecting the evolution of their massive star progenitors. In low-metallicity environments, typically found at earlier cosmic times, reduced stellar-wind mass loss allows massive stars to retain more mass, leading to the formation of heavier black holes. This process improves the efficiency of BBH formation but if the delay time is long, by the time they merge, the host galaxies may have become larger and more metal-rich[287]. Studies, such as those by [409] and [411], indicate that low-metallicity progenitors are often associated with shorter delay times between formation and merger, resulting in a higher incidence of BBH mergers at lower redshifts. On the other hand if delay times are short, in higher-metallicity environments prevalent at later cosmic times, increased mass loss through stellar winds reduces the masses of progenitor

stars, thereby decreasing the BBH formation rates and subsequently the merger rate. This interplay between metallicity, star formation, and galaxy evolution accounts for the observed variations in BBH merger rates across different epochs, as discussed in [414] and [416]. For more detail see [331, 414, 415]. In this chapter, analogous to stellar mass and SFR, we represent the influence of metallicity by a broken power law for the Z host-galaxy probability function as following,

$$P_{\rm host}({\rm GW|Z}) = \begin{cases} A_Z 10^{(\log(Z)-\log(Z_{\mathcal{K}}))/\zeta_l}, & \text{for } -4 \leq \log(Z) \leq \log(Z_{\mathcal{K}}), \\ A_Z 10^{(\log(Z_{\mathcal{K}})-\log(Z))/\zeta_h}, & \text{for } \log(Z) > \log(Z_{\mathcal{K}}), \end{cases} \tag{7.16}$$

where $Z_{\mathcal{K}}$ is the pivot point and Z is the metallicity of galaxies read from the catalog. A_Z is the normalization factor. In this study, we consider $Z_{\mathcal{K}} \in [0.01, 0.1]$, and the set of values of $\zeta_l \in \{0.5, 1, 2, 5, 10\}$, and $\zeta_h \in \{0.5, 1, 2, 5, 10\}$.

Finally, by taking the product of all three host-galaxy probability functions defined above and directly extracting the values of SFR, Z, and M_* from the catalog (without relying on observational scaling relationships), we can calculate the overall probability of host galaxies based on their combined astrophysical properties using Eq. (7.13). This provides a very direct and comprehensive strategy for the selection of host galaxies such that the complex underlying physics of the BH-galaxy relation is encoded in the shape of these three host-galaxy probability functions. By varying the parameters governing their shape, we are able to cover different BH formation scenarios, and investigate how much GW bias changes for different parameters. Furthermore, this formulation also allows us to consider scenarios in which only one of these three astrophysical properties are important. In particular, we also consider a scenario in which we marginalize over SFR and Z, populating the BBHs in the catalog solely based on $P_{\rm host}({\rm GW}|M_*)$, to compare the result with Chapter 6.

7.5 3D power spectrum estimation

7.5.1 Measuring the anisotropic power spectrum and covariance

In this section, we measure the 3D power spectrum for the sample from the SDSS DR7 catalog described in Section 7.3. The power spectra of the siren and galaxy catalogs are computed using `pypower` [6] (from the `Cosmodesi` package), which provides functions and classes for computing and analyzing the power spectrum of the large-scale structure in the universe using DESI data[7]. It is a software tool designed for the calculation of galaxy auto-

[6] https://pypower.readthedocs.io/en/latest/

[7] https://github.com/cosmodesi/pypower

and cross-power spectra, along with the estimation of associated window functions. `pypower` can calculate the power spectra of both periodic boxes and of realistic survey geometries, operating in either flat-sky or plane-parallel RSD configurations, and incorporating first-order odd wide-angle corrections when computing the window-convolved power spectrum from a galaxy survey. More specifically, we used the `CatalogFFTPower` algorithm, which takes the galaxy catalog (with position coordinates provided in terms of right ascension, declination, and redshift) as an input and provides the associated power spectrum multipoles using the Yamamoto estimator [451]. This algorithm requires the redshift-dependent number density for the data catalog[8]. We first calculate the redshift histogram of the randoms catalog (using the `RedshiftHistogram` algorithm) and then adjust the redshift distribution to match the data catalog. Next, we should also account for the "FKP weight" which refers to the weight assigned to each galaxy or object in the catalog based on the FKP (Feldman, Kaiser, and Peacock) method [452]. We compute the corresponding FKP weight using the following formula

$$w_{FKP} = \frac{1}{1 + n(z) \times P_0}, \tag{7.17}$$

where $P_0 = 10^4 h^{-3}$ Mpc^3. With these ingredients in place, one can proceed to calculate the 3D power spectrum using the `CatalogFFTPower` algorithm. Specifically, because of redshift-space distortions, isotropy is broken in the observed power spectrum and it is therefore a function of both wavenumber k and angle to the line-of-sight, μ, or can be equivalently expressed in the basis of Legendre multipoles of the line-of-sight. More details are provided in the Pypower[9] and Nbodykit[10] documentations.

The window-convolved power spectrum covariance is calculated analytically using `thecov` software [453, 454].[11] We compute the window-convolved Gaussian contribution to the covariance, using the appropriate randoms for the SDSS DR7 geometry. We find good agreement between the analytic covariance and the mock-based covariances from the SDSS DR7 clustering analysis of [381].

As the last remark of this section, to calculate the power spectra of sirens, as previously mentioned, we first need to generate the siren catalog from the galaxy catalog. To do so, we first calculate the number of mergers in each redshift bin using Eq. (7.12). Note that since the number of sirens is much smaller than galaxies in each bin for an order of one year or even a decade of observing time, we multiply the number of sirens by an amplitude to

[8]Note that, the SDSS DR7 galaxy sample has data and random files. Data files contain the actual observed data of galaxies. Random files are used to correct for observational biases and to account for the survey's selection functions.

[9]https://pypower.readthedocs.io/en/latest/

[10]https://nbodykit.readthedocs.io/en/latest/cookbook/convpower.html

[11]https://github.com/cosmodesi/thecov

obtain a meaningful theoretical prediction. We then aim to populate the galaxy catalog with potential host for binary mergers using M_* and M_*-Z-SFR host-galaxy probably models mentioned in section (7.4.2). The amplitude values are varied depending on the choice of $M_{\mathcal{K}}$, $SFR_{\mathcal{K}}$, or $Z_{\mathcal{K}}$, reflecting the diversity in stellar mass, star formation rate (SFR), and metallicity distributions among galaxies. The selection of the amplitude thresholds is carefully calibrated to ensure that the number of selected standard sirens remains below the total number of galaxies and also prevent shot-noise domination and to avoid the trivial case in which all galaxies are designated as sirens. With the siren catalog in hand, we can then proceed to compute the corresponding 3D power spectrum. We also compute the theoretical power spectrum as explained in Section 7.5.2 below. Having both observationally-measured and theoretically-measured power spectra, we are able to find the best-fit value of the GW bias parameter.

7.5.2 Measuring the GW bias from the power spectrum multipoles

Galaxies are observed in redshift space, and their observed positions are influenced by peculiar velocities and gravitational infall in addition to cosmological expansion. This causes two types of anisotropic clustering along the line of sight: apparent elongation due to random motion on small scales (Fingers of God (FoG)) and large-scale squashing along the line-of-sight (Kaiser effect) due to galaxy motions toward over-densities (see Ref. [455]). Taking into account the redshift space distortion due to the Kaiser effect, the power spectrum in redshift space can be shown as [267]

$$P^s_{XX}(k,\mu) = P^r_{XX} + 2\mu^2 P^r_{X\theta} + \mu^4 P^r_{\theta\theta}, \tag{7.18}$$

where subscript $X = g$ refers to all the galaxies and $X = GW$ is the case of GW host galaxies, while the s, r superscripts indicate the redshift-space and real-space field, respectively. θ is the divergence of the peculiar velocity field, and μ is the cosine of the angle between wave vector k and the line of sight. Using $\delta_X = b_X \delta_m$ and $\theta = -f\delta_m$ in the linear regime, we find

$$P_{XX}(k,\mu) = b_X^2 P_m + 2\mu^2 b_X f P_m + \mu^4 f^2 P_m\,, \tag{7.19}$$

where b_X is the galaxy(or GW sources) bias, f is the growth rate, and P_m is the matter power spectrum. Note that, this does not yet include the small-scale nonlinear redshift space distortion (*i.e.* the Fingers of God). Accounting for small-scale redshift-space distortions (FoG effect), $P(k,\mu)$ is commonly modeled as [267][12]

$$P_{XX}(k,\mu) = (b_X^2 P_m + 2\mu^2 b_X f P_m + \mu^4 f^2 P_m)(1 + \frac{k^2\mu^2\sigma_p^2}{2})^{-2}, \tag{7.20}$$

[12]This approximation may fail in quasi linear regime.

where σ_p, the velocity dispersion parameter is estimated to be $\sigma_p = 400$ km s^{-1} based on constraints from SDSS MGS [456, 457]. It is convenient to decompose $P_{XX}(k,\mu)$ into the Legendre basis including the monopole ($P_{XX,0}$), quadrupole ($P_{XX,2}$), and hexadecapole ($P_{XX,4}$). The FoG term affects the damping of the power spectrum, particularly the quadrupole-to-monopole ratio. Higher σ_p leads to stronger suppression of anisotropies, while lower values cause overestimation. We tested variations around 400 km s^{-1} to assess their impact on GW bias estimates. The signal-to-noise on the hexadecapole is poor, so we will only consider the monopole and quadrupole hereafter. Then $P_{XX}(k,\mu)$ weighted by the corresponding Legendre polynomial for μ from 0 to 1, the monopole and quadrupole of the above power spectrum can be calculated by

$$P_{XX,0}(k) = \int (b_X^2 P_m + 2\mu^2 b_X f P_m + \mu^4 f^2 P_m)(1 + \frac{k^2\mu^2\sigma_p^2}{2})^{-2} d\mu. \tag{7.21}$$

$$P_{XX,2}(k) = \int \frac{1}{2}(3\mu^2 - 1)(b_X^2 P_m + 2\mu^2 b_X f P_m + \mu^4 f^2 P_m)(1 + \frac{k^2\mu^2\sigma_p^2}{2})^{-2} d\mu. \tag{7.22}$$

in which X can again be either galaxies or GW sources. In what follows, to calculate GW bias or galaxy bias, we first theoretically compute the monopole and quadropole power spectrum (Eq. (7.21) and Eq. (7.22)) of galaxies using the dark matter power spectra (measured by the Code for Anisotropies in the Microwave Background, `CAMB`). In the next step, we measure the corresponding poles of the observed power spectra for both galaxy and GW sources catalog (details were explained in sec 7.5.1). We then use χ^2 model to find the value of the GW bias parameter that provides the best fit between the theoretically computed power spectrum and the observed one.

7.6 Discussion on the GW bias parameter results

In this section, we show our results for the GW bias. As a sanity check, we first test the framework by randomly selecting sirens from galaxy samples rather than using any host-galaxy probability. The details are shown in the Appendix I. We then use the M_* host-galaxy probability (see 7.6.1) and M_*-Z-SFR (see 7.6.2) to populate the GW sources from the galaxy catalog, calculate the GW bias, and study its dependency on the parameters of the host-galaxy probability function.

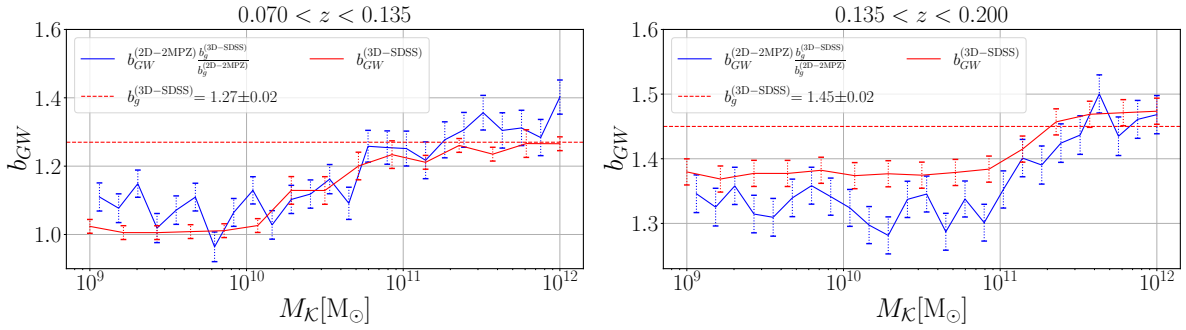


Figure 7.6: Comparison between our results and those of Chapter 6. Here we compare the bias computed using angular galaxy clustering in 2MPZ to the bias computed using 3D galaxy power in SDSS. We use two different redshift bins, $0.07 \leq z < 0.135$ and $0.135 \leq z < 0.2$. The 2MPZ bias (blue line) is rescaled to match the overall galaxy bias of SDSS (red line) as explained in the text. The agreement between GW hosts populated in 2MPZ and SDSS is good, up to small differences in the second redshift bin at $M_{\mathcal{K}} < 10^{11}\ M_\odot$ due to differences in the galaxy survey completeness. Here we use a GW host-galaxy probability function which depends on M_* only with $\delta_l = 4$, and $\delta_h = 0.5$.

7.6.1 GW bias parameter exclusively derived from M_* host-galaxy probability (M_* model)

In order to compare our results with those of Chapter 6 (published in [2]), we start with a simpler model where the GW host-galaxy probability function only depends on the stellar mass, to match the model used in [2]. In other words, we only consider $P(\mathrm{GW}|M*)$ in Eq (7.13) and marginalize over SFR, and Z. When populating sirens according to this model, Fig 7.6 shows a good agreement between the computed GW bias from the 3D power spectrum using sirens populated SDSS DR7, and the 2D power spectrum sirens populated in 2MPZ (following the methods of [2]). Note that the average galaxy biases are different between SDSS and 2MPZ, with biases of 1.27 and 1.34 for SDSS and 2MPZ, respectively, in the first bin, and 1.45 and 1.78 for SDSS and 2MPZ, respectively, in the second bin. Due to this offset in the galaxy biases, we scale the GW bias from 2MPZ by the ratio of galaxy biases from the two catalogs, which is shown in Fig. 7.6. Overall, the GW bias obtained from the two methods are in good agreement in the 1-2σ range. The increase in GW bias parameter with $M_{\mathcal{K}}$ (also described in [2]) is due to the fact that for lower values of $M_{\mathcal{K}}$, the host-galaxy probability function will pick up less massive galaxies, which have smaller halo masses and thus lower biases.

7.6.2 Results for the GW bias parameter using the entire M_*-Z-SFR model

GW bias dependency on the parameters of $P(\mathrm{GW}|M_*)$

We start by studying the GW bias dependence on the parameters of $P(\mathrm{GW}|M_*)$, i.e., δ_l, δ_h, and $M_{\mathcal{K}}$ while fixing the parameters corresponding to $P(\mathrm{GW}|\mathrm{SFR})$, and $P(\mathrm{GW}|Z)$ (ϵ_l, ϵ_h, $\mathrm{SFR}_{\mathcal{K}}$ and ζ_l, ζ_h, $Z_{\mathcal{K}}$, see bottom panel of Fig. 7.7) and show how b_{GW} varies as a function of $M_{\mathcal{K}}$, for some selections of δ_l and δ_h. The results are shown in Fig. 7.7. The list of findings are

- **General behavior:** Even though $P(\mathrm{GW}|\mathrm{SFR})$, and $P(\mathrm{GW}|Z)$ are not uniform the over all trends are still consistent with the results of [2] and M_* model in Fig. 7.6, which show increasing $M_{\mathcal{K}}$ increases GW bias. This is the consequence of the relationship between stellar mass and halo mass [352, 354], i.e., galaxies with more stellar mass live in more massive and biased dark matter halos. In all the panels of Fig 7.6, the b_{GW}-$M_{\mathcal{K}}$ lines show a similar pattern. At lower values of $M_{\mathcal{K}}$, GW bias is almost constant up to a point where GW bias starts growing. However, at high stellar mass values GW bias is almost constant again.

  The constant b_{GW} at low $M_{\mathcal{K}}$ is due to incompleteness at the low-mass end of the stellar mass distribution in observations. At the low-mass end, once $M_{\mathcal{K}}$ drops far below the peak of stellar mass distribution, further lowering $M_{\mathcal{K}}$ has less effect on the selection, i.e. galaxies are chosen with the same power law slope above $M_{\mathcal{K}}$, as there are few galaxies below $M_{\mathcal{K}}$. This regime approximately corresponds to $M_{\mathcal{K}} < 10^{10}\ M_{\odot}$ in the first redshift bin and $2 \times 10^{10}\ M_{\odot}$ in the second bin. These are within the estimated stellar mass completeness limit extrapolated from the results of [448], $1.7 \times 10^{10}\ M_{\odot}$ at $z = 0.07$ and $6.2 \times 10^{10}\ M_{\odot}$ at $z = 0.135$. We therefore emphasize that this feature at the lower $M_{\mathcal{K}}$ is purely an observational effect based on our selection of GW sources from a stellar mass-limited sample, and does not correspond to a physical transition in the relationship between GW bias and $M_{\mathcal{K}}$.

  The constant b_{GW} at the high-mass end corresponds to the transition scale at $\hat{M}_\star$ in the Schechter function between the power-law low mass regime and the exponentially dropping high-mass end of the mass function. Once $M_{\mathcal{K}}$ exceeds $\hat{M}_\star$, the number of galaxies above the pivot mass is exponentially suppressed and the bias asymptotes to a constant value. This transition occurs at about $2 \times 10^{11}\ M_{\odot}$ in both redshift bins; this is close to $\hat{M}_\star = 1.3 \times 10^{11}\ M_{\odot}$ [458].

- **Effect of different $\mathrm{SFR}_{\mathcal{K}}$ and $Z_{\mathcal{K}}$ on the GW bias:** We test the dependence on SFR and metallicity by showing the results for three pairs of $\mathrm{SFR}_{\mathcal{K}}$ and $Z_{\mathcal{K}}$ values: $\mathrm{SFR}_{\mathcal{K}} = 0.193\ M_{\odot}/\mathrm{yr}$ and $Z_{\mathcal{K}} = 0.26$; $\mathrm{SFR}_{\mathcal{K}} = 3.72\ M_{\odot}/\mathrm{yr}$ and $Z_{\mathcal{K}} = 0.26$; and $\mathrm{SFR}_{\mathcal{K}} = 3.72\ M_{\odot}/\mathrm{yr}$ and $Z_{\mathcal{K}} = 0.025$. The values for $\mathrm{SFR}_{\mathcal{K}}$ and $Z_{\mathcal{K}}$ are chosen such that they lie in the regions with different SFR and Z distributions. This way, we explore two different regimes of SFR and Z dependency in the host-galaxy probability function. Changing $\mathrm{SFR}_{\mathcal{K}}$ can impact the b_{GW}-$M_{\mathcal{K}}$ relationship (compare the upper and middle panels in Fig. 7.7). The shape of b_{GW} as a function of $M_{\mathcal{K}}$ is nearly the same, but increasing $\mathrm{SFR}_{\mathcal{K}}$ leads to an overall lower bias. As $SFR_{\mathcal{K}}$ increases, sirens increasingly populate galaxies with higher SFR, which tend to have lower stellar mass and thus smaller galaxy bias. In contrast, changing the metallicity parameter $Z_{\mathcal{K}}$ has almost no impact (compare the middle and bottom panels in Fig. 7.7) since metallicity is weakly correlated with stellar mass.

- **Effect of slopes for M_* selection on GW bias:** An increase in δ_h (compare dark solid red lines and light solid red lines in Fig 7.6) results in a shallower decaying slope in the host-galaxy probability function for galaxies with $M_* > M_{\mathcal{K}}$, less preference for selecting galaxies with lower mass, and subsequently a higher GW bias. However, changing δ_h is important only in the low-$M_{\mathcal{K}}$ regime. This is because if $M_{\mathcal{K}}$ is above the mean of stellar mass distribution, changing δ_h changes the host-galaxy probability for fewer number of galaxies with masses above $M_{\mathcal{K}}$, and therefore has minimal impact on the GW bias. In contrast, δ_l plays the opposite role, with slightly higher bias for lower δ_l values in the high-$M_{\mathcal{K}}$ part (compare solid and dashed lines of the same color in Fig 7.6).

- **Redshift dependency:** Furthermore, the analysis revealed a dependency of the GW bias on redshift, with higher values of b_{GW} at the upper redshift range ($0.135 \leq z < 0.2$). This is similar to the findings of [2] which is the dependence of GW bias on redshift mostly reflects the trend observed in the redshift dependence of galaxy bias, meaning that GW bias increases as redshift increases. This is a purely observational effect due to increasing incompleteness at higher redshifts, which preferentially picks out higher-mass and thus higher-bias galaxies. This observation seems to hold true for all selections of host-galaxy probabilities, but there can be some mild dependence in redshift evolution of b_{GW}/b_g depending on the parameters of host-galaxy probability. We leave this for future studies.

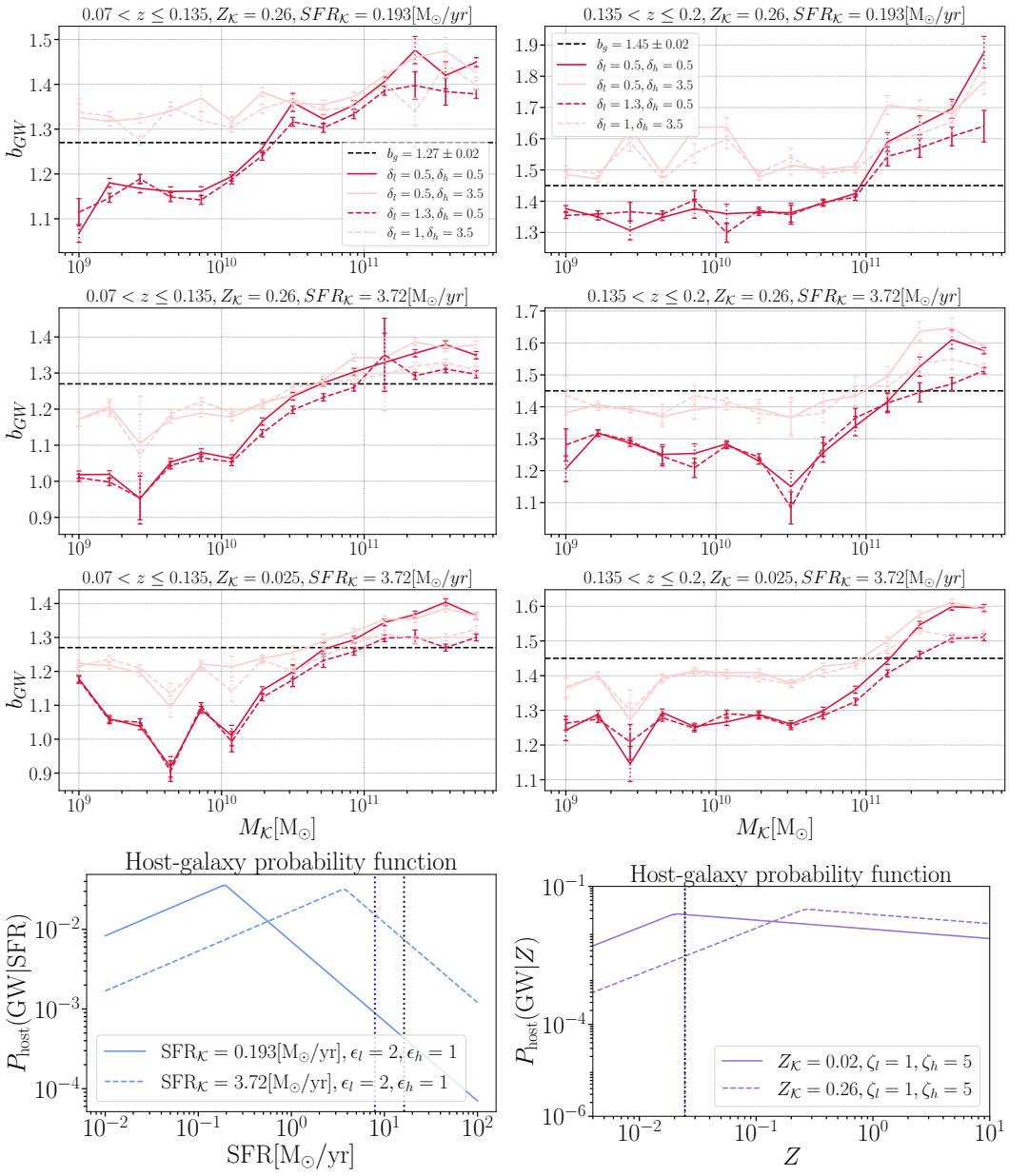


Figure 7.7: GW bias parameter as a function of stellar mass break $M_{\mathcal{K}}$ within redshift $0.07 \leq z < 0.135$ (left) and $0.135 \leq z < 0.2$ (right). Within each plot, we show the impact of changing the upper (δ_h) and lower (δ_l) slopes of $P(\mathrm{GW}|M_*)$. The rows show the impact of changing the break in star formation rate, $\mathrm{SFR}_{\mathcal{K}}$, and metallicity, $Z_{\mathcal{K}}$, changing to the values indicated in the plot titles. The slopes of the SFR and metallicity portions of the GW host-galaxy probability are fixed at $\epsilon_l = 2$, $\epsilon_h = 1$, $\zeta_l = 1$, $\zeta_h = 5$. The last row shows fixed $P(\mathrm{GW}|\mathrm{SFR})$, and $P(\mathrm{GW}|Z)$ functions used here. In the bottom-left plot, the dotted vertical blue and black lines show the mean SFR of the galaxies ($\{7.9, 16.28\}$ $[M_\odot/\mathrm{yr}]$ respectively) in the first and second redshift bins. In the bottom-right plot, the blue and black vertical dotted lines show the mean metallicity of galaxies ($\{0.0245, 0.0246\}$, respectively) in the first and second redshift bins.

GW bias dependency on the parameters of $P(\mathrm{GW}|\mathrm{SFR})$

In this section, we look into the impact of varying the shape of $P(\mathrm{GW}|\mathrm{SFR})$ i.e., ϵ_l, ϵ_h and $\mathrm{SFR}_{\mathcal{K}}$ for several distinct but fixed configurations of $P(\mathrm{GW}|M_*)$, and $P(\mathrm{GW}|Z)$ which is for specific choices of δ_l, δ_h, $M_{\mathcal{K}}$, ζ_l, ζ_h, and $Z_{\mathcal{K}}$, see bottom panel of Fig. 7.8). We vary the parameters of $P(\mathrm{GW}|\mathrm{SFR})$ i.e., ϵ_l, ϵ_h and $\mathrm{SFR}_{\mathcal{K}}$. The results are shown in Fig. 7.8 which illustrates the following points:

- **General behavior:** Based on our discussion in section (7.4.2) in low redshifts we expect the merger events to be more probable in galaxies with low SFR. This is supported by the findings of the population synthesis simulations of [449], where the power-law fit of the host-galaxy probability for SFR has a negative power of around -0.051. Therefore, for analysis, we focus on varying the values of $\mathrm{SFR}_{\mathcal{K}}$ on the lower end, which means mainly assessing the impact of the negative to flat slopes of $P(\mathrm{GW}|\mathrm{SFR})$. Consequently the general trend as expected is for $\epsilon_h = \infty$, (which describes host selection always dominated by a uniform SFR selection [13] for $\mathrm{SRF} > \mathrm{SFR}_{\mathcal{K}}$) b_{GW} does not vary much (dashed and solid light blue lines in Fig. 7.8). On the other hand when we set $\epsilon_h = 1$ (i.e. linearly decaying) as $\mathrm{SFR}_{\mathcal{K}}$ increases, b_{GW} decreases. This is because for low $\mathrm{SFR}_{\mathcal{K}}$, the selection of host galaxies is also dominated by those with low SFR. However, as it increases to higher values, galaxies with higher SFR begin to be selected. Galaxies with higher SFR observed today tend to have lower stellar mass, and are thus less clustered. As SFR decreases, galaxies transition from low-mass, low bias, highly star-forming galaxies to higher mass, higher bias, quenched galaxies.

- **Effect of different $M_{\mathcal{K}}$, and $Z_{\mathcal{K}}$ on the GW bias:** Overall, GW bias seems to have a stronger correlation with $M_{\mathcal{K}}$, but it does not show a significant correlation with $Z_{\mathcal{K}}$. Although the values of $Z_{\mathcal{K}}$ that we have considered are each at the extreme end of the distribution. The impact of $M_{\mathcal{K}}$ is more distinct when we consider $\epsilon_h = \infty$ or uniform SFR host selection as in this case similar to Fig. 7.7, at lower values of $M_{\mathcal{K}}$ (i.e. 10^{10} $M_\odot$ in the top two panels of Fig. 7.8), b_{GW} is less than the average galaxy bias. This causes the nearly constant trends at high ϵ_h to have a lower b_{GW} (solid light blue line). However, when $M_{\mathcal{K}}$ increases to 10^{11} $M_\odot$ which is close to the mean stellar mass of the galaxies(bottom panel of Fig. 7.8), b_{GW} approaches the average value of the galaxy bias.

[13] Note in our parametrization the slope is $-1/\epsilon_h$.

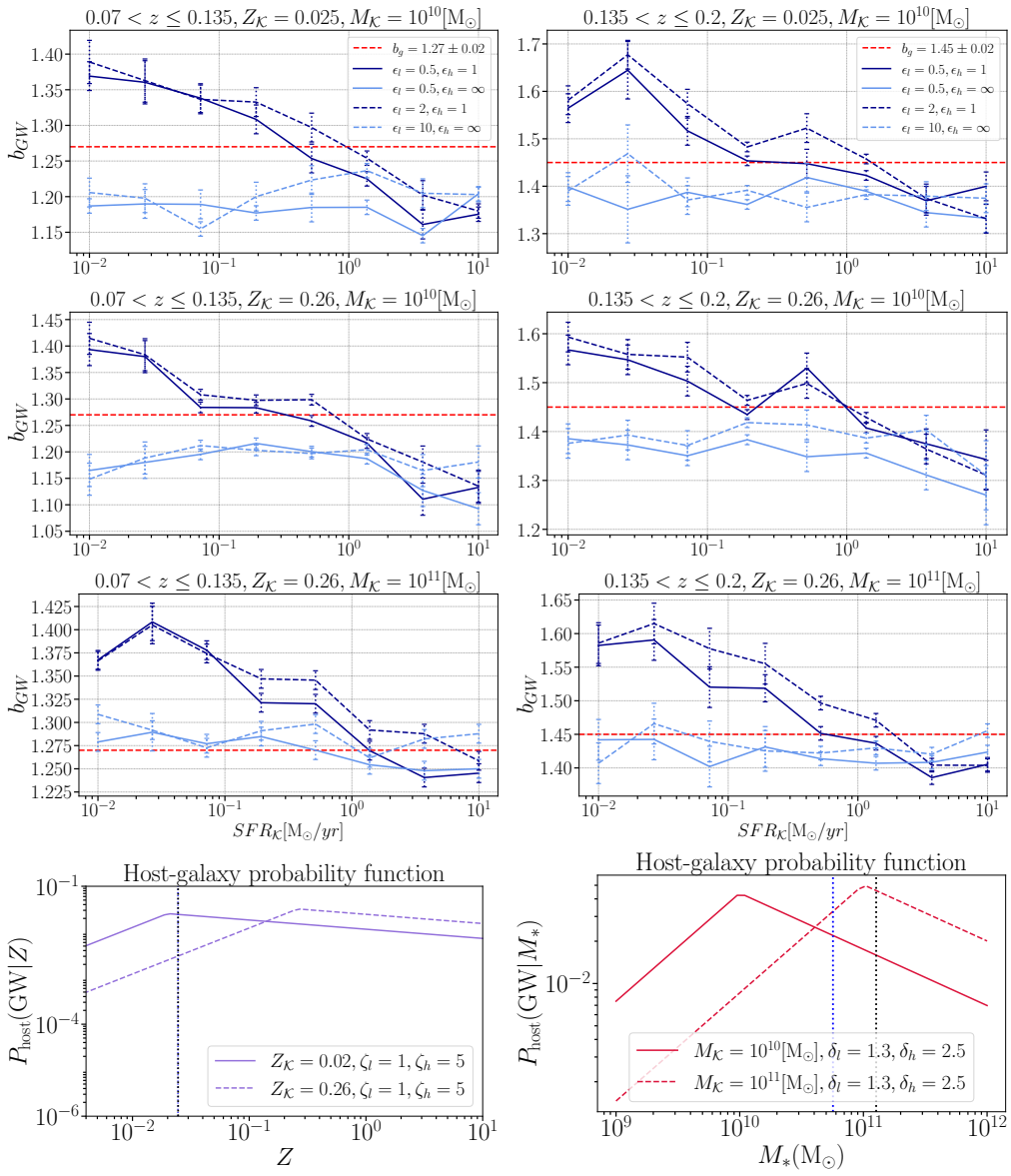


Figure 7.8: GW bias parameter as a function of star-formation rate turnover, $\mathrm{SFR}_{\mathcal{K}}$, within redshift $0.07 \leq z < 0.135$ (left) and $0.135 \leq z < 0.2$ (right). Within each panel, the different line styles and colors indicate different values of the SFR slopes ϵ_l and ϵ_h. The solid blue lines correspond to $\epsilon_l = 0.5$ and $\epsilon_h = 1$, and the solid light blue line indicates a model where $\epsilon_l = 0.5$ but the inverse of the falling slope, ϵ_h, goes to ∞ (i.e. there is no SFR-dependence in the host-galaxy probability function above the turnover). The dashed lines display the impact of changing ϵ_l, compared to solid lines of the same color. Different rows of panels show the impact of varying the transition points for stellar mass $M_{\mathcal{K}}$ and the metallicity $Z_{\mathcal{K}}$. We have fixed the other parameters to $\delta_l = 1.3$, $\delta_h = 2.5$, $\zeta_l = 1$, and $\zeta_h = 5$. The last row shows fixed $P(\mathrm{GW}|M_*)$, and $P(\mathrm{GW}|Z)$ functions used here. In the bottom-left plot, the dotted vertical blue and black lines show the mean metallicity of the galaxies ($\{0.0245, 0.0246\}$ respectively) in the first and second redshift bins. In the bottom-right plot, the blue and black vertical dotted lines show the mean stellar mass of galaxies ($\{5.67, 12.68\}$ $[10^{10} M_\odot]$ respectively) in the first and second redshift bins.

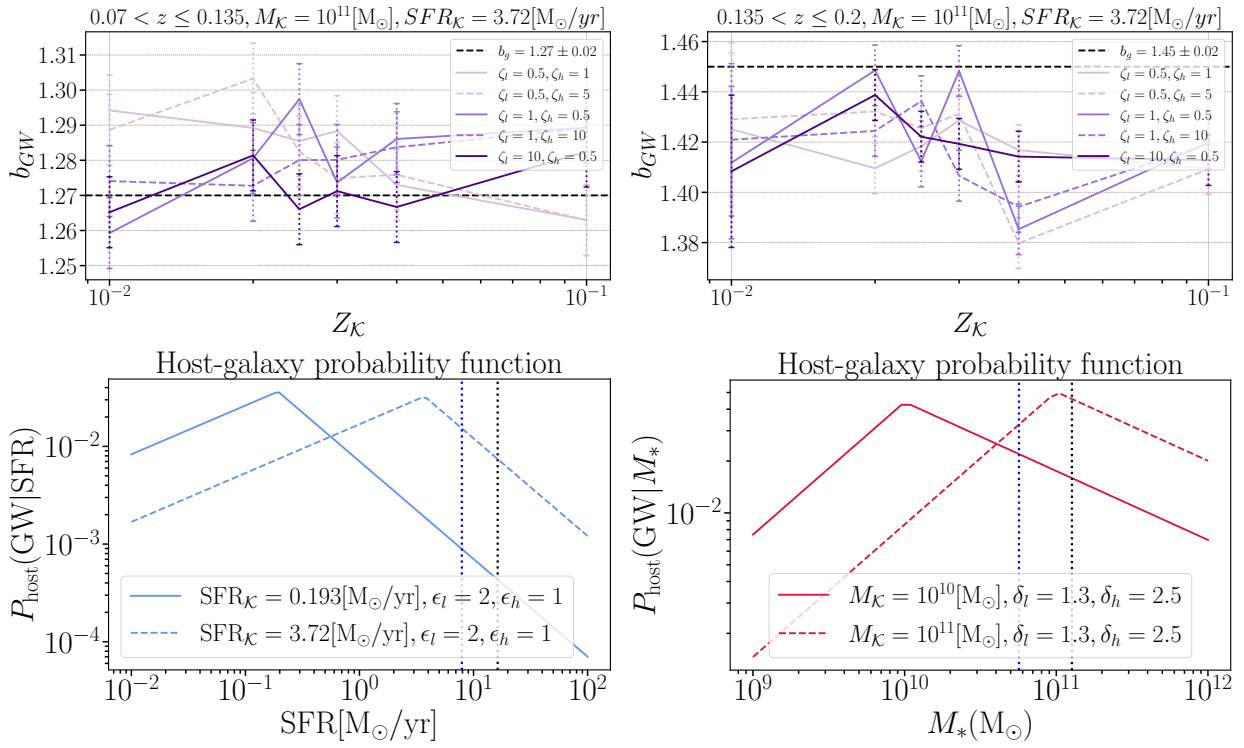


Figure 7.9: GW bias parameter as a function of $Z_{\mathcal{K}}$ for the $0.07 \leq z < 0.135$ (left) and $0.135 \leq z < 0.2$ (right) bins. Here we only consider a single value of $M_{\mathcal{K}}$ and $\mathrm{SFR}_{\mathcal{K}}$ as the plots are similar when these change. The different colors and linestyles show the impact of changing the metallicity slopes ζ_l and ζ_h. Note that, $\delta_l = 1.3, \delta_h = 2.5, \epsilon_l = 2, \epsilon_h = 1$. The second row shows fixed $P(\mathrm{GW}|M_*)$, and $P(\mathrm{GW}|\mathrm{SFR})$ functions used here

GW bias dependency on the parameters of $P(\mathrm{GW}|Z)$

Last, we fix parameters of $P(\mathrm{GW}|M_*)$, and $P(\mathrm{GW}|\mathrm{SFR})$ (i.e., δ_h, δ_l, $M_{\mathcal{K}}$ and ϵ_l,ϵ_h, $SFR_{\mathcal{K}}$, see bottom panel of Fig. 7.9) and vary parameters of $P(\mathrm{GW}|Z)$ (i.e., ζ_l, ζ_h, $Z_{\mathcal{K}}$). Figure 7.9 shows the GW bias parameter across different values of $Z_{\mathcal{K}}$. We set $\mathrm{SFR}_{\mathcal{K}} = 3.72$ $[M_\odot/\mathrm{yr}]$, and $M_{\mathcal{K}} = 10^{11}[M_\odot]$. As can be seen, the bias exhibits some fluctuations within both redshift ranges, however, the results show an insignificant correlation between GW bias and metallicity. The various line styles and colors represent different slopes considered in the M_*-Z-SFR model (ζ_l and ζ_h), and show that the GW bias is not very sensitive to these slopes either. Our result suggests that the GW bias is primarily sensitive to stellar mass, and next SFR and but the least to metallicity in the M_*-Z-SFR model.

7.7 Conclusion

This study presents the very first attempts in modeling and calculating the GW bias parameter for BBHs. This research, as a complement to our previous study [2], further develops the modeling and computation of the GW bias parameter for simulated BBH (siren) catalogs based on spectroscopic galaxy surveys. In particular, we created a phenomenological model by expressing host-galaxy probabilities through the astrophysical characteristics of galaxies, specifically their stellar mass, star formation rate, and metallicity (M_*-Z-SFR model). We then generated siren catalogs from the original SDSS DR7 galaxy catalog. Subsequently, we evaluated the best linear fit for gravitational wave bias parameter, b_{GW}, for different selections of the parameters in (M_*-Z-SFR) model, by fitting the power spectrum form the mocked catalog to the theoretical siren power spectrum derived from CAMB. This allowed us to assess the impact and sensitivity of the GW bias parameter to various astrophysical model parameters for BBH merger events, particularly in relation to the astrophysical properties of their host galaxies, such as stellar mass, star formation rate (SFR), and metallicity.

Notably, our result indicated that the stellar mass correlates more strongly with the GW bias, reflecting the influence of the stellar mass-halo mass relationships. Our analysis also suggest relationship between gravitational wave (GW) bias and star formation rate (SFR). We observed that when host-galaxy probability has preference in selecting high-SFR galaxies, it induces a notable decrease in GW bias. This intriguing phenomenon can be attributed to the characteristic of galaxies with higher SFRs, which tend to exhibit lower stellar masses. As galaxies with enhanced SFR are often associated with vigorous ongoing star formation, they typically harbor younger stellar populations and consequently possess lower stellar masses. Meanwhile the dependency of gravitational wave (GW) bias on the metallicity of the host galaxy did not show a significant correlation. We wish to emphasize that the stellar mass, star formation rate (SFR), and metallicity of galaxies are not independent properties, and a scaling relationship among them is indeed anticipated. Therefore, even though we concentrated solely on a discrete set of parameters within the 9-dimensional parameter space related to the shape of the host-galaxy probability, a comprehensive exploration might reveal some degenerate surfaces in this space. Another prospective avenue for future research could involve the computation of the gravitational wave (GW) bias parameter for binary neutron stars or primordial black holes.

Chapter 8

Conclusion

In this thesis, we aimed to present a comprehensive overview of the cosmos. First, in an attempt to explain the origin of the universe, we introduced cosmological perturbation theory in Chapter 2. This framework assumes a homogeneous and isotropic universe with small perturbations and is particularly powerful because it provides observational quantities that can be directly compared with data. We then demonstrated how early universe models can be constructed to align with current observational constraints. Beyond these constraints, we argued that theoretical considerations — most notably, the avoidance of singularities — play a crucial role in evaluating such models. In light of this criterion, bouncing cosmology models are proposed in which there is a contracting phase before the expanding universe.

Formulating a viable bouncing cosmology is generally challenging. First, the contracting phase must smooth out inhomogeneities while simultaneously generating a scale-invariant power spectrum — two objectives that are difficult to achieve in single-field bouncing models. Additionally, a smooth transition between contracting and expanding phases is required, which is generally non-trivial and can lead to instabilities. To address these challenges, we propose a bouncing model known as the Cuscuton bounce in Chapter 3. The Cuscuton bounce is realized within a modified gravity theory of the same name, which introduces a scalar field into the action without adding new degrees of freedom. We demonstrated that this model successfully avoids space-time singularities and achieves a smooth, stable bounce without violating the null energy condition in the matter sector. Furthermore, it can generate scale-invariant entropy perturbations, which are later converted into curvature perturbations. We then extended the perturbation analysis to third order, aiming to demonstrate that the model remains perturbatively well-behaved and does not exhibit strong coupling issues. Additionally, we calculated the non-Gaussianities produced in this model for comparison with current observational constraints. While our analysis does not reveal a significant non-Gaussianity signal, additional non-Gaussianity could arise during the conversion process,

potentially reaching levels comparable to current observational bounds.

Bouncing cosmology, and in particular the Cuscuton model, has recently garnered significant interest (e.g., [251–257]). In principle, these models can be tested and distinguished using future CMB experiments. Several upcoming CMB experiments are either already under construction, actively collecting data, or planned for launch in the near future. Broadly, these experiments fall into three categories: (I) Ground-based experiments (AdvACT, Polarbear, CLASS, Keck/BICEP3, Simons Array, and SPT3G), (II) Balloon-borne experiments (EBEX 10k, ExploreR (Piper), and Spider), and (III) Satellite missions (CMBPol, COrE, PIXIE, PRISM, and LiteBIRD) [40–43]. Notably, while this thesis was being written, the Atacama Cosmology Telescope (ACT) released new results, which remain consistent with the standard ΛCDM model [459, 460]. We believe the field of primordial cosmology has a promising future, with an influx of new data expected in the coming years. In particular, the detection of CMB B-mode polarization—an indicator of primordial tensor perturbations—or the discovery of specific features in the primordial power spectrum could mark the beginning of a new era in cosmology.

Having developed a fairly successful theory to explain the generation of small perturbations in the early universe, a natural question arises: how have these perturbations evolved throughout cosmic history to form the large inhomogeneities we observe today, such as galaxies and stars? We addressed this question in Chapter 4, where we discussed the formation of large-scale structures. In particular, we demonstrated that the statistical distribution of dark matter halos and galaxies within nonlinear evolution can be characterized by some bias parameters. The set of tracers that map the matter distribution extends beyond dark matter halos and galaxies, allowing for the definition of new bias parameters for gravitational sources. Gravitational waves, explored in Chapter 5, provide a powerful observational window, offering access to objects that are otherwise unobservable through electromagnetic signals.

In Chapter 6, we investigated the dependence of the gravitational-wave (GW) bias parameter for binary black holes (BBHs) on the GW host-galaxy probability using angular power spectrum measurements of mock sirens in galaxy catalogs. By modeling BBH host-galaxy probabilities with a broken power law in stellar mass, the study explores how the turnover point, $M_{\mathcal{K}}$, affects the GW bias. This is linked to the transition between star-forming and quiescent galaxies and how long the delay time is compared to the quenching time scale. In Chapter 7, we extended this analysis to the 3D power spectrum using spectroscopic galaxy surveys. These surveys provide information on each galaxy's star formation rate (SFR), and metallicity in addition to stellar mass. Accordingly, we updated the GW host-galaxy probability model to depend on all three parameters, making it more realistic.

Our results remained consistent with previous findings and additionally revealed that GW bias is generally higher for galaxies with low SFR, as they tend to be older, more massive, and more clustered. Furthermore, we found no significant correlation between GW bias and metallicity. Even though this approach, which uses observed galaxy catalogs rather than simulations, is limited by catalog incompleteness and observational constraints, it provides flexibility and larger volumes. This methodology offers a novel way to connect observed galaxies to GW sources, enabling future studies to refine GW bias models and improve cosmological parameter inference, particularly for the Hubble constant and dark energy.

To conclude this thesis, we believe that cosmology may have entered a golden era, driven by a wealth of upcoming experiments and observational data. With the advent of artificial intelligence (AI) tools, processing these vast datasets has become more feasible than ever, enabling deeper insights and more precise analyses [461]. In the near future, galaxy catalogs from DESI [433], Euclid [434], the Vera Rubin Observatory [435], and the Roman Space Telescope [436] will provide unprecedented insights into the universe's large-scale structure. Additionally, next-generation gravitational wave detectors such as Cosmic Explorer, the Einstein Telescope, and LISA [293, 294, 296, 297] promise exciting new discoveries, particularly in the study of extreme objects like supermassive black holes and strong gravitational fields. These observations could even reveal possible modifications to general relativity. As an example of this rapidly evolving field, while this thesis was being written, new data from DESI was released, providing updated constraints on neutrino masses and, more importantly, strong evidence for a dynamical dark energy component, as opposed to a simple cosmological constant [462]. Any deviation from the standard model of particle physics, Einstein's theory of gravity, or the ΛCDM model that emerges from these new experiments could open an entirely new window for both theoretical and observational cosmology.

References

[1] Amir Dehghani, Ghazal Geshnizjani, and Jerome Quintin. Cuscuton bounce beyond the linear regime: Bispectrum and strong coupling constraints, 2025.

[2] Amir Dehghani, J. Leo Kim, Dorsa Sadat Hosseini, Alex Krolewski, Suvodip Mukherjee, and Ghazal Geshnizjani. The gravitational wave bias parameter from angular power spectra: Bridging between galaxies and binary black holes, 2024.

[3] D. S. Hosseini, A. Dehghani, J. L. Kim, A. Krolewski, S. Mukherjee, and G. Geshnizjani. Gravitational Wave Bias from 3D Power Spectrum: Tracing Binary Black Hole Mergers Through Galaxy Surveys. , In Preparation.

[4] Yi Wang. MathGR: a tensor and GR computation package to keep it simple. *arXiv e-prints*, 6 2013.

[5] Astropy Collaboration. Astropy: A community Python package for astronomy. *aap*, 558:A33, October 2013.

[6] Astropy Collaboration. The Astropy Project: Building an Open-science Project and Status of the v2.0 Core Package. , 156(3):123, September 2018.

[7] Astropy Collaboration. The Astropy Project: Sustaining and Growing a Community-oriented Open-source Project and the Latest Major Release (v5.0) of the Core Package. , 935(2):167, August 2022.

[8] Origins of the expanding universe: 1912-1932, 2013.

[9] D. J. Fixsen. The temperature of the cosmic microwave background. *The Astrophysical Journal*, 707(2):916–920, November 2009.

[10] N. Aghanim et al. Planck 2018 results. I. Overview and the cosmological legacy of Planck. *Astron. Astrophys.*, 641:A1, 2020.

[11] N. Aghanim et al. Planck 2018 results. VI. Cosmological parameters. *Astron. Astrophys.*, 641:A6, 2020. [Erratum: Astron.Astrophys. 652, C4 (2021)].

[12] Hideo Kodama and Misao Sasaki. Cosmological Perturbation Theory. *Prog. Theor. Phys. Suppl.*, 78:1–166, 1984.

[13] David Langlois. Inflation, quantum fluctuations and cosmological perturbations, 2004.

[14] Juan Maldacena. Non-gaussian features of primordial fluctuations in single field inflationary models. *Journal of High Energy Physics*, 2003(05):013–013, May 2003.

[15] V. Mukhanov. *Physical Foundations of Cosmology*. Cambridge University Press, Oxford, 2005.

[16] Viatcheslav F. Mukhanov, H. A. Feldman, and Robert H. Brandenberger. Theory of cosmological perturbations. Part 1. Classical perturbations. Part 2. Quantum theory of perturbations. Part 3. Extensions. *Phys. Rept.*, 215:203–333, 1992.

[17] Richard Arnowitt, Stanley Deser, and Charles W. Misner. Republication of: The dynamics of general relativity. *General Relativity and Gravitation*, 40(9):1997–2027, August 2008.

[18] Jerome Quintin. *Topics in Pre-Big Bang Cosmology*. PhD thesis, McGill U., 2019.

[19] D. H. Lyth. Large-scale energy-density perturbations and inflation. *Phys. Rev. D*, 31:1792–1798, Apr 1985.

[20] Eric Gourgoulhon. 3+1 formalism and bases of numerical relativity, 2007.

[21] Rishabh Jha. Introduction to hamiltonian formulation of general relativity and homogeneous cosmologies. *SciPost Physics Lecture Notes*, August 2023.

[22] N. D. Birrell and P. C. W. Davies. *Quantum Fields in Curved Space*. Cambridge Monographs on Mathematical Physics. Cambridge University Press, Cambridge, UK, 1982.

[23] Sean M. Carroll. *Spacetime and Geometry: An Introduction to General Relativity*. Cambridge University Press, 7 2019.

[24] Ulf H. Danielsson. Note on inflation and trans-planckian physics. *Physical Review D*, 66(2), July 2002.

[25] T. S. Bunch and P. C. W. Davies. Quantum Field Theory in de Sitter Space: Renormalization by Point Splitting. *Proc. Roy. Soc. Lond. A*, 360:117–134, 1978.

[26] Daniel Babich, Paolo Creminelli, and Matias Zaldarriaga. The shape of non-gaussianities. *Journal of Cosmology and Astroparticle Physics*, 2004(08):009–009, August 2004.

[27] Y. Akrami et al. Planck 2018 results. IX. Constraints on primordial non-Gaussianity. *Astron. Astrophys.*, 641:A9, 2020.

[28] Leonardo Senatore, Kendrick M Smith, and Matias Zaldarriaga. Non-gaussianities in single field inflation and their optimal limits from the wmap 5-year data. *Journal of Cosmology and Astroparticle Physics*, 2010(01):028–028, January 2010.

[29] Paolo Creminelli, Alberto Nicolis, Leonardo Senatore, Max Tegmark, and Matias Zaldarriaga. Limits on non-gaussianities from wmap data. *Journal of Cosmology and Astroparticle Physics*, 2006(05):004–004, May 2006.

[30] Juan Martin Maldacena. Non-Gaussian features of primordial fluctuations in single field inflationary models. *JHEP*, 05:013, 2003.

[31] Xingang Chen. Primordial Non-Gaussianities from Inflation Models. *Adv. Astron.*, 2010:638979, 2010.

[32] Yi Wang. Inflation, Cosmic Perturbations and Non-Gaussianities. *Commun. Theor. Phys.*, 62:109–166, 2014.

[33] Jerome Martin, Christophe Ringeval, and Vincent Vennin. Encyclopædia Inflationaris: Opiparous Edition. *Phys. Dark Univ.*, 5-6:75–235, 2014.

[34] Jérôme Martin, Christophe Ringeval, Roberto Trotta, and Vincent Vennin. The best inflationary models after planck. *Journal of Cosmology and Astroparticle Physics*, 2014(03):039–039, March 2014.

[35] Jinn-Ouk Gong. Multi-field inflation and cosmological perturbations. *International Journal of Modern Physics D*, 26(01):1740003, January 2017.

[36] Sergei D. Odintsov, Vasilis K. Oikonomou, Ifigeneia Giannakoudi, Fotis P. Fronimos, and Eirini C. Lymperiadou. Recent advances in inflation. *Symmetry*, 15(9):1701, September 2023.

[37] Y. Akrami et al. Planck 2018 results. X. Constraints on inflation. *Astron. Astrophys.*, 641:A10, 2020.

[38] P. A. R. Ade et al. Improved Constraints on Primordial Gravitational Waves using Planck, WMAP, and BICEP/Keck Observations through the 2018 Observing Season. *Phys. Rev. Lett.*, 127(15):151301, 2021.

[39] Josquin Errard, Stephen M. Feeney, Hiranya V. Peiris, and Andrew H. Jaffe. Robust forecasts on fundamental physics from the foreground-obscured, gravitationally-lensed CMB polarization. , 2016(3):052, March 2016.

[40] Paolo Creminelli, Diana López Nacir, Marko Simonović, Gabriele Trevisan, and Matias Zaldarriaga. Detecting primordial B-modes after Planck. , 2015(11):031, November 2015.

[41] M. C. Guzzetti, N. Bartolo, M. Liguori, and S. Matarrese. Gravitational waves from inflation. *Riv. Nuovo Cim.*, 39(9):399–495, 2016.

[42] E. Allys et al. Probing Cosmic Inflation with the LiteBIRD Cosmic Microwave Background Polarization Survey. *PTEP*, 2023(4):042F01, 2023.

[43] J. A. Sobrin et al. The Design and Integrated Performance of SPT-3G. *Astrophys. J. Supp.*, 258(2):42, 2022.

[44] E. Komatsu, D. N. Spergel, and B. D. Wandelt. Measuring primordial non-gaussianity in the cosmic microwave background. *The Astrophysical Journal*, 634(1):14–19, November 2005.

[45] Xingang Chen, Min-xin Huang, Shamit Kachru, and Gary Shiu. Observational signatures and non-Gaussianities of general single field inflation. *JCAP*, 01:002, 2007.

[46] David Wands. Local non-gaussianity from inflation. *Classical and Quantum Gravity*, 27(12):124002, May 2010.

[47] Justin Khoury, Burt A. Ovrut, Paul J. Steinhardt, and Neil Turok. Ekpyrotic universe: Colliding branes and the origin of the hot big bang. *Physical Review D*, 64(12), November 2001.

[48] Sergei Dubovsky, Thomas Grégoire, Alberto Nicolis, and Riccardo Rattazzi. Null energy condition and superluminal propagation. *JHEP*, 2006(3), March 2006.

[49] Claudia de Rham and Scott Melville. Unitary null energy condition violation in P(X) cosmologies. *Phys. Rev. D*, 95(12):123523, June 2017.

[50] Damien A. Easson, Ignacy Sawicki, and Alexander Vikman. G-bounce. *JCAP*, 2011(11):021, November 2011.

[51] Taotao Qiu, Jarah Evslin, Yi-Fu Cai, Mingzhe Li, and Xinmin Zhang. Bouncing Galileon cosmologies. *JCAP*, 2011(10):036, October 2011.

[52] Arvind Borde, Alan H. Guth, and Alexander Vilenkin. Inflationary Spacetimes Are Incomplete in Past Directions. *Phys. Rev. Lett.*, 90(15):151301, April 2003.

[53] Ghazal Geshnizjani, Eric Ling, and Jerome Quintin. On the initial singularity and extendibility of flat quasi-de Sitter spacetimes. *JHEP*, 10:182, 2023.

[54] Robert M. Wald. *General Relativity*. Chicago Univ. Pr., Chicago, USA, 1984.

[55] Alan H. Guth. The Inflationary Universe: A Possible Solution to the Horizon and Flatness Problems. *Phys. Rev. D*, 23:347–356, 1981.

[56] Andrei D. Linde. A New Inflationary Universe Scenario: A Possible Solution of the Horizon, Flatness, Homogeneity, Isotropy and Primordial Monopole Problems. *Phys. Lett. B*, 108:389–393, 1982.

[57] Viatcheslav F. Mukhanov and G. V. Chibisov. Quantum Fluctuations and a Nonsingular Universe. *JETP Lett.*, 33:532–535, 1981.

[58] James M. Bardeen, Paul J. Steinhardt, and Michael S. Turner. Spontaneous Creation of Almost Scale-Free Density Perturbations in an Inflationary Universe. *Phys. Rev. D*, 28:679, 1983.

[59] Robert Brandenberger. Do we have a Theory of Early Universe Cosmology? *Stud. Hist. Phil. Sci. B*, 46:109–121, 2014.

[60] Anna Ijjas, Paul J. Steinhardt, and Abraham Loeb. Inflationary paradigm in trouble after Planck2013. *Phys. Lett. B*, 723:261–266, 2013.

[61] Anna Ijjas, Paul J. Steinhardt, and Abraham Loeb. Inflationary schism. *Phys. Lett. B*, 736:142–146, 2014.

[62] Prateek Agrawal, Georges Obied, Paul J. Steinhardt, and Cumrun Vafa. On the Cosmological Implications of the String Swampland. *Phys. Lett. B*, 784:271–276, 2018.

[63] Alice Di Tucci, Job Feldbrugge, Jean-Luc Lehners, and Neil Turok. Quantum Incompleteness of Inflation. *Phys. Rev. D*, 100(6):063517, 2019.

[64] Alek Bedroya, Robert Brandenberger, Marilena Loverde, and Cumrun Vafa. Trans-Planckian Censorship and Inflationary Cosmology. *Phys. Rev. D*, 101(10):103502, 2020.

[65] Arvind Borde, Alan H. Guth, and Alexander Vilenkin. Inflationary space-times are incomplete in past directions. *Phys. Rev. Lett.*, 90:151301, 2003.

[66] Daisuke Yoshida and Jerome Quintin. Maximal extensions and singularities in inflationary spacetimes. *Class. Quant. Grav.*, 35(15):155019, 2018.

[67] D. Battefeld and Patrick Peter. A Critical Review of Classical Bouncing Cosmologies. *Phys. Rept.*, 571:1–66, 2015.

[68] Robert Brandenberger and Patrick Peter. Bouncing Cosmologies: Progress and Problems. *Found. Phys.*, 47(6):797–850, 2017.

[69] M. Gasperini, M. Maggiore, and G. Veneziano. Towards a nonsingular pre-big bang cosmology. *Nucl. Phys. B*, 494:315–330, 1997.

[70] M. Gasperini, Massimo Giovannini, and G. Veneziano. Perturbations in a nonsingular bouncing universe. *Phys. Lett. B*, 569:113–122, 2003.

[71] M. Gasperini, Massimo Giovannini, and G. Veneziano. Cosmological perturbations across a curvature bounce. *Nucl. Phys. B*, 694:206–238, 2004.

[72] Jerome Quintin, Robert H. Brandenberger, Maurizio Gasperini, and Gabriele Veneziano. Stringy black-hole gas in α'-corrected dilaton gravity. *Phys. Rev. D*, 98(10):103519, 2018.

[73] Olaf Hohm and Barton Zwiebach. Duality invariant cosmology to all orders in α'. *Phys. Rev. D*, 100(12):126011, 2019.

[74] Peng Wang, Houwen Wu, Haitang Yang, and Shuxuan Ying. Non-singular string cosmology via α' corrections. *JHEP*, 10:263, 2019.

[75] Peng Wang, Houwen Wu, Haitang Yang, and Shuxuan Ying. Construct α' corrected or loop corrected solutions without curvature singularities. *JHEP*, 01:164, 2020.

[76] Jerome Quintin, Heliudson Bernardo, and Guilherme Franzmann. Cosmology at the top of the α' tower. *JHEP*, 07:149, 2021.

[77] M. Gasperini and G. Veneziano. Non-singular pre-big bang scenarios from all-order α' corrections. *JHEP*, 07:144, 2023.

[78] Ghanashyam Date and Golam Mortuza Hossain. Genericity of big bounce in isotropic loop quantum cosmology. *Phys. Rev. Lett.*, 94:011302, 2005.

[79] Parampreet Singh, Kevin Vandersloot, and G. V. Vereshchagin. Non-singular bouncing universes in loop quantum cosmology. *Phys. Rev. D*, 74:043510, 2006.

[80] Jaume de Haro. Does loop quantum cosmology replace the big rip singularity by a non-singular bounce? *JCAP*, 11:037, 2012.

[81] Edward Wilson-Ewing. The Matter Bounce Scenario in Loop Quantum Cosmology. *JCAP*, 03:026, 2013.

[82] Edward Wilson-Ewing. Ekpyrotic loop quantum cosmology. *JCAP*, 08:015, 2013.

[83] Yi-Fu Cai and Edward Wilson-Ewing. Non-singular bounce scenarios in loop quantum cosmology and the effective field description. *JCAP*, 03:026, 2014.

[84] Edward Wilson-Ewing. The loop quantum cosmology bounce as a Kasner transition. *Class. Quant. Grav.*, 35(6):065005, 2018.

[85] Niayesh Afshordi, Daniel J. H. Chung, and Ghazal Geshnizjani. Cuscuton: A Causal Field Theory with an Infinite Speed of Sound. *Phys. Rev. D*, 75:083513, 2007.

[86] Niayesh Afshordi, Daniel J. H. Chung, Michael Doran, and Ghazal Geshnizjani. Cuscuton Cosmology: Dark Energy meets Modified Gravity. *Phys. Rev. D*, 75:123509, 2007.

[87] Jishnu Bhattacharyya, Andrew Coates, Mattia Colombo, A. Emir Gumrukcuoglu, and Thomas P. Sotiriou. Revisiting the cuscuton as a Lorentz-violating gravity theory. *Phys. Rev. D*, 97(6):064020, 2018.

[88] Henrique Gomes and Daniel C. Guariento. Hamiltonian analysis of the cuscuton. *Phys. Rev. D*, 95(10):104049, 2017.

[89] Supranta S. Boruah, Hyung J. Kim, and Ghazal Geshnizjani. Theory of Cosmological Perturbations with Cuscuton. *JCAP*, 07:022, 2017.

[90] Chunshan Lin and Shinji Mukohyama. A Class of Minimally Modified Gravity Theories. *JCAP*, 10:033, 2017.

[91] Aya Iyonaga, Kazufumi Takahashi, and Tsutomu Kobayashi. Extended Cuscuton: Formulation. *JCAP*, 12:002, 2018.

[92] Shinji Mukohyama and Karim Noui. Minimally Modified Gravity: a Hamiltonian Construction. *JCAP*, 07:049, 2019.

[93] Xian Gao and Zhi-Bang Yao. Spatially covariant gravity theories with two tensorial degrees of freedom: the formalism. *Phys. Rev. D*, 101(6):064018, 2020.

[94] Katsuki Aoki, Francesco Di Filippo, and Shinji Mukohyama. Non-uniqueness of massless transverse-traceless graviton. *JCAP*, 05:071, 2021.

[95] Antonio De Felice, Kei-ichi Maeda, Shinji Mukohyama, and Masroor C. Pookkillath. Comparison of two theories of Type-IIa minimally modified gravity. *Phys. Rev. D*, 106(2):024028, 2022.

[96] Maria Mylova and Niayesh Afshordi. Effective cuscuton theory. *JHEP*, 04:144, 2024.

[97] Chunshan Lin, Jerome Quintin, and Robert H. Brandenberger. Massive gravity and the suppression of anisotropies and gravitational waves in a matter-dominated contracting universe. *JCAP*, 01:011, 2018.

[98] Supranta S. Boruah, Hyung J. Kim, Michael Rouben, and Ghazal Geshnizjani. Cuscuton bounce. *JCAP*, 08:031, 2018.

[99] Jerome Quintin and Daisuke Yoshida. Cuscuton gravity as a classically stable limiting curvature theory. *JCAP*, 02:016, 2020.

[100] Yuki Sakakihara, Daisuke Yoshida, Kazufumi Takahashi, and Jerome Quintin. Theories with limited extrinsic curvature and a nonsingular anisotropic universe. *Phys. Rev. D*, 102(8):084004, 2020.

[101] Chunshan Lin, Robert H. Brandenberger, and Laurence Perreault Levasseur. A Matter Bounce By Means of Ghost Condensation. *JCAP*, 04:019, 2011.

[102] Damien A. Easson, Ignacy Sawicki, and Alexander Vikman. G-Bounce. *JCAP*, 11:021, 2011.

[103] Taotao Qiu, Jarah Evslin, Yi-Fu Cai, Mingzhe Li, and Xinmin Zhang. Bouncing Galileon Cosmologies. *JCAP*, 10:036, 2011.

[104] Yi-Fu Cai, Damien A. Easson, and Robert Brandenberger. Towards a Nonsingular Bouncing Cosmology. *JCAP*, 08:020, 2012.

[105] Yi-Fu Cai, Evan McDonough, Francis Duplessis, and Robert H. Brandenberger. Two Field Matter Bounce Cosmology. *JCAP*, 10:024, 2013.

[106] Damien A. Easson, Ignacy Sawicki, and Alexander Vikman. When Matter Matters. *JCAP*, 07:014, 2013.

[107] Lorenzo Battarra, Michael Koehn, Jean-Luc Lehners, and Burt A. Ovrut. Cosmological Perturbations Through a Non-Singular Ghost-Condensate/Galileon Bounce. *JCAP*, 07:007, 2014.

[108] Anna Ijjas and Paul J. Steinhardt. Classically stable nonsingular cosmological bounces. *Phys. Rev. Lett.*, 117(12):121304, 2016.

[109] Anna Ijjas and Paul J. Steinhardt. Fully stable cosmological solutions with a non-singular classical bounce. *Phys. Lett. B*, 764:289–294, 2017.

[110] M. Libanov, S. Mironov, and V. Rubakov. Generalized Galileons: instabilities of bouncing and Genesis cosmologies and modified Genesis. *JCAP*, 08:037, 2016.

[111] Tsutomu Kobayashi. Generic instabilities of nonsingular cosmologies in Horndeski theory: A no-go theorem. *Phys. Rev. D*, 94(4):043511, 2016.

[112] Claudia de Rham and Scott Melville. Unitary null energy condition violation in P(X) cosmologies. *Phys. Rev. D*, 95(12):123523, 2017.

[113] David A. Dobre, Andrei V. Frolov, José T. Gálvez Ghersi, Sabir Ramazanov, and Alexander Vikman. Unbraiding the Bounce: Superluminality around the Corner. *JCAP*, 03:020, 2018.

[114] Anna Ijjas. Space-time slicing in Horndeski theories and its implications for non-singular bouncing solutions. *JCAP*, 02:007, 2018.

[115] Shingo Akama and Tsutomu Kobayashi. Generalized multi-Galileons, covariantized new terms, and the no-go theorem for nonsingular cosmologies. *Phys. Rev. D*, 95(6):064011, 2017.

[116] Shreya Banerjee, Yi-Fu Cai, and Emmanuel N. Saridakis. Evading the theoretical no-go theorem for nonsingular bounces in Horndeski/Galileon cosmology. *Class. Quant. Grav.*, 36(13):135009, 2019.

[117] Sergey Mironov. Mathematical Formulation of the No-Go Theorem in Horndeski Theory. *Universe*, 5(2):52, 2019.

[118] Gen Ye and Yun-Song Piao. Bounce in general relativity and higher-order derivative operators. *Phys. Rev. D*, 99(8):084019, 2019.

[119] Gen Ye and Yun-Song Piao. Implication of GW170817 for cosmological bounces. *Commun. Theor. Phys.*, 71(4):427, 2019.

[120] Paolo Creminelli, David Pirtskhalava, Luca Santoni, and Enrico Trincherini. Stability of Geodesically Complete Cosmologies. *JCAP*, 11:047, 2016.

[121] Yong Cai, Youping Wan, Hai-Guang Li, Taotao Qiu, and Yun-Song Piao. The Effective Field Theory of nonsingular cosmology. *JHEP*, 01:090, 2017.

[122] Yong Cai, Hai-Guang Li, Taotao Qiu, and Yun-Song Piao. The Effective Field Theory of nonsingular cosmology: II. *Eur. Phys. J. C*, 77(6):369, 2017.

[123] S. Mironov and V. Volkova. Non-singular cosmological scenarios in scalar-tensor theories and their stability: a review. , 9 2024.

[124] Lara Bohnenblust, Serena Giardino, Lavinia Heisenberg, and Nadine Nussbaumer. To bounce or not to bounce in generalized Proca theory and beyond. , 12 2024.

[125] Ok Song An, Jin U. Kang, Yong Jin Kim, Ui Ri Mun, and Un Gyong Ri. Fully viable DHOST bounce with extra scalar. , 1 2025.

[126] R. Kolevatov, S. Mironov, N. Sukhov, and V. Volkova. Cosmological bounce and Genesis beyond Horndeski. *JCAP*, 08:038, 2017.

[127] Yong Cai and Yun-Song Piao. A covariant Lagrangian for stable nonsingular bounce. *JHEP*, 09:027, 2017.

[128] S. Mironov, V. Rubakov, and V. Volkova. Bounce beyond Horndeski with GR asymptotics and γ-crossing. *JCAP*, 10:050, 2018.

[129] R. Kolevatov, S. Mironov, V. Rubakov, N. Sukhov, and V. Volkova. Cosmological bounce in Horndeski theory and beyond. *EPJ Web Conf.*, 191:07013, 2018.

[130] V. E. Volkova, S. A. Mironov, and V. A. Rubakov. Cosmological Scenarios with Bounce and Genesis in Horndeski Theory and Beyond. *J. Exp. Theor. Phys.*, 129(4):553–565, 2019.

[131] S. Mironov, V. Rubakov, and V. Volkova. Subluminal cosmological bounce beyond Horndeski. *JCAP*, 05:024, 2020.

[132] S. Mironov, V. Rubakov, and V. Volkova. Cosmological scenarios with bounce and Genesis in Horndeski theory and beyond: An essay in honor of I.M. Khalatnikov on the occasion of his 100th birthday. , 6 2019.

[133] Amara Ilyas, Mian Zhu, Yunlong Zheng, Yi-Fu Cai, and Emmanuel N. Saridakis. DHOST Bounce. *JCAP*, 09:002, 2020.

[134] Mian Zhu, Amara Ilyas, Yunlong Zheng, Yi-Fu Cai, and Emmanuel N. Saridakis. Scalar and tensor perturbations in DHOST bounce cosmology. *JCAP*, 11(11):045, 2021.

[135] S. Mironov and V. Volkova. Stable nonsingular cosmologies in beyond Horndeski theory and disformal transformations. *Int. J. Mod. Phys. A*, 37(14):2250088, 2022.

[136] Michael Koehn, Jean-Luc Lehners, and Burt Ovrut. Nonsingular bouncing cosmology: Consistency of the effective description. *Phys. Rev. D*, 93(10):103501, 2016.

[137] Daniel Baumann, Leonardo Senatore, and Matias Zaldarriaga. Scale-Invariance and the Strong Coupling Problem. *JCAP*, 05:004, 2011.

[138] Yuliya Aleksandrovna Ageeva and Pavel Konstantinovich Petrov. Unitarity relation and unitarity bounds for scalars with different sound speeds. *Phys. Usp.*, 66(11):1134–1141, 2023.

[139] Yong Cai, Ji Xu, Shuai Zhao, and Siyi Zhou. Perturbative unitarity and NEC violation in genesis cosmology. *JHEP*, 10:140, 2022. [Erratum: JHEP 11, 063 (2022)].

[140] Yulia Ageeva, Pavel Petrov, and Valery Rubakov. Generating cosmological perturbations in non-singular Horndeski cosmologies. *JHEP*, 01:026, 2023.

[141] Austin Joyce and Justin Khoury. Strong Coupling Problem with Time-Varying Sound Speed. *Phys. Rev. D*, 84:083514, 2011.

[142] Y. A. Ageeva, O. A. Evseev, O. I. Melichev, and V. A. Rubakov. Horndeski Genesis: strong coupling and absence thereof. *EPJ Web Conf.*, 191:07010, 2018.

[143] Y. Ageeva, O. Evseev, O. Melichev, and V. Rubakov. Toward evading the strong coupling problem in Horndeski genesis. *Phys. Rev. D*, 102(2):023519, 2020.

[144] Yulia Ageeva, Pavel Petrov, and Valery Rubakov. Horndeski genesis: consistency of classical theory. *JHEP*, 12:107, 2020.

[145] Y. Ageeva, P. Petrov, and V. Rubakov. Nonsingular cosmological models with strong gravity in the past. *Phys. Rev. D*, 104(6):063530, 2021.

[146] Yulia Ageeva and Pavel Petrov. On the strong coupling problem in cosmologies with "strong gravity in the past". *Mod. Phys. Lett. A*, 37(26):2250171, 2022.

[147] Shingo Akama and Shin'ichi Hirano. Primordial non-Gaussianity from Galilean Genesis without strong coupling problem. *Phys. Rev. D*, 107(6):063504, 2023.

[148] S. Dubovsky, T. Gregoire, A. Nicolis, and R. Rattazzi. Null energy condition and superluminal propagation. *JHEP*, 03:025, 2006.

[149] S. Mironov, V. Rubakov, and V. Volkova. Superluminality in beyond Horndeski theory with extra scalar field. *Phys. Scripta*, 95(8):084002, 2020.

[150] S. Mironov, V. Rubakov, and V. Volkova. Superluminality in DHOST theory with extra scalar. *JHEP*, 04:035, 2021.

[151] David Wands. Duality invariance of cosmological perturbation spectra. *Phys. Rev. D*, 60:023507, 1999.

[152] Fabio Finelli and Robert Brandenberger. On the generation of a scale invariant spectrum of adiabatic fluctuations in cosmological models with a contracting phase. *Phys. Rev. D*, 65:103522, 2002.

[153] Robert H. Brandenberger. The Matter Bounce Alternative to Inflationary Cosmology. , 6 2012.

[154] Justin Khoury, Burt A. Ovrut, Paul J. Steinhardt, and Neil Turok. The Ekpyrotic universe: Colliding branes and the origin of the hot big bang. *Phys. Rev. D*, 64:123522, 2001.

[155] Evgeny I. Buchbinder, Justin Khoury, and Burt A. Ovrut. New Ekpyrotic cosmology. *Phys. Rev. D*, 76:123503, 2007.

[156] Jean-Luc Lehners. Ekpyrotic and Cyclic Cosmology. *Phys. Rept.*, 465:223–263, 2008.

[157] M. Gasperini and G. Veneziano. Pre-big bang in string cosmology. *Astropart. Phys.*, 1:317–339, 1993.

[158] M. Gasperini and G. Veneziano. The pre-big bang scenario in string cosmology. *Phys. Rept.*, 373:1–212, 2003.

[159] Maurizio Gasperini. *Elements of string cosmology.* Cambridge University Press, Cambridge, UK, 2007.

[160] Jerome Quintin, Zeinab Sherkatghanad, Yi-Fu Cai, and Robert H. Brandenberger. Evolution of cosmological perturbations and the production of non-Gaussianities through a nonsingular bounce: Indications for a no-go theorem in single field matter bounce cosmologies. *Phys. Rev. D*, 92(6):063532, 2015.

[161] Yu-Bin Li, Jerome Quintin, Dong-Gang Wang, and Yi-Fu Cai. Matter bounce cosmology with a generalized single field: non-Gaussianity and an extended no-go theorem. *JCAP*, 03:031, 2017.

[162] Yi-Fu Cai, Robert Brandenberger, and Patrick Peter. Anisotropy in a Nonsingular Bounce. *Class. Quant. Grav.*, 30:075019, 2013.

[163] Aaron M. Levy. Fine-tuning challenges for the matter bounce scenario. *Phys. Rev. D*, 95(2):023522, 2017.

[164] Chandrima Ganguly and Jerome Quintin. Microphysical manifestations of viscosity and consequences for anisotropies in the very early universe. *Phys. Rev. D*, 105(2):023532, 2022.

[165] Shingo Akama, Shin'ichi Hirano, and Tsutomu Kobayashi. Primordial non-Gaussianities of scalar and tensor perturbations in general bounce cosmology: Evading the no-go theorem. *Phys. Rev. D*, 101(4):043529, 2020.

[166] Shingo Akama and Mian Zhu. Parity violation in primordial tensor non-Gaussianities from matter bounce cosmology. *JCAP*, 07:039, 2024.

[167] Shingo Akama, Giorgio Orlando, and Paola C. M. Delgado. Towards testing the general bounce cosmology with the CMB B-mode auto-bispectrum. *JCAP*, 09:055, 2024.

[168] Shingo Akama. Primordial full bispectra from the general bounce cosmology. , 2 2025.

[169] Angelika Fertig, Jean-Luc Lehners, and Enno Mallwitz. Ekpyrotic Perturbations With Small Non-Gaussian Corrections. *Phys. Rev. D*, 89(10):103537, 2014.

[170] Anna Ijjas, Jean-Luc Lehners, and Paul J. Steinhardt. General mechanism for producing scale-invariant perturbations and small non-Gaussianity in ekpyrotic models. *Phys. Rev. D*, 89(12):123520, 2014.

[171] Jerome Quintin, Xingang Chen, and Reza Ebadi. Fingerprints of a non-inflationary universe from massive fields. *JCAP*, 09:026, 2024.

[172] J. Leo Kim and Ghazal Geshnizjani. Spectrum of Cuscuton Bounce. *JCAP*, 03:104, 2021.

[173] Joel K. Erickson, Daniel H. Wesley, Paul J. Steinhardt, and Neil Turok. Kasner and mixmaster behavior in universes with equation of state $w \geq 1$. *Phys. Rev. D*, 69:063514, 2004.

[174] A. A. Coley and W. C. Lim. Asymptotic analysis of spatially inhomogeneous stiff and ultra-stiff cosmologies. *Class. Quant. Grav.*, 22:3073–3082, 2005.

[175] James E. Lidsey. Cosmic no hair for collapsing universes. *Class. Quant. Grav.*, 23:3517–3530, 2006.

[176] David Garfinkle, Woei Chet Lim, Frans Pretorius, and Paul J. Steinhardt. Evolution to a smooth universe in an ekpyrotic contracting phase with $w > 1$. *Phys. Rev. D*, 78:083537, 2008.

[177] John D Barrow and Kei Yamamoto. Anisotropic Pressures at Ultra-stiff Singularities and the Stability of Cyclic Universes. *Phys. Rev. D*, 82:063516, 2010.

[178] J. Mark Heinzle and Patrik Sandin. The initial singularity of ultrastiff perfect fluid spacetimes without symmetries. *Commun. Math. Phys.*, 313:385–403, 2012.

[179] John D. Barrow and Chandrima Ganguly. Evolution of initially contracting Bianchi Class A models in the presence of an ultra-stiff anisotropic pressure fluid. *Class. Quant. Grav.*, 33(12):125004, 2016.

[180] William G. Cook, Iryna A. Glushchenko, Anna Ijjas, Frans Pretorius, and Paul J. Steinhardt. Supersmoothing through Slow Contraction. *Phys. Lett. B*, 808:135690, 2020.

[181] Angelika Fertig and Jean-Luc Lehners. The Non-Minimal Ekpyrotic Trispectrum. *JCAP*, 01:026, 2016.

[182] Angelika Fertig, Jean-Luc Lehners, Enno Mallwitz, and Edward Wilson-Ewing. Converting entropy to curvature perturbations after a cosmic bounce. *JCAP*, 10:005, 2016.

[183] Viatcheslav F. Mukhanov and Robert H. Brandenberger. A Nonsingular universe. *Phys. Rev. Lett.*, 68:1969–1972, 1992.

[184] Robert H. Brandenberger, Viatcheslav F. Mukhanov, and A. Sornborger. A Cosmological theory without singularities. *Phys. Rev. D*, 48:1629–1642, 1993.

[185] Daisuke Yoshida, Jerome Quintin, Masahide Yamaguchi, and Robert H. Brandenberger. Cosmological perturbations and stability of nonsingular cosmologies with limiting curvature. *Phys. Rev. D*, 96(4):043502, 2017.

[186] Aaron M. Levy, Anna Ijjas, and Paul J. Steinhardt. Scale-invariant perturbations in ekpyrotic cosmologies without fine-tuning of initial conditions. *Phys. Rev. D*, 92(6):063524, 2015.

[187] Anna Ijjas, William G. Cook, Frans Pretorius, Paul J. Steinhardt, and Elliot Y. Davies. Robustness of slow contraction to cosmic initial conditions. *JCAP*, 08:030, 2020.

[188] Anna Ijjas, Andrew P. Sullivan, Frans Pretorius, Paul J. Steinhardt, and William G. Cook. Ultralocality and slow contraction. *JCAP*, 06:013, 2021.

[189] Anna Ijjas, Frans Pretorius, Paul J. Steinhardt, and Andrew P. Sullivan. The effects of multiple modes and reduced symmetry on the rapidity and robustness of slow contraction. *Phys. Lett. B*, 820:136490, 2021.

[190] Anna Ijjas, Frans Pretorius, Paul J. Steinhardt, and David Garfinkle. Dynamical attractors in contracting spacetimes dominated by kinetically coupled scalar fields. *JCAP*, 12(12):030, 2021.

[191] Anna Ijjas. Numerical Relativity as a New Tool for Fundamental Cosmology. *MDPI Physics*, 4(1):301–314, 2022.

[192] Timo Kist and Anna Ijjas. The robustness of slow contraction and the shape of the scalar field potential. *JCAP*, 08(08):046, 2022.

[193] Anna Ijjas. Slow Contraction and the Weyl Curvature Hypothesis. , 4 2023.

[194] Anna Ijjas, Paul J. Steinhardt, David Garfinkle, and William G. Cook. Smoothing and flattening the universe through slow contraction versus inflation. *JCAP*, 07:077, 2024.

[195] Jean-Luc Lehners. Cosmic Bounces and Cyclic Universes. *Class. Quant. Grav.*, 28:204004, 2011.

[196] Andrew R. Liddle and D. H. Lyth. *Cosmological inflation and large scale structure*. Cambridge University Press, Cambridge, UK, 2000.

[197] Jean-Luc Lehners. Small-Field and Scale-Free: Inflation and Ekpyrosis at their Extremes. *JCAP*, 11:001, 2018.

[198] Heliudson Bernardo and Robert Brandenberger. Contracting cosmologies and the swampland. *JHEP*, 07:206, 2021.

[199] Gary Shiu, Flavio Tonioni, and Hung V. Tran. Collapsing universe before time. *JCAP*, 05:124, 2024.

[200] Yu-ichi Takamizu and Kei-ichi Maeda. Collision of domain walls and reheating of the brane universe. *Phys. Rev. D*, 70:123514, 2004.

[201] Jean-Luc Lehners and Paul J. Steinhardt. Non-Gaussian density fluctuations from entropically generated curvature perturbations in Ekpyrotic models. *Phys. Rev. D*, 77:063533, 2008. [Erratum: Phys.Rev.D 79, 129903 (2009)].

[202] Thorsten Battefeld. Modulated Perturbations from Instant Preheating after new Ekpyrosis. *Phys. Rev. D*, 77:063503, 2008.

[203] Jean-Luc Lehners and Paul J. Steinhardt. Intuitive understanding of non-gaussianity in ekpyrotic and cyclic models. *Phys. Rev. D*, 78:023506, 2008. [Erratum: Phys.Rev.D 79, 129902 (2009)].

[204] Jean-Luc Lehners and Paul J. Steinhardt. Non-Gaussianity Generated by the Entropic Mechanism in Bouncing Cosmologies Made Simple. *Phys. Rev. D*, 80:103520, 2009.

[205] Jean-Luc Lehners and Sebastien Renaux-Petel. Multifield Cosmological Perturbations at Third Order and the Ekpyrotic Trispectrum. *Phys. Rev. D*, 80:063503, 2009.

[206] Jean-Luc Lehners. Ekpyrotic Non-Gaussianity: A Review. *Adv. Astron.*, 2010:903907, 2010.

[207] Jerome Quintin, Yi-Fu Cai, and Robert H. Brandenberger. Matter creation in a non-singular bouncing cosmology. *Phys. Rev. D*, 90(6):063507, 2014.

[208] W. S. Hipolito-Ricaldi, Robert Brandenberger, Elisa G. M. Ferreira, and L. L. Graef. Particle Production in Ekpyrotic Scenarios. *JCAP*, 11:024, 2016.

[209] Anna Ijjas and Roman Kolevatov. Sourcing curvature modes with entropy perturbations in non-singular bouncing cosmologies. *JCAP*, 06:012, 2021.

[210] Anna Ijjas and Roman Kolevatov. Nearly scale-invariant curvature modes from entropy perturbations during the graceful exit phase. *Phys. Rev. D*, 103(10):L101302, 2021.

[211] David Langlois, Sebastien Renaux-Petel, Daniele A. Steer, and Takahiro Tanaka. Primordial perturbations and non-Gaussianities in DBI and general multi-field inflation. *Phys. Rev. D*, 78:063523, 2008.

[212] Stefano Camera, Mario G. Santos, and Roy Maartens. Probing primordial non-Gaussianity with SKA galaxy redshift surveys: a fully relativistic analysis. *Mon. Not. Roy. Astron. Soc.*, 448(2):1035–1043, 2015. [Erratum: Mon.Not.Roy.Astron.Soc. 467, 1505–1506 (2017)].

[213] Dionysios Karagiannis, Andrei Lazanu, Michele Liguori, Alvise Raccanelli, Nicola Bartolo, and Licia Verde. Constraining primordial non-Gaussianity with bispectrum and power spectrum from upcoming optical and radio surveys. *Mon. Not. Roy. Astron. Soc.*, 478(1):1341–1376, 2018.

[214] Peter Ade et al. The Simons Observatory: Science goals and forecasts. *JCAP*, 02:056, 2019.

[215] P. Daniel Meerburg et al. Primordial Non-Gaussianity. *Bull. Am. Astron. Soc.*, 51(3):107, 2019.

[216] Dionysios Karagiannis, Anže Slosar, and Michele Liguori. Forecasts on Primordial non-Gaussianity from 21 cm Intensity Mapping experiments. *JCAP*, 11:052, 2020.

[217] Matteo Biagetti, Juan Calles, Lina Castiblanco, Alex Cole, and Jorge Noreña. Fisher forecasts for primordial non-Gaussianity from persistent homology. *JCAP*, 10:002, 2022.

[218] Xian Gao, Marc Lilley, and Patrick Peter. Production of non-gaussianities through a positive spatial curvature bouncing phase. *JCAP*, 07:010, 2014.

[219] Xian Gao, Marc Lilley, and Patrick Peter. Non-Gaussianity excess problem in classical bouncing cosmologies. *Phys. Rev. D*, 91(2):023516, 2015.

[220] Hael Collins. Primordial non-Gaussianities from inflation. , 1 2011.

[221] Louis Leblond and Sarah Shandera. Simple Bounds from the Perturbative Regime of Inflation. *JCAP*, 08:007, 2008.

[222] Carlos Duaso Pueyo, Harry Goodhew, Ciaran McCulloch, and Enrico Pajer. Perturbative unitarity bounds from momentum-space entanglement. , 10 2024.

[223] Ghazal Geshnizjani and Nahid Ahmadi. Can non-local or higher derivative theories provide alternatives to inflation? *JCAP*, 11:029, 2013.

[224] N. Bartolo, E. Komatsu, Sabino Matarrese, and A. Riotto. Non-Gaussianity from inflation: Theory and observations. *Phys. Rept.*, 402:103–266, 2004.

[225] Eiichiro Komatsu and David N. Spergel. Acoustic signatures in the primary microwave background bispectrum. *Phys. Rev. D*, 63:063002, 2001.

[226] Yi-Fu Cai, Wei Xue, Robert Brandenberger, and Xinmin Zhang. Non-Gaussianity in a Matter Bounce. *JCAP*, 05:011, 2009.

[227] Nicola Bartolo, Alexander Ganz, and Sabino Matarrese. Cuscuton inflation. *JCAP*, 05(05):008, 2022.

[228] Alexander Ganz, Paul Martens, Shinji Mukohyama, and Ryo Namba. Bispectrum from inflation/bouncing Universe in VCDM. , 7 2024.

[229] Paola C. M. Delgado, Ruth Durrer, and Nelson Pinto-Neto. The CMB bispectrum from bouncing cosmologies. *JCAP*, 11:024, 2021.

[230] Bartjan van Tent, Paola C. M. Delgado, and Ruth Durrer. Constraining the Bispectrum from Bouncing Cosmologies with Planck. *Phys. Rev. Lett.*, 130(19):191002, 2023.

[231] Y. Ageeva, M. Kotenko, and P. Petrov. The primordial non-Gaussianities for non-singular Horndeski cosmologies. , 10 2024.

[232] Alexander Ganz, Paul Martens, Shinji Mukohyama, and Ryo Namba. Bouncing cosmology in VCDM. *JCAP*, 04:060, 2023.

[233] Justin Khoury and Paul J. Steinhardt. Adiabatic Ekpyrosis: Scale-Invariant Curvature Perturbations from a Single Scalar Field in a Contracting Universe. *Phys. Rev. Lett.*, 104:091301, 2010.

[234] Justin Khoury and Paul J. Steinhardt. Generating Scale-Invariant Perturbations from Rapidly-Evolving Equation of State. *Phys. Rev. D*, 83:123502, 2011.

[235] Ghazal Geshnizjani and William H. Kinney. Theoretical implications of detecting gravitational waves. *JCAP*, 08:008, 2015.

[236] Oliver H. E. Philcox. Searching for Inflationary Physics with the CMB Trispectrum: 1. Primordial Theory & Optimal Estimators. , 2 2025.

[237] Oliver H. E. Philcox. Searching for Inflationary Physics with the CMB Trispectrum: 2. Code & Validation. , 2 2025.

[238] Oliver H. E. Philcox. Searching for Inflationary Physics with the CMB Trispectrum: 3. Constraints from Planck. , 2 2025.

[239] Xingang Chen. Primordial Features as Evidence for Inflation. *JCAP*, 01:038, 2012.

[240] Xingang Chen. Fingerprints of Primordial Universe Paradigms as Features in Density Perturbations. *Phys. Lett. B*, 706:111–115, 2011.

[241] Xingang Chen, Mohammad Hossein Namjoo, and Yi Wang. Quantum Primordial Standard Clocks. *JCAP*, 02:013, 2016.

[242] Xingang Chen, Abraham Loeb, and Zhong-Zhi Xianyu. Unique Fingerprints of Alternatives to Inflation in the Primordial Power Spectrum. *Phys. Rev. Lett.*, 122(12):121301, 2019.

[243] Guillem Domènech, Xingang Chen, Marc Kamionkowski, and Abraham Loeb. Planck residuals anomaly as a fingerprint of alternative scenarios to inflation. *JCAP*, 10:005, 2020.

[244] B. J. Carr and A. A. Coley. Persistence of black holes through a cosmological bounce. *Int. J. Mod. Phys. D*, 20:2733–2738, 2011.

[245] Jerome Quintin and Robert H. Brandenberger. Black hole formation in a contracting universe. *JCAP*, 11:029, 2016.

[246] Jie-Wen Chen, Junyu Liu, Hao-Lan Xu, and Yi-Fu Cai. Tracing Primordial Black Holes in Nonsingular Bouncing Cosmology. *Phys. Lett. B*, 769:561–568, 2017.

[247] Timothy Clifton, Bernard Carr, and Alan Coley. Persistent Black Holes in Bouncing Cosmologies. *Class. Quant. Grav.*, 34(13):135005, 2017.

[248] Maxence Corman, William E. East, and Justin L. Ripley. Evolution of black holes through a nonsingular cosmological bounce. *JCAP*, 09:063, 2022.

[249] Jie-Wen Chen, Mian Zhu, Sheng-Feng Yan, Qing-Qing Wang, and Yi-Fu Cai. Enhance primordial black hole abundance through the non-linear processes around bounce point. *JCAP*, 01:015, 2023.

[250] Yi-Fu Cai, Chengfeng Tang, Geyu Mo, Sheng-Feng Yan, Chao Chen, Xiao-Han Ma, Bo Wang, Wentao Luo, Damien A. Easson, and Antonino Marciano. Primordial black hole mass functions as a probe of cosmic origin. *Sci. China Phys. Mech. Astron.*, 67(5):259512, 2024.

[251] Javier Chagoya and Gianmassimo Tasinato. A geometrical approach to degenerate scalar-tensor theories. *JHEP*, 02:113, 2017.

[252] Claudia de Rham and Hayato Motohashi. Caustics for Spherical Waves. *Phys. Rev. D*, 95(6):064008, 2017.

[253] Antonio De Felice, David Langlois, Shinji Mukohyama, Karim Noui, and Anzhong Wang. Generalized instantaneous modes in higher-order scalar-tensor theories. *Phys. Rev. D*, 98(8):084024, 2018.

[254] Enrico Pajer and David Stefanyszyn. Symmetric Superfluids. *JHEP*, 06:008, 2019.

[255] Tanguy Grall, Sadra Jazayeri, and Enrico Pajer. Symmetric Scalars. *JCAP*, 05:031, 2020.

[256] Valerio Faraoni, Andrea Giusti, Sonia Jose, and Serena Giardino. Peculiar thermal states in the first-order thermodynamics of gravity. *Phys. Rev. D*, 106(2):024049, 2022.

[257] Marcello Miranda, Daniele Vernieri, Salvatore Capozziello, and Valerio Faraoni. Fluid nature constrains Horndeski gravity. *Gen. Rel. Grav.*, 55(7):84, 2023.

[258] Niayesh Afshordi. Cuscuton and low energy limit of Horava-Lifshitz gravity. *Phys. Rev. D*, 80:081502, 2009.

[259] Petr Horava. Quantum Gravity at a Lifshitz Point. *Phys. Rev. D*, 79:084008, 2009.

[260] Shinji Mukohyama and Ryo Namba. Partial UV Completion of $P(X)$ from a Curved Field Space. *JCAP*, 02:001, 2021.

[261] Guillermo Lara, Miguel Bezares, and Enrico Barausse. UV completions, fixing the equations, and nonlinearities in k-essence. *Phys. Rev. D*, 105(6):064058, 2022.

[262] Houjun Mo, Frank C. van den Bosch, and Simon White. *Galaxy Formation and Evolution*. Cambridge University Press, 2010.

[263] A COORAY and R SHETH. Halo models of large scale structure. *Physics Reports*, 372(1):1–129, December 2002.

[264] U. Seljak. Analytic model for galaxy and dark matter clustering. *Monthly Notices of the Royal Astronomical Society*, 318(1):203–213, October 2000.

[265] J. A. Peacock and R. E. Smith. Halo occupation numbers and galaxy bias. *Monthly Notices of the Royal Astronomical Society*, 318(4):1144–1156, November 2000.

[266] Vincent Desjacques, Donghui Jeong, and Fabian Schmidt. Large-scale galaxy bias. *Physics Reports*, 733:1–193, February 2018.

[267] Will J. Percival. Large scale structure observations, 2013.

[268] Daniel J. Eisenstein and Wayne Hu. Power spectra for cold dark matter and its variants. *The Astrophysical Journal*, 511(1):5–15, January 1999.

[269] Daniel J. Eisenstein and Wayne Hu. Baryonic features in the matter transfer function. *The Astrophysical Journal*, 496(2):605–614, April 1998.

[270] William H. Press and Paul Schechter. Formation of Galaxies and Clusters of Galaxies by Self-Similar Gravitational Condensation. *apj*, 187:425–438, February 1974.

[271] J. R. Bond, S. Cole, G. Efstathiou, and N. Kaiser. Excursion Set Mass Functions for Hierarchical Gaussian Fluctuations. *apj*, 379:440, October 1991.

[272] Michele Maggiore. *Gravitational Waves. Vol. 1: Theory and Experiments*. Oxford University Press, 2007.

[273] Michele Maggiore. *Gravitational Waves. Vol. 2: Astrophysics and Cosmology*. Oxford University Press, 3 2018.

[274] E. Poisson and C.M. Will. *Gravity: Newtonian, Post-Newtonian, Relativistic*. Cambridge University Press, 2014.

[275] Eanna E. Flanagan and Scott A. Hughes. The Basics of gravitational wave theory. *New J. Phys.*, 7:204, 2005.

[276] Charles W. Misner, K. S. Thorne, and J. A. Wheeler. *Gravitation*. W. H. Freeman, San Francisco, 1973.

[277] Neha Singh, Brynmor Haskell, Dipanjan Mukherjee, and Tomasz Bulik. Asymmetric accretion and thermal 'mountains' in magnetized neutron star crusts. *Mon. Not. Roy. Astron. Soc.*, 493(3):3866–3878, 2020.

[278] Christopher P. L. Berry, Scott A. Hughes, Carlos F. Sopuerta, Alvin J. K. Chua, Anna Heffernan, Kelly Holley-Bockelmann, Deyan P. Mihaylov, M. Coleman Miller, and Alberto Sesana. The unique potential of extreme mass-ratio inspirals for gravitational-wave astronomy, 2019.

[279] Nelson Christensen. Stochastic Gravitational Wave Backgrounds. *Rept. Prog. Phys.*, 82(1):016903, 2019.

[280] Gabriella Agazie et al. The NANOGrav 15 yr Data Set: Evidence for a Gravitational-wave Background. *Astrophys. J. Lett.*, 951(1):L8, 2023.

[281] B. P. Abbott et al. Observation of Gravitational Waves from a Binary Black Hole Merger. *Phys. Rev. Lett.*, 116(6):061102, 2016.

[282] Jahed Abedi, Hannah Dykaar, and Niayesh Afshordi. Echoes from the abyss: Tentative evidence for planck-scale structure at black hole horizons. *Physical Review D*, 96(8), October 2017.

[283] Krzysztof Belczynski, Vassiliki Kalogera, and Tomasz Bulik. A Comprehensive study of binary compact objects as gravitational wave sources: Evolutionary channels, rates, and physical properties. *Astrophys. J.*, 572:407–431, 2001.

[284] Coenraad J. Neijssel, Alejandro Vigna-Gómez, Simon Stevenson, Jim W. Barrett, Sebastian M. Gaebel, Floor S. Broekgaarden, Selma E. de Mink, Dorottya Szécsi, Serena Vinciguerra, and Ilya Mandel. The effect of the metallicity-specific star formation history on double compact object mergers. , 490(3):3740–3759, December 2019.

[285] Michal Dominik, Emanuele Berti, Richard O'Shaughnessy, Ilya Mandel, Krzysztof Belczynski, Christopher Fryer, Daniel E. Holz, Tomasz Bulik, and Francesco Pannarale. Double Compact Objects III: Gravitational Wave Detection Rates. *Astrophys. J.*, 806(2):263, 2015.

[286] Michela Mapelli, Nicola Giacobbo, Emanuele Ripamonti, and Mario Spera. The cosmic merger rate of stellar black hole binaries from the Illustris simulation. *Mon. Not. Roy. Astron. Soc.*, 472(2):2422–2435, 2017.

[287] Filippo Santoliquido, Michela Mapelli, M. Celeste Artale, and Lumen Boco. Modelling the host galaxies of binary compact object mergers with observational scaling relations. *Mon. Not. Roy. Astron. Soc.*, 516(3):3297–3317, 2022.

[288] B. P. Abbott et al. Prospects for observing and localizing gravitational-wave transients with Advanced LIGO, Advanced Virgo and KAGRA. *Living Rev. Rel.*, 19:1, 2016.

[289] J. M. Weisberg, J. H. Taylor, and L. A. Fowler. Gravitational waves from an orbiting pulsar. *Scientific American*, 245:74–82, October 1981.

[290] J. M. Weisberg, D. J. Nice, and J. H. Taylor. Timing measurements of the relativistic binary pulsar psr b1913+16. *The Astrophysical Journal*, 722(2):1030–1034, September 2010.

[291] B. Willke et al. The GEO 600 gravitational wave detector. *Class. Quant. Grav.*, 19:1377–1387, 2002.

[292] R. Abbott et al. Population of Merging Compact Binaries Inferred Using Gravitational Waves through GWTC-3. *Phys. Rev. X*, 13(1):011048, 2023.

[293] Matthew Evans et al. A horizon study for cosmic explorer: Science, observatories, and community, 2021.

[294] M. Punturo et al. The Einstein Telescope: A third-generation gravitational wave observatory. *Class. Quant. Grav.*, 27:194002, 2010.

[295] Marica Branchesi et al. Science with the Einstein Telescope: a comparison of different designs. *JCAP*, 07:068, 2023.

[296] Pau Amaro Seoane et al. Astrophysics with the Laser Interferometer Space Antenna. *Living Rev. Rel.*, 26(1):2, 2023.

[297] Pierre Auclair et al. Cosmology with the Laser Interferometer Space Antenna. *Living Rev. Rel.*, 26(1):5, 2023.

[298] J. Aasi et al. Advanced LIGO. *Class. Quant. Grav.*, 32:074001, 2015.

[299] F. Acernese et al. Advanced Virgo: a second-generation interferometric gravitational wave detector. *Class. Quant. Grav.*, 32(2):024001, 2015.

[300] Yoichi Aso, Yuta Michimura, Kentaro Somiya, Masaki Ando, Osamu Miyakawa, Takanori Sekiguchi, Daisuke Tatsumi, and Hiroaki Yamamoto. Interferometer design of the kagra gravitational wave detector. *Phys. Rev. D*, 88:043007, Aug 2013.

[301] B. P. Abbott et al. Binary Black Hole Population Properties Inferred from the First and Second Observing Runs of Advanced LIGO and Advanced Virgo. *Astrophys. J. Lett.*, 882(2):L24, 2019.

[302] R. Abbott et al. Population Properties of Compact Objects from the Second LIGO-Virgo Gravitational-Wave Transient Catalog. *Astrophys. J. Lett.*, 913(1):L7, 2021.

[303] R. Abbott et al. Constraints on the Cosmic Expansion History from GWTC–3. *Astrophys. J.*, 949(2):76, 2023.

[304] Péter Mészáros, Derek B. Fox, Chad Hanna, and Kohta Murase. Multi-Messenger Astrophysics. *Nature Rev. Phys.*, 1:585–599, 2019.

[305] Emanuele Berti, Vitor Cardoso, Zoltán Haiman, Daniel E. Holz, Emil Mottola, Suvodip Mukherjee, Bangalore Sathyaprakash, Xavier Siemens, and Nicolás Yunes. Snowmass2021 Cosmic Frontier White Paper: Fundamental Physics and Beyond the Standard Model. In *Snowmass 2021*, 3 2022.

[306] Rana X. Adhikari et al. Report of the Topical Group on Cosmic Probes of Fundamental Physics for for Snowmass 2021. , 9 2022.

[307] B. P. Abbott et al. GWTC-1: A Gravitational-Wave Transient Catalog of Compact Binary Mergers Observed by LIGO and Virgo during the First and Second Observing Runs. *Phys. Rev. X*, 9(3):031040, 2019.

[308] R. Abbott et al. GWTC-2: Compact Binary Coalescences Observed by LIGO and Virgo During the First Half of the Third Observing Run. *Phys. Rev. X*, 11:021053, 2021.

[309] R. Abbott et al. GWTC-3: Compact Binary Coalescences Observed by LIGO and Virgo During the Second Part of the Third Observing Run. , 11 2021.

[310] Bernard F. Schutz. Determining the Hubble Constant from Gravitational Wave Observations. *Nature*, 323:310–311, 1986.

[311] Masamune Oguri. Measuring the distance-redshift relation with the cross-correlation of gravitational wave standard sirens and galaxies. *Phys. Rev. D*, 93(8):083511, 2016.

[312] B. P. Abbott et al. A gravitational-wave standard siren measurement of the Hubble constant. *Nature*, 551(7678):85–88, 2017.

[313] Hsin-Yu Chen, Maya Fishbach, and Daniel E. Holz. A two per cent Hubble constant measurement from standard sirens within five years. *Nature*, 562(7728):545–547, 2018.

[314] M. Fishbach et al. A Standard Siren Measurement of the Hubble Constant from GW170817 without the Electromagnetic Counterpart. *Astrophys. J. Lett.*, 871(1):L13, 2019.

[315] Suvodip Mukherjee and Benjamin D. Wandelt. Beyond the classical distance-redshift test: cross-correlating redshift-free standard candles and sirens with redshift surveys. , 8 2018.

[316] M. Soares-Santos et al. First Measurement of the Hubble Constant from a Dark Standard Siren using the Dark Energy Survey Galaxies and the LIGO/Virgo Binary–Black-hole Merger GW170814. *Astrophys. J. Lett.*, 876(1):L7, 2019.

[317] Suvodip Mukherjee, Benjamin D. Wandelt, Samaya M. Nissanke, and Alessandra Silvestri. Accurate precision Cosmology with redshift unknown gravitational wave sources. *Phys. Rev. D*, 103(4):043520, 2021.

[318] Sayantani Bera, Divya Rana, Surhud More, and Sukanta Bose. Incompleteness Matters Not: Inference of H_0 from Binary Black Hole–Galaxy Cross-correlations. *Astrophys. J.*, 902(1):79, 2020.

[319] Suvodip Mukherjee, Alex Krolewski, Benjamin D. Wandelt, and Joseph Silk. Cross-correlating dark sirens and galaxies: measurement of H_0 from GWTC-3 of LIGO-Virgo-KAGRA. , 3 2022.

[320] Samsuzzaman Afroz and Suvodip Mukherjee. Prospect of Precision Cosmology and Testing General Relativity using Binary Black Holes- Galaxies Cross-correlation. , 7 2024.

[321] J. Antoniadis et al. The second data release from the European Pulsar Timing Array I. The dataset and timing analysis. *Astron. Astrophys.*, 6 2023.

[322] Heng Xu et al. Searching for the Nano-Hertz Stochastic Gravitational Wave Background with the Chinese Pulsar Timing Array Data Release I. *Res. Astron. Astrophys.*, 23(7):075024, 2023.

[323] Mohit Raj Sah and Suvodip Mukherjee. Discovering the Cosmic Evolution of Supermassive Black Holes using Nano-Hertz Gravitational Waves and Galaxy Surveys. , 7 2024.

[324] Mohit Raj Sah, Suvodip Mukherjee, Vida Saeedzadeh, Arif Babul, Michael Tremmel, and Thomas R. Quinn. Imprints of Supermassive Black Hole Evolution on the Spectral and Spatial Anisotropy of Nano-Hertz Stochastic Gravitational-Wave Background. , 4 2024.

[325] R. O'Shaughnessy, V. Kalogera, and K. Belczynski. Binary Compact Object Coalescence Rates: The Role of Elliptical Galaxies. *Astrophys. J.*, 716:615–633, 2010.

[326] Nicola Giacobbo, Michela Mapelli, and Mario Spera. Merging black hole binaries: the effects of progenitor's metallicity, mass-loss rate and Eddington factor. *Mon. Not. Roy. Astron. Soc.*, 474(3):2959–2974, 2018.

[327] Liang Cao, Youjun Lu, and Yuetong Zhao. Host galaxy properties of mergers of stellar binary black holes and their implications for advanced LIGO gravitational wave sources. , 474(4):4997–5007, March 2018.

[328] Maya Fishbach, Daniel E. Holz, and Will M. Farr. Does the Black Hole Merger Rate Evolve with Redshift? *Astrophys. J. Lett.*, 863(2):L41, 2018.

[329] Michal Dominik, Krzysztof Belczynski, Christopher Fryer, Daniel Holz, Emanuele Berti, Tomasz Bulik, Ilya Mandel, and Richard O'Shaughnessy. Double Compact Objects I: The Significance of the Common Envelope on Merger Rates. *Astrophys. J.*, 759:52, 2012.

[330] Mattia Toffano, Michela Mapelli, Nicola Giacobbo, Maria Celeste Artale, and Giancarlo Ghirlanda. The host galaxies of double compact objects across cosmic time. *Mon. Not. Roy. Astron. Soc.*, 489(4):4622–4631, 2019.

[331] M. Celeste Artale, Michela Mapelli, Yann Bouffanais, Nicola Giacobbo, Mario Spera, and Mario Pasquato. Mass and star formation rate of the host galaxies of compact binary mergers across cosmic time. *Mon. Not. Roy. Astron. Soc.*, 491(3):3419–3434, 2020.

[332] Kevin S. McCarthy, Zheng Zheng, and Enrico Ramirez-Ruiz. Constraining delay time distribution of binary neutron star mergers from host galaxy properties. , 499(4):5220–5229, December 2020.

[333] Samsuzzaman Afroz and Suvodip Mukherjee. Phase Space of Binary Black Holes from Gravitational Wave Observations to Unveil its Formation History. , 11 2024.

[334] Davide Gerosa and Emanuele Berti. Are merging black holes born from stellar collapse or previous mergers? *prd*, 95(12):124046, June 2017.

[335] Y. Yang, I. Bartos, V. Gayathri, K. E. S. Ford, Z. Haiman, S. Klimenko, B. Kocsis, S. Márka, Z. Márka, B. McKernan, and R. O'Shaughnessy. Hierarchical Black Hole Mergers in Active Galactic Nuclei. *prl*, 123(18):181101, November 2019.

[336] Z. Doctor, D. Wysocki, R. O'Shaughnessy, D. E. Holz, and B. Farr. Black Hole Coagulation: Modeling Hierarchical Mergers in Black Hole Populations. , 893(1):35, April 2020.

[337] Bill Paxton, Lars Bildsten, Aaron Dotter, Falk Herwig, Pierre Lesaffre, and Frank Timmes. Modules for Experiments in Stellar Astrophysics (MESA). *apjs*, 192(1):3, January 2011.

[338] Mark Vogelsberger, Shy Genel, Volker Springel, Paul Torrey, Debora Sijacki, Dandan Xu, Greg Snyder, Dylan Nelson, and Lars Hernquist. Introducing the Illustris Project: simulating the coevolution of dark and visible matter in the Universe. *mnras*, 444(2):1518–1547, October 2014.

[339] Philip F. Hopkins et al. FIRE-3: updated stellar evolution models, yields, and microphysics and fitting functions for applications in galaxy simulations. *mnras*, 519(2):3154–3181, February 2023.

[340] R. Farmer, M. Renzo, S. E. de Mink, P. Marchant, and S. Justham. Mind the Gap: The Location of the Lower Edge of the Pair-instability Supernova Black Hole Mass Gap. , 887(1):53, December 2019.

[341] Suvodip Mukherjee. The redshift dependence of black hole mass distribution: is it reliable for standard sirens cosmology? , 515(4):5495–5505, October 2022.

[342] Christos Karathanasis, Suvodip Mukherjee, and Simone Mastrogiovanni. Binary black holes population and cosmology in new lights: signature of PISN mass and formation channel in GWTC-3. , 523(3):4539–4555, August 2023.

[343] Risa H. Wechsler and Jeremy L. Tinker. The Connection between Galaxies and their Dark Matter Halos. *Ann. Rev. Astron. Astrophys.*, 56:435–487, 2018.

[344] A. J. Benson, S. Cole, C. S. Frenk, C. M. Baugh, and Cedric G. Lacey. The Nature of galaxy bias and clustering. *Mon. Not. Roy. Astron. Soc.*, 311:793–808, 2000.

[345] Ravi K. Sheth and Giuseppe Tormen. Large scale bias and the peak background split. *Mon. Not. Roy. Astron. Soc.*, 308:119, 1999.

[346] Shaun Cole, Cedric G. Lacey, Carlton M. Baugh, and Carlos S. Frenk. Hierarchical galaxy formation. , 319(1):168–204, November 2000.

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

[348] A. Vale and J. P. Ostriker. Linking halo mass to galaxy luminosity. , 353(1):189–200, September 2004.

[349] Liang Gao and Simon D. M. White. Assembly bias in the clustering of dark matter haloes. *Mon. Not. Roy. Astron. Soc.*, 377:L5–L9, 2007.

[350] Darren J. Croton, Liang Gao, and Simon D. M. White. Halo assembly bias and its effects on galaxy clustering. *Mon. Not. Roy. Astron. Soc.*, 374:1303–1309, 2007.

[351] Neal Dalal, Martin White, J. Richard Bond, and Alexander Shirokov. Halo Assembly Bias in Hierarchical Structure Formation. , 687(1):12–21, November 2008.

[352] Benjamin P. Moster, Rachel S. Somerville, Christian Maulbetsch, Frank C. van den Bosch, Andrea V. Macciò, Thorsten Naab, and Ludwig Oser. Constraints on the Relationship between Stellar Mass and Halo Mass at Low and High Redshift. , 710(2):903–923, February 2010.

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

[354] Peter S. Behroozi, Risa H. Wechsler, and Charlie Conroy. The Average Star Formation Histories of Galaxies in Dark Matter Halos from z = 0-8. , 770(1):57, June 2013.

[355] Peter Behroozi, Risa H. Wechsler, Andrew P. Hearin, and Charlie Conroy. UNIVERSEMACHINE: The correlation between galaxy growth and dark matter halo assembly from z = 0-10. , 488(3):3143–3194, September 2019.

[356] Giulio Scelfo, Nicola Bellomo, Alvise Raccanelli, Sabino Matarrese, and Licia Verde. GW×LSS: chasing the progenitors of merging binary black holes. *JCAP*, 09:039, 2018.

[357] Suvodip Mukherjee and Joseph Silk. Time-dependence of the astrophysical stochastic gravitational wave background. *Mon. Not. Roy. Astron. Soc.*, 491(4):4690–4701, 2020.

[358] Francesca Calore, Alessandro Cuoco, Tania Regimbau, Surabhi Sachdev, and Pasquale Dario Serpico. Cross-correlating galaxy catalogs and gravitational waves: a tomographic approach. *Phys. Rev. Res.*, 2:023314, 2020.

[359] Cristina Cigarran Diaz and Suvodip Mukherjee. Mapping the cosmic expansion history from LIGO-Virgo-KAGRA in synergy with DESI and SPHEREx. *Mon. Not. Roy. Astron. Soc.*, 511(2):2782–2795, 2022.

[360] Sarah Libanore, Maria Celeste Artale, Dionysios Karagiannis, Michele Liguori, Nicola Bartolo, Yann Bouffanais, Michela Mapelli, and Sabino Matarrese. Clustering of Gravitational Wave and Supernovae events: a multitracer analysis in Luminosity Distance Space. *JCAP*, 02(02):003, 2022.

[361] Thorsten Naab and Jeremiah P. Ostriker. Theoretical challenges in galaxy formation. *Annual Review of Astronomy and Astrophysics*, 55(1):59–109, August 2017.

[362] Kentaro Nagamine. Feedback models in galaxy simulations and probing their impact by cosmological hydrodynamic simulations. In Tony Wong and Woong-Tae Kim, editors, *Resolving the Rise and Fall of Star Formation in Galaxies*, volume 373 of *IAU Symposium*, pages 283–292, January 2023.

[363] Laura V. Sales, Andrew Wetzel, and Azadeh Fattahi. Baryonic solutions and challenges for cosmological models of dwarf galaxies. *Nature Astron.*, 6(8):897–910, 2022.

[364] S. Libanore, M. C. Artale, D. Karagiannis, M. Liguori, N. Bartolo, Y. Bouffanais, N. Giacobbo, M. Mapelli, and S. Matarrese. Gravitational wave mergers as tracers of large scale structures. *Journal of Cosmology and Astroparticle Physics*, 2021(02):035–035, February 2021.

[365] Matteo Peron, Andrea Ravenni, Sarah Libanore, Michele Liguori, and Maria Celeste Artale. Clustering of binary black hole mergers: a detailed analysis of the EAGLE + MOBSE simulation. , 530(1):1129–1143, May 2024.

[366] Giulio Scelfo, Lumen Boco, Andrea Lapi, and Matteo Viel. Exploring galaxies-gravitational waves cross-correlations as an astrophysical probe. *JCAP*, 10:045, 2020.

[367] Suvodip Mukherjee and Azadeh Moradinezhad Dizgah. Toward a Precision Measurement of Binary Black Holes Formation Channels Using Gravitational Waves and Emission Lines. *Astrophys. J. Lett.*, 937(2):L27, 2022.

[368] Toshiya Namikawa, Atsushi Nishizawa, and Atsushi Taruya. Detecting black-hole binary clustering via the second-generation gravitational-wave detectors. , 94(2):024013, July 2016.

[369] Suvodip Mukherjee, Benjamin D. Wandelt, and Joseph Silk. Probing the theory of gravity with gravitational lensing of gravitational waves and galaxy surveys. *Mon. Not. Roy. Astron. Soc.*, 494(2):1956–1970, 2020.

[370] Suvodip Mukherjee, Benjamin D. Wandelt, and Joseph Silk. Testing the general theory of relativity using gravitational wave propagation from dark standard sirens. *Mon. Not. Roy. Astron. Soc.*, 502(1):1136–1144, 2021.

[371] S. Libanore, M.C. Artale, D. Karagiannis, M. Liguori, N. Bartolo, Y. Bouffanais, M. Mapelli, and S. Matarrese. Clustering of gravitational wave and supernovae events: a multitracer analysis in luminosity distance space. *Journal of Cosmology and Astroparticle Physics*, 2022(02):003, February 2022.

[372] E. L. Gagnon, D. Anbajagane, J. Prat, C. Chang, and J. Frieman. Cosmological Constraints from Combining Galaxy Surveys and Gravitational Wave Observatories. *arXiv e-prints*, page arXiv:2312.16289, December 2023.

[373] Alvise Raccanelli, Ely D. Kovetz, Simeon Bird, Ilias Cholis, and Julian B. Muñoz. Determining the progenitors of merging black-hole binaries. , 94(2):023516, July 2016.

[374] Giulio Scelfo, Nicola Bellomo, Alvise Raccanelli, Sabino Matarrese, and Licia Verde. GW×LSS: chasing the progenitors of merging binary black holes. *"jcap"*, 2018(9):039, September 2018.

[375] Sarah Libanore, Michele Liguori, and Alvise Raccanelli. Signatures of primordial black holes in gravitational wave clustering. *Journal of Cosmology and Astroparticle Physics*, 2023(08):055, August 2023.

[376] Anthony Challinor and Antony Lewis. Linear power spectrum of observed source number counts. , 84(4):043516, August 2011.

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

[378] Cullan Howlett, Antony Lewis, Alex Hall, and Anthony Challinor. Cmb power spectrum parameter degeneracies in the era of precision cosmology. *Journal of Cosmology and Astroparticle Physics*, 2012(04):027–027, April 2012.

[379] Anthony Challinor and Antony Lewis. Lensed CMB power spectra from all-sky correlation functions. , 71(10):103010, May 2005.

[380] Piero Madau and Mark Dickinson. Cosmic star-formation history. *Annual Review of Astronomy and Astrophysics*, 52:415–486, 2014.

[381] Ashley J. Ross, Lado Samushia, Cullan Howlett, Will J. Percival, Angela Burden, and Marc Manera. The clustering of the SDSS DR7 main galaxy sample – i. a 4 per cent distance measure at z = 0.15. *Monthly Notices of the Royal Astronomical Society*, 449(1):835–847, mar 2015.

[382] G. Dálya et al. GLADE+ : an extended galaxy catalogue for multimessenger searches with advanced gravitational-wave detectors. *Mon. Not. Roy. Astron. Soc.*, 514(1):1403–1411, 2022.

[383] Maciej Bilicki, Thomas H. Jarrett, John A. Peacock, Michelle E. Cluver, and Louise Steward. 2MASS Photometric Redshift catalog: a comprehensive three-dimensional census of the whole sky. *Astrophys. J. Suppl.*, 210:9, 2014.

[384] T. H. Jarrett, T. Chester, R. Cutri, S. Schneider, M. Skrutskie, and J. P. Huchra. 2mass extended source catalog: overview and algorithms. *Astron. J.*, 119:2498–2531, 2000.

[385] Skrutskie, and others. The Two Micron All Sky Survey (2MASS). , 131(2):1163–1183, February 2006.

[386] Wright, and others. The Wide-field Infrared Survey Explorer (WISE): Mission Description and Initial On-orbit Performance. , 140(6):1868–1881, December 2010.

[387] R. M. Cutri and et al. VizieR Online Data Catalog: WISE All-Sky Data Release (Cutri+ 2012). *VizieR Online Data Catalog*, page II/311, April 2012.

[388] Nigel Hambly et al. The supercosmos sky survey. paper I: introduction and description. *Mon. Not. Roy. Astron. Soc.*, 326:1279, 2001.

[389] Nigel Hambly, Mike Irwin, and Harvey MacGillivray. The supercosmos sky survey. paper. 2. Image detection, parameterisation, classification and photometry. *Mon. Not. Roy. Astron. Soc.*, 326:1295, 2001.

[390] Nigel Hambly, Clive Davenhall, Mike Irwin, and Harvey MacGillivray. The supercosmos sky survey. paper. 3. Astrometry. *Mon. Not. Roy. Astron. Soc.*, 326:1315, 2001.

[391] Adrian A. Collister and Ofer Lahav. ANNz: Estimating Photometric Redshifts Using Artificial Neural Networks. , 116(818):345–351, April 2004.

[392] Christopher P Ahn, Rachael Alexandroff, Carlos Allende Prieto, Scott F Anderson, Timothy Anderton, Brett H Andrews, Éric Aubourg, Stephen Bailey, Eduardo Balbinot, Rory Barnes, et al. The Ninth Data Release of the Sloan Digital Sky Survey: First Spectroscopic Data from the SDSS-III Baryon Oscillation Spectroscopic Survey. , 203(2):21, December 2012.

[393] Jones, and others. The 6dF Galaxy Survey: final redshift release (DR3) and southern large-scale structures. , 399(2):683–698, October 2009.

[394] Matthew Colless et al. The 2dF Galaxy Redshift Survey: Spectra and redshifts. *Mon. Not. Roy. Astron. Soc.*, 328:1039, 2001.

[395] Thomas Jarrett. Large Scale Structure in the Local Universe - The 2MASS Galaxy Catalog. *Publ. Astron. Soc. Austral.*, 21(4):396–403, January 2004.

[396] Eric F. Bell, Daniel H. McIntosh, Neal Katz, and Martin D. Weinberg. The optical and near-infrared properties of galaxies. 1. Luminosity and stellar mass functions. *Astrophys. J. Suppl.*, 149:289, 2003.

[397] M. Bilicki et al. WISE x SuperCOSMOS photometric redshift catalog: 20 million galaxies over 3pi steradians. *Astrophys. J. Suppl.*, 225:5, 2016.

[398] et.al. Cutri. VizieR Online Data Catalog: AllWISE Data Release (Cutri+ 2013). *VizieR Online Data Catalog*, page II/328, February 2021.

[399] Simon P Driver, Peder Norberg, Ivan K Baldry, Steven P Bamford, Andrew M Hopkins, Jochen Liske, Jon Loveday, and John A Peacock. GAMA: towards a physical understanding of galaxy formation. *Astronomy and Geophysics*, 50(5):5.12–5.19, October 2009.

[400] Liske, and others. Galaxy And Mass Assembly (GAMA): end of survey report and data release 2. , 452(2):2087–2126, September 2015.

[401] K. M. Górski, E. Hivon, A. J. Banday, B. D. Wandelt, F. K. Hansen, M. Reinecke, and M. Bartelmann. HEALPix: A Framework for High-Resolution Discretization and Fast Analysis of Data Distributed on the Sphere. , 622:759–771, April 2005.

[402] Andrea Zonca, Leo Singer, Daniel Lenz, Martin Reinecke, Cyrille Rosset, Eric Hivon, and Krzysztof Gorski. healpy: equal area pixelization and spherical harmonics transforms for data on the sphere in python. *Journal of Open Source Software*, 4(35):1298, March 2019.

[403] A. Balaguera-Antolínez, M. Bilicki, E. Branchini, and A. Postiglione. Extracting cosmological information from the angular power spectrum of the 2MASS Photometric Redshift catalogue. *Mon. Not. Roy. Astron. Soc.*, 476(1):1050–1070, 2018.

[404] H. S. Xavier, M. V. Costa-Duarte, A. Balaguera-Antolínez, and M. Bilicki. All-sky angular power spectra from cleaned WISE×SuperCOSMOS galaxy number counts. *JCAP*, 08:037, 2019.

[405] T. Kettlety, J. Hesling, S. Phillipps, M. N. Bremer, M. E. Cluver, E. N. Taylor, J. Bland-Hawthorn, S. Brough, R. De Propris, S. P. Driver, B. W. Holwerda, L. S. Kelvin, W. Sutherland, and A. H. Wright. Galaxy and mass assembly (GAMA): the consistency of GAMA and WISE derived mass-to-light ratios. , 473(1):776–783, January 2018.

[406] Edward L. Fitzpatrick. Correcting for the Effects of Interstellar Extinction. , 111(755):63–75, January 1999.

[407] Sambaran Banerjee, Holger Baumgardt, and Pavel Kroupa. Stellar-mass black holes in star clusters: implications for gravitational wave radiation. , 402(1):371–380, February 2010.

[408] Suvodip Mukherjee, Tom Broadhurst, Jose M Diego, Joseph Silk, and George F Smoot. Impact of astrophysical binary coalescence time-scales on the rate of lensed gravitational wave events. *Mon. Not. Roy. Astron. Soc.*, 506(3):3751–3759, 2021.

[409] Maya Fishbach and Vicky Kalogera. The Time Delay Distribution and Formation Metallicity of LIGO-Virgo's Binary Black Holes. *Astrophys. J. Lett.*, 914(2):L30, 2021.

[410] B. P. Abbott et al. GW190425: Observation of a Compact Binary Coalescence with Total Mass $\sim 3.4 M_{\odot}$. *Astrophys. J. Lett.*, 892(1):L3, 2020.

[411] Martyna Chruslinska, Gijs Nelemans, and Krzysztof Belczynski. The influence of the distribution of cosmic star formation at different metallicities on the properties of

merging double compact objects. *Mon. Not. Roy. Astron. Soc.*, 482(4):5012–5017, 2019.

[412] Martyna Chruślińska. Chemical Evolution of the Universe and its Consequences for Gravitational-Wave Astrophysics. *Annalen Phys.*, 536(2):2200170, 2024.

[413] M Celeste Artale, Yann Bouffanais, Michela Mapelli, Nicola Giacobbo, Nadeen B Sabha, Filippo Santoliquido, Mario Pasquato, and Mario Spera. An astrophysically motivated ranking criterion for low-latency electromagnetic follow-up of gravitational wave events. *Monthly Notices of the Royal Astronomical Society*, 495(2):1841–1852, May 2020.

[414] Liana Rauf, Cullan Howlett, Tamara M. Davis, and Claudia D. P. Lagos. Exploring binary black hole mergers and host galaxies with SHARK and COMPAS. , 523(4):5719–5737, August 2023.

[415] Salim, and others. UV Star Formation Rates in the Local Universe. , 173(2):267–292, December 2007.

[416] F. Mannucci, G. Cresci, R. Maiolino, A. Marconi, and A. Gnerucci. A fundamental relation between mass, star formation rate and metallicity in local and high-redshift galaxies. , 408(4):2115–2127, November 2010.

[417] Peng, and others. Mass and Environment as Drivers of Galaxy Evolution in SDSS and zCOSMOS and the Origin of the Schechter Function. , 721(1):193–221, September 2010.

[418] Michela Mapelli, Nicola Giacobbo, Mattia Toffano, Emanuele Ripamonti, Alessandro Bressan, Mario Spera, and Marica Branchesi. The host galaxies of double compact objects merging in the local Universe. , 481(4):5324–5330, December 2018.

[419] Qi Guo, Simon White, Michael Boylan-Kolchin, Gabriella De Lucia, Guinevere Kauffmann, Gerard Lemson, Cheng Li, Volker Springel, and Simone Weinmann. From dwarf spheroidals to cD galaxies: simulating the galaxy population in a ΛCDM cosmology. , 413(1):101–131, May 2011.

[420] Pavel Kroupa. On the variation of the initial mass function. , 322(2):231–246, April 2001.

[421] Jeffrey E. McClintock and Ronald A. Remillard. Black hole binaries. In *Compact stellar X-ray sources*, volume 39, pages 157–213. , 2006.

[422] K. M. Górski, E. Hivon, A. J. Banday, B. D. Wandelt, F. K. Hansen, M. Reinecke, and M. Bartelmann. HEALPix: A Framework for High-Resolution Discretization and Fast Analysis of Data Distributed on the Sphere. , 622(2):759–771, April 2005.

[423] G. Efstathiou. A maximum likelihood analysis of the low cosmic microwave background multipoles from the Wilkinson Microwave Anisotropy Probe. , 348(3):885–896, March 2004.

[424] Lloyd Knox. Determination of inflationary observables by cosmic microwave background anisotropy experiments. , 52(8):4307–4318, October 1995.

[425] Hebertt Leandro, Valerio Marra, and Riccardo Sturani. Measuring the hubble constant with black sirens. *Physical Review D*, 105(2), January 2022.

[426] David H. Weinberg, Romeel Davé, Neal Katz, and Lars Hernquist. Galaxy Clustering and Galaxy Bias in a ΛCDM Universe. , 601(1):1–21, January 2004.

[427] J. A. Peacock and M. Bilicki. Wide-area tomography of CMB lensing and the growth of cosmological density fluctuations. , 481(1):1133–1148, November 2018.

[428] A. Balaguera-Antolínez, M. Bilicki, E. Branchini, and A. Postiglione. Extracting cosmological information from the angular power spectrum of the 2MASS Photometric Redshift catalogue. , 476(1):1050–1070, May 2018.

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

[430] Peter S. Behroozi, Charlie Conroy, and Risa H. Wechsler. A Comprehensive Analysis of Uncertainties Affecting the Stellar Mass-Halo Mass Relation for $0 < z < 4$. , 717(1):379–403, July 2010.

[431] P. Schechter. An analytic expression for the luminosity function for galaxies. , 203:297–306, January 1976.

[432] Cole, Shaun and others. The 2dF galaxy redshift survey: near-infrared galaxy luminosity functions. , 326(1):255–273, September 2001.

[433] Arjun Dey, David J Schlegel, Dustin Lang, Robert Blum, Kaylan Burleigh, Xiaohui Fan, Joseph R Findlay, Doug Finkbeiner, David Herrera, Stéphanie Juneau, et al. Overview of the desi legacy imaging surveys. *The Astronomical Journal*, 157(5):168, 2019.

[434] Euclid Collaboration, R Scaramella, et al. Euclid preparation-i. the euclid wide survey. *A&A*, 662:A112, 2022.

[435] Željko Ivezić, Steven M Kahn, J Anthony Tyson, Bob Abel, Emily Acosta, Robyn Allsman, David Alonso, Yusra AlSayyad, Scott F Anderson, John Andrew, et al. Lsst: from science drivers to reference design and anticipated data products. *The Astrophysical Journal*, 873(2):111, 2019.

[436] Olivier Doré et al. WFIRST: The Essential Cosmology Space Observatory for the Coming Decade. , 4 2019.

[437] J. E. Gunn et al. The Sloan Digital Sky Survey Photometric Camera. *aj*, 116(6):3040–3081, December 1998.

[438] J. E. Gunn et al. The 2.5 m Telescope of the Sloan Digital Sky Survey. *aj*, 131(4):2332–2359, April 2006.

[439] Nikhil Padmanabhan, David J Schlegel, Douglas P Finkbeiner, JC Barentine, Michael R Blanton, Howard J Brewington, James E Gunn, Michael Harvanek, David W Hogg, Željko Ivezić, et al. An improved photometric calibration of the sloan digital sky survey imaging data. *The Astrophysical Journal*, 674(2):1217, 2008.

[440] K. N. Abazajian et al. The Seventh Data Release of the Sloan Digital Sky Survey. *apjs*, 182(2):543–558, June 2009.

[441] M. A. Strauss et al. Spectroscopic Target Selection in the Sloan Digital Sky Survey: The Main Galaxy Sample. *aj*, 124(3):1810–1824, September 2002.

[442] David J Schlegel, Douglas P Finkbeiner, and Marc Davis. Maps of dust infrared emission for use in estimation of reddening and cosmic microwave background radiation foregrounds. *The Astrophysical Journal*, 500(2):525, 1998.

[443] Michael A Strauss, David H Weinberg, Robert H Lupton, Vijay K Narayanan, James Annis, Mariangela Bernardi, Michael Blanton, Scott Burles, AJ Connolly, Julianne Dalcanton, et al. Spectroscopic target selection in the sloan digital sky survey: the main galaxy sample. *The Astronomical Journal*, 124(3):1810, 2002.

[444] M. R. Blanton et al. New York University Value-Added Galaxy Catalog: A Galaxy Catalog Based on New Public Surveys. *aj*, 129(6):2562–2578, June 2005.

[445] Michael R. Blanton, David J. Schlegel, Michael A. Strauss, J. Brinkmann, Douglas Finkbeiner, Masataka Fukugita, James E. Gunn, David W. Hogg, Željko Ivezić, G. R.

Knapp, Robert H. Lupton, Jeffrey A. Munn, Donald P. Schneider, Max Tegmark, and Idit Zehavi. New york university value-added galaxy catalog: A galaxy catalog based on new public surveys. *The Astronomical Journal*, 129(6):2562–2578, June 2005.

[446] Michael R. Blanton and Sam Roweis. K-corrections and filter transformations in the ultraviolet, optical, and near infrared. *Astron. J.*, 133:734–754, 2007.

[447] Yu-Yen Chang, Arjen van der Wel, Elisabete da Cunha, and Hans-Walter Rix. Stellar Masses and Star Formation Rates for 1M Galaxies from SDSS+WISE. *apjs*, 219(1):8, July 2015.

[448] Anna K. Weigel, Kevin Schawinski, and Claudio Bruderer. Stellar mass functions: methods, systematics and results for the local Universe. *mnras*, 459(2):2150–2187, June 2016.

[449] M. Celeste Artale, Michela Mapelli, Nicola Giacobbo, Nadeen B. Sabha, Mario Spera, Filippo Santoliquido, and Alessandro Bressan. Host galaxies of merging compact objects: mass, star formation rate, metallicity, and colours. , 487(2):1675–1688, August 2019.

[450] Aditya Vijaykumar, Maya Fishbach, Susmita Adhikari, and Daniel E. Holz. Inferring host galaxy properties of LIGO-Virgo-KAGRA's black holes. *arXiv e-prints*, page arXiv:2312.03316, December 2023.

[451] Kazuhiro Yamamoto, Masashi Nakamichi, Akinari Kamino, Bruce A. Bassett, and Hiroaki Nishioka. A Measurement of the Quadrupole Power Spectrum in the Clustering of the 2dF QSO Survey. *pasj*, 58:93–102, February 2006.

[452] Hume A. Feldman, Nick Kaiser, and John A. Peacock. Power-spectrum analysis of three-dimensional redshift surveys. *The Astrophysical Journal*, 426:23, May 1994.

[453] Digvijay Wadekar and Román Scoccimarro. Galaxy power spectrum multipoles covariance in perturbation theory. *prd*, 102(12):123517, December 2020.

[454] Otavio Alves and DESI Collaboration. Analytical covariance matrices of desi galaxy power spectrum multipoles. (in prep.), 2024.

[455] Nick Kaiser. Clustering in real space and in redshift space. *mnras*, 227:1–21, July 1987.

[456] Cullan Howlett, Ashley J. Ross, Lado Samushia, Will J. Percival, and Marc Manera. The clustering of the sdss main galaxy sample – ii. mock galaxy catalogues and a

measurement of the growth of structure from redshift space distortions at z = 0.15. *Monthly Notices of the Royal Astronomical Society*, 449(1):848–866, March 2015.

[457] A. J. Ross et al. The clustering of galaxies in the completed sdss-iii baryon oscillation spectroscopic survey: observational systematics and baryon acoustic oscillations in the correlation function. *Monthly Notices of the Royal Astronomical Society*, 464(1):1168–1191, September 2016.

[458] Karun Thanjavur, Luc Simard, Asa F. L. Bluck, and Trevor Mendel. Stellar mass functions of galaxies, discs and spheroids at $z \sim 0.1$. *Monthly Notices of the Royal Astronomical Society*, 459(1):44–69, March 2016.

[459] Thibaut Louis et al. The Atacama Cosmology Telescope: DR6 Power Spectra, Likelihoods and ΛCDM Parameters. , 3 2025.

[460] Erminia Calabrese et al. The Atacama Cosmology Telescope: DR6 Constraints on Extended Cosmological Models. , 3 2025.

[461] Kana Moriwaki, Takahiro Nishimichi, and Naoki Yoshida. Machine learning for observational cosmology. *Reports on Progress in Physics*, 86(7):076901, May 2023.

[462] M. Abdul Karim et al. DESI DR2 Results II: Measurements of Baryon Acoustic Oscillations and Cosmological Constraints. , 3 2025.

[463] D. J. Fixsen, E. S. Cheng, J. M. Gales, John C. Mather, R. A. Shafer, and E. L. Wright. The Cosmic Microwave Background spectrum from the full COBE FIRAS data set. *Astrophys. J.*, 473:576, 1996.

[464] G. Rakavy and G. Shaviv. Instabilities in Highly Evolved Stellar Models. *apj*, 148:803, June 1967.

[465] R. Abbott et al. Gw190521: A binary black hole merger with a total mass of 150 $M_{\odot}$. *Phys. Rev. Lett.*, 125:101102, Sep 2020.

APPENDICES

Appendix A

Three-point functions for two locally non-Gaussian scalar fields

Here we show how (3.131) is derived from (3.129). In order to make the derivation easy to follow, we take into account the terms in (3.129) one by one and derive the final expression of (3.131) step by step.

A.1 One field

Let us consider the following[1],

$$\zeta(\boldsymbol{x}) = \zeta_{\rm G}(\boldsymbol{x}) + f_{\rm NL}^{\zeta}\left(\zeta_{\rm G}(\boldsymbol{x})^2 - \langle\zeta_{\rm G}(\boldsymbol{x})^2\rangle\right) + \cdots\,, \tag{A.1}$$

and evaluate $\zeta_{\boldsymbol{k}}$, the Fourier transform of ζ,

$$\begin{aligned}\zeta_{\boldsymbol{k}} = \int \mathrm{d}^3\boldsymbol{x}\,\zeta(\boldsymbol{x})e^{-i\boldsymbol{k}\cdot\boldsymbol{x}} = \int \mathrm{d}^3\boldsymbol{x}\,\zeta_{\rm G}(\boldsymbol{x})e^{-i\boldsymbol{k}\cdot\boldsymbol{x}} \\ + f_{\rm NL}^{\zeta}\left(\int \mathrm{d}^3\boldsymbol{x}\,\zeta_{\rm G}(\boldsymbol{x})^2 e^{-i\boldsymbol{k}\cdot\boldsymbol{x}} - \int \mathrm{d}^3\boldsymbol{x}\,\langle\zeta_{\rm G}(\boldsymbol{x})^2\rangle e^{-i\boldsymbol{k}\cdot\boldsymbol{x}}\right).\end{aligned} \tag{A.2}$$

[1]Note that $\langle\zeta_{\rm G}(\boldsymbol{x})^2\rangle$ is essentially a constant ensuring that $\langle\zeta(\boldsymbol{x})\rangle = 0$ to quadratic order, which fixes the gauge freedom in defining ζ. However, when $\zeta_{\rm G}$ has a scale-invariant or red power spectrum, the real-space 2-point function presents an IR divergence, resulting in $\langle\zeta_{\rm G}(\boldsymbol{x})^2\rangle$ being infinite. As shown below in the bispectrum calculation, this term matters only for $k = 0$. Yet, both theoretically and observationally, we always have to consider an IR cut-off.

The first term simply evaluates to $\zeta^{\rm G}_{\boldsymbol{k}}$, the Fourier transform of $\zeta_{\rm G}(\boldsymbol{x})$. For the second term, after inserting $\zeta_{\rm G}(\boldsymbol{x}) = \int_{\boldsymbol{p}} \zeta^{\rm G}_{\boldsymbol{p}} \exp(i\boldsymbol{p}\cdot\boldsymbol{x})$, we have

$$\begin{aligned}\int \mathrm{d}^3\boldsymbol{x}\, \zeta_{\rm G}(\boldsymbol{x})^2 e^{-i\boldsymbol{k}\cdot\boldsymbol{x}} &= \int \mathrm{d}^3\boldsymbol{x} \int_{\boldsymbol{p}} \zeta^{\rm G}_{\boldsymbol{p}} e^{i\boldsymbol{p}\cdot\boldsymbol{x}} \int_{\boldsymbol{q}} \zeta^{\rm G}_{\boldsymbol{q}} e^{i\boldsymbol{q}\cdot\boldsymbol{x}} e^{-i\boldsymbol{k}\cdot\boldsymbol{x}} = \int_{\boldsymbol{p}}\int_{\boldsymbol{q}} \zeta^{\rm G}_{\boldsymbol{p}}\zeta^{\rm G}_{\boldsymbol{q}} \int \mathrm{d}^3\boldsymbol{x}\, e^{i(\boldsymbol{p}+\boldsymbol{q}-\boldsymbol{k})\cdot\boldsymbol{x}} \\ &= \int_{\boldsymbol{p}}\int_{\boldsymbol{q}} \zeta^{\rm G}_{\boldsymbol{p}}\zeta^{\rm G}_{\boldsymbol{q}} (2\pi)^3\delta^{(3)}(\boldsymbol{p}+\boldsymbol{q}-\boldsymbol{k}) = \int_{\boldsymbol{p}} \zeta^{\rm G}_{\boldsymbol{p}}\zeta^{\rm G}_{\boldsymbol{k}-\boldsymbol{p}}\,, \end{aligned} \tag{A.3}$$

and for the last term,

$$\int \mathrm{d}^3\boldsymbol{x}\, \langle\zeta_{\rm G}(\boldsymbol{x})^2\rangle e^{-i\boldsymbol{k}\cdot\boldsymbol{x}} = \langle\zeta_{\rm G}(\boldsymbol{x})^2\rangle (2\pi)^3\delta^{(3)}(\boldsymbol{k})\,, \tag{A.4}$$

which as mentioned before drops out for all values of k except for $k = 0$. Putting all the terms together, we can write

$$\zeta_{\boldsymbol{k}} = \zeta^{\rm G}_{\boldsymbol{k}} + \zeta^{\rm NL}_{\boldsymbol{k}}\,, \qquad \zeta^{\rm NL}_{\boldsymbol{k}} \equiv f^{\zeta}_{\rm NL}\int_{\boldsymbol{p}} \zeta^{\rm G}_{\boldsymbol{p}}\zeta^{\rm G}_{\boldsymbol{k}-\boldsymbol{p}} - (2\pi)^3 f^{\zeta}_{\rm NL}\langle\zeta_{\rm G}(\boldsymbol{x})^2\rangle\delta^{(3)}(\boldsymbol{k})\,. \tag{A.5}$$

Now let us calculate $\langle\zeta_{\boldsymbol{k}_1}\zeta_{\boldsymbol{k}_2}\zeta_{\boldsymbol{k}_3}\rangle$. Inserting the above, we have

$$\begin{aligned}\langle\zeta_{\boldsymbol{k}_1}\zeta_{\boldsymbol{k}_2}\zeta_{\boldsymbol{k}_3}\rangle &= \langle(\zeta^{\rm G}_{\boldsymbol{k}_1} + \zeta^{\rm NL}_{\boldsymbol{k}_1})(\zeta^{\rm G}_{\boldsymbol{k}_2} + \zeta^{\rm NL}_{\boldsymbol{k}_2})(\zeta^{\rm G}_{\boldsymbol{k}_3} + \zeta^{\rm NL}_{\boldsymbol{k}_3})\rangle \\ &= \langle\zeta^{\rm G}_{\boldsymbol{k}_1}\zeta^{\rm G}_{\boldsymbol{k}_2}\zeta^{\rm NL}_{\boldsymbol{k}_3}\rangle + [\boldsymbol{k}_2\leftrightarrow\boldsymbol{k}_3] + [\boldsymbol{k}_1\leftrightarrow\boldsymbol{k}_3] + \mathcal{O}\left((\zeta^{\rm NL}_{\boldsymbol{k}})^2\right)\,, \end{aligned} \tag{A.6}$$

where we set $\langle\zeta^{\rm G}_{\boldsymbol{k}_1}\zeta^{\rm G}_{\boldsymbol{k}_2}\zeta^{\rm G}_{\boldsymbol{k}_3}\rangle = 0$ based on Wick's theorem. Let us then focus on evaluating $\langle\zeta^{\rm G}_{\boldsymbol{k}_1}\zeta^{\rm G}_{\boldsymbol{k}_2}\zeta^{\rm NL}_{\boldsymbol{k}_3}\rangle$. Inserting (A.5), we have

$$\begin{aligned}\langle\zeta^{\rm G}_{\boldsymbol{k}_1}\zeta^{\rm G}_{\boldsymbol{k}_2}\zeta^{\rm NL}_{\boldsymbol{k}_3}\rangle &= f^{\zeta}_{\rm NL}\int_{\boldsymbol{p}} \langle\zeta^{\rm G}_{\boldsymbol{k}_1}\zeta^{\rm G}_{\boldsymbol{k}_2}\zeta^{\rm G}_{\boldsymbol{p}}\zeta^{\rm G}_{\boldsymbol{k}_3-\boldsymbol{p}}\rangle - \langle\zeta^{\rm G}_{\boldsymbol{k}_1}\zeta^{\rm G}_{\boldsymbol{k}_2}\rangle(2\pi)^3 f^{\zeta}_{\rm NL}\langle\zeta_{\rm G}(\boldsymbol{x})^2\rangle\delta^{(3)}(\boldsymbol{k}_3) \\ &= f^{\zeta}_{\rm NL}\int_{\boldsymbol{p}} \left(\langle\zeta^{\rm G}_{\boldsymbol{k}_1}\zeta^{\rm G}_{\boldsymbol{p}}\rangle\langle\zeta^{\rm G}_{\boldsymbol{k}_2}\zeta^{\rm G}_{\boldsymbol{k}_3-\boldsymbol{p}}\rangle + \langle\zeta^{\rm G}_{\boldsymbol{k}_1}\zeta^{\rm G}_{\boldsymbol{k}_3-\boldsymbol{p}}\rangle\langle\zeta^{\rm G}_{\boldsymbol{k}_2}\zeta^{\rm G}_{\boldsymbol{p}}\rangle + \langle\zeta^{\rm G}_{\boldsymbol{k}_1}\zeta^{\rm G}_{\boldsymbol{k}_2}\rangle\langle\zeta^{\rm G}_{\boldsymbol{p}}\zeta^{\rm G}_{\boldsymbol{k}_3-\boldsymbol{p}}\rangle\right) \\ &\quad - \langle\zeta^{\rm G}_{\boldsymbol{k}_1}\zeta^{\rm G}_{\boldsymbol{k}_2}\rangle(2\pi)^3 f^{\zeta}_{\rm NL}\langle\zeta_{\rm G}(\boldsymbol{x})^2\rangle\,\delta^{(3)}(\boldsymbol{k}_3) \\ &= 2(2\pi)^3 f^{\zeta}_{\rm NL}P_{\zeta}(k_1)P_{\zeta}(k_2)\,\delta^{(3)}(\boldsymbol{k}_1+\boldsymbol{k}_2+\boldsymbol{k}_3) \\ &\quad + (2\pi)^6 f^{\zeta}_{\rm NL}P_{\zeta}(k_1)\,\delta^{(3)}(\boldsymbol{k}_1+\boldsymbol{k}_2)\delta^{(3)}(\boldsymbol{k}_3)\left(\int_{\boldsymbol{p}} P_{\zeta}(p) - \langle\zeta_{\rm G}(\boldsymbol{x})^2\rangle\right)\,, \end{aligned} \tag{A.7}$$

where the second line follows from Wick's theorem, and in the last two lines, the definition of the power spectrum from (3.132) is used. Noting that $\int_{\boldsymbol{p}} P_{\zeta}(p) = \langle\zeta_{\rm G}(\boldsymbol{x})^2\rangle$ is the Fourier transform of the 2-point function, the last term of (A.7) vanishes regardless of the value of $\boldsymbol{k}_3$. Together with (A.6), we thus have

$$\langle\zeta_{\boldsymbol{k}_1}\zeta_{\boldsymbol{k}_2}\zeta_{\boldsymbol{k}_3}\rangle = 2(2\pi)^3 f^{\zeta}_{\rm NL}\left[P_{\zeta}(k_1)P_{\zeta}(k_2) + P_{\zeta}(k_1)P_{\zeta}(k_3) + P_{\zeta}(k_2)P_{\zeta}(k_3)\right]\delta^{(3)}\left(\sum_{n=1}^{3}\boldsymbol{k}_n\right)\,. \tag{A.8}$$

This gives the first line of (3.131).

A.2 Two fields

Next, we consider the full definition of $\zeta(\boldsymbol{x})$ and $\chi(\boldsymbol{x})$ with local-type non-Gaussianities considered in (3.129). This updates (A.5) as

$$\begin{aligned}
\zeta_{\boldsymbol{k}} &= \zeta^{\rm G}_{\boldsymbol{k}} + \zeta^{\rm NL}_{\boldsymbol{k}}\,, \quad \zeta^{\rm NL}_{\boldsymbol{k}} \equiv f^{\zeta}_{\rm NL}\int_{\boldsymbol{p}} \zeta^{\rm G}_{\boldsymbol{p}}\zeta^{\rm G}_{\boldsymbol{k}-\boldsymbol{p}} - (2\pi)^3 f^{\zeta}_{\rm NL}\langle \zeta_{\rm G}(\boldsymbol{x})^2\rangle \delta^{(3)}(\boldsymbol{k}) + C^{\zeta}_{\rm NL}\int_{\boldsymbol{p}} \zeta^{\rm G}_{\boldsymbol{p}}\chi^{\rm G}_{\boldsymbol{k}-\boldsymbol{p}} \\
\chi_{\boldsymbol{k}} &= \chi^{\rm G}_{\boldsymbol{k}} + \chi^{\rm NL}_{\boldsymbol{k}}\,, \quad \chi^{\rm NL}_{\boldsymbol{k}} \equiv f^{\chi}_{\rm NL}\int_{\boldsymbol{p}} \chi^{\rm G}_{\boldsymbol{p}}\chi^{\rm G}_{\boldsymbol{k}-\boldsymbol{p}} - (2\pi)^3 f^{\chi}_{\rm NL}\langle \chi_{\rm G}(\boldsymbol{x})^2\rangle \delta^{(3)}(\boldsymbol{k}) + C^{\chi}_{\rm NL}\int_{\boldsymbol{p}} \chi^{\rm G}_{\boldsymbol{p}}\zeta^{\rm G}_{\boldsymbol{k}-\boldsymbol{p}}\,.
\end{aligned} \tag{A.9}$$

We then need to repeat the derivation of $\langle \zeta^{\rm G}_{\boldsymbol{k}_1}\zeta^{\rm G}_{\boldsymbol{k}_2}\zeta^{\rm NL}_{\boldsymbol{k}_3}\rangle$ in (A.7). The presence of $C^{\zeta}_{\rm NL}$ brings an extra term compared to (A.7),

$$\begin{aligned}
\langle \zeta^{\rm G}_{\boldsymbol{k}_1}\zeta^{\rm G}_{\boldsymbol{k}_2}\zeta^{\rm NL}_{\boldsymbol{k}_3}\rangle &\supset C^{\zeta}_{\rm NL}\int_{\boldsymbol{p}} \langle \zeta^{\rm G}_{\boldsymbol{k}_1}\zeta^{\rm G}_{\boldsymbol{k}_2}\zeta^{\rm G}_{\boldsymbol{p}}\chi^{\rm G}_{\boldsymbol{k}_3-\boldsymbol{p}}\rangle \\
&= C^{\zeta}_{\rm NL}\int_{\boldsymbol{p}} \left(\langle \zeta^{\rm G}_{\boldsymbol{k}_1}\zeta^{\rm G}_{\boldsymbol{p}}\rangle\langle \zeta^{\rm G}_{\boldsymbol{k}_2}\chi^{\rm G}_{\boldsymbol{k}_3-\boldsymbol{p}}\rangle + \langle \zeta^{\rm G}_{\boldsymbol{k}_1}\chi^{\rm G}_{\boldsymbol{k}_3-\boldsymbol{p}}\rangle\langle \zeta^{\rm G}_{\boldsymbol{k}_2}\zeta^{\rm G}_{\boldsymbol{p}}\rangle + \langle \zeta^{\rm G}_{\boldsymbol{k}_1}\zeta^{\rm G}_{\boldsymbol{k}_2}\rangle\langle \zeta^{\rm G}_{\boldsymbol{p}}\chi^{\rm G}_{\boldsymbol{k}_3-\boldsymbol{p}}\rangle\right) = 0\,,
\end{aligned} \tag{A.10}$$

since $\langle \zeta^{\rm G}_{\boldsymbol{k}}\chi^{\rm G}_{\boldsymbol{p}}\rangle = 0$ for any $\boldsymbol{k}$ and $\boldsymbol{p}$; recall (3.135) and (3.132). Therefore, (A.8) still holds.

To derive the second line of (3.131), we should start from the correlator $\langle \zeta_{\boldsymbol{k}_1}\zeta_{\boldsymbol{k}_2}\chi_{\boldsymbol{k}_3}\rangle$:

$$\begin{aligned}
\langle \zeta_{\boldsymbol{k}_1}\zeta_{\boldsymbol{k}_2}\chi_{\boldsymbol{k}_3}\rangle &= \langle (\zeta^{\rm G}_{\boldsymbol{k}_1} + \zeta^{\rm NL}_{\boldsymbol{k}_1})(\zeta^{\rm G}_{\boldsymbol{k}_2} + \zeta^{\rm NL}_{\boldsymbol{k}_2})(\chi^{\rm G}_{\boldsymbol{k}_3} + \chi^{\rm NL}_{\boldsymbol{k}_3})\rangle \\
&= \langle \zeta^{\rm G}_{\boldsymbol{k}_1}\zeta^{\rm NL}_{\boldsymbol{k}_2}\chi^{\rm G}_{\boldsymbol{k}_3}\rangle + [\boldsymbol{k}_1 \leftrightarrow \boldsymbol{k}_2] + \langle \zeta^{\rm G}_{\boldsymbol{k}_1}\zeta^{\rm G}_{\boldsymbol{k}_2}\chi^{\rm NL}_{\boldsymbol{k}_3}\rangle + \mathcal{O}\left((\zeta^{\rm NL}_{\boldsymbol{k}})^2, \zeta^{\rm NL}_{\boldsymbol{k}}\chi^{\rm NL}_{\boldsymbol{k}}\right)\,.
\end{aligned} \tag{A.11}$$

First, we show that $\langle \zeta^{\rm G}_{\boldsymbol{k}_1}\zeta^{\rm G}_{\boldsymbol{k}_2}\chi^{\rm NL}_{\boldsymbol{k}_3}\rangle = 0$, hence only the first two terms will contribute. Indeed, using (A.9),

$$\begin{aligned}
\langle \zeta^{\rm G}_{\boldsymbol{k}_1}\zeta^{\rm G}_{\boldsymbol{k}_2}\chi^{\rm NL}_{\boldsymbol{k}_3}\rangle &= f^{\chi}_{\rm NL}\int_{\boldsymbol{p}} \langle \zeta^{\rm G}_{\boldsymbol{k}_1}\zeta^{\rm G}_{\boldsymbol{k}_2}\chi^{\rm G}_{\boldsymbol{p}}\chi^{\rm G}_{\boldsymbol{k}_3-\boldsymbol{p}}\rangle - \langle \zeta^{\rm G}_{\boldsymbol{k}_1}\zeta^{\rm G}_{\boldsymbol{k}_2}\rangle (2\pi)^3 f^{\chi}_{\rm NL}\langle \chi_{\rm G}(\boldsymbol{x})^2\rangle \delta^{(3)}(\boldsymbol{k}_3) \\
&\quad + C^{\chi}_{\rm NL}\int_{\boldsymbol{p}} \langle \zeta^{\rm G}_{\boldsymbol{k}_1}\zeta^{\rm G}_{\boldsymbol{k}_2}\chi^{\rm G}_{\boldsymbol{p}}\zeta^{\rm G}_{\boldsymbol{k}_3-\boldsymbol{p}}\rangle \\
&= f^{\chi}_{\rm NL}\langle \zeta^{\rm G}_{\boldsymbol{k}_1}\zeta^{\rm G}_{\boldsymbol{k}_2}\rangle \left(\int_{\boldsymbol{p}} \langle \chi^{\rm G}_{\boldsymbol{p}}\chi^{\rm G}_{\boldsymbol{k}_3-\boldsymbol{p}}\rangle - (2\pi)^3\langle \chi_{\rm G}(\boldsymbol{x})^2\rangle \delta^{(3)}(\boldsymbol{k}_3)\right) \\
&= (2\pi)^3 f^{\chi}_{\rm NL}\langle \zeta^{\rm G}_{\boldsymbol{k}_1}\zeta^{\rm G}_{\boldsymbol{k}_2}\rangle \delta^{(3)}(\boldsymbol{k}_3)\left(\int_{\boldsymbol{p}} P_{\chi}(p) - \langle \chi_{\rm G}(\boldsymbol{x})^2\rangle\right) = 0\,,
\end{aligned} \tag{A.12}$$

where in the first equality the last term vanishes in a similar fashion to (A.10). Lastly by similar arguments,

$$\langle \zeta^{\rm G}_{\boldsymbol{k}_1}\zeta^{\rm NL}_{\boldsymbol{k}_2}\chi^{\rm G}_{\boldsymbol{k}_3}\rangle = f^{\zeta}_{\rm NL}\int_{\boldsymbol{p}} \langle \zeta^{\rm G}_{\boldsymbol{k}_1}\zeta^{\rm G}_{\boldsymbol{p}}\zeta^{\rm G}_{\boldsymbol{k}_2-\boldsymbol{p}}\chi^{\rm G}_{\boldsymbol{k}_3}\rangle - (2\pi)^3 f^{\zeta}_{\rm NL}\langle \zeta_{\rm G}(\boldsymbol{x})^2\rangle \delta^{(3)}(\boldsymbol{k}_3)\langle \zeta^{\rm G}_{\boldsymbol{k}_1}\chi^{\rm G}_{\boldsymbol{k}_3}\rangle$$

$$+ C^{\zeta}_{\rm NL} \int_{\boldsymbol{p}} \langle \zeta^{\rm G}_{\boldsymbol{k}_1} \zeta^{\rm G}_{\boldsymbol{p}} \chi^{\rm G}_{\boldsymbol{k}_2-\boldsymbol{p}} \chi^{\rm G}_{\boldsymbol{k}_3} \rangle$$

$$= C^{\zeta}_{\rm NL} \int_{\boldsymbol{p}} \langle \zeta^{\rm G}_{\boldsymbol{k}_1} \zeta^{\rm G}_{\boldsymbol{p}} \rangle \langle \chi^{\rm G}_{\boldsymbol{k}_2-\boldsymbol{p}} \chi^{\rm G}_{\boldsymbol{k}_3} \rangle = (2\pi)^3 C^{\zeta}_{\rm NL} P_\zeta(k_1) P_\chi(k_3) \delta^{(3)}\left(\sum_{n=1}^{3} \boldsymbol{k}_n \right) . \tag{A.13}$$

Therefore,

$$\langle \zeta_{\boldsymbol{k}_1} \zeta_{\boldsymbol{k}_2} \chi_{\boldsymbol{k}_3} \rangle = (2\pi)^3 C^{\zeta}_{\rm NL} \left[P_\zeta(k_1) P_\chi(k_3) + P_\zeta(k_2) P_\chi(k_3) \right] \delta^{(3)}\left(\sum_{n=1}^{3} \boldsymbol{k}_n \right) . \tag{A.14}$$

This gives the second line of (3.131). By the symmetry $\zeta \leftrightarrow \chi$, the last two lines of (3.131) also follow.

Appendix B

Redshift conversion

We make use of the following formula to convert from the object's redshift to its Hubble velocity

$$v(z) = c \int_0^z \frac{dz'}{E(z')}, \tag{B.1}$$

where $E(z)$ is defined as

$$E(z) \equiv \sqrt{\Omega_{m0}(1+z)^3 + \Omega_{r0}(1+z)^4 + \Omega_\Lambda}. \tag{B.2}$$

The quantities Ω_{x0} represent the standard density parameters for species x today, which are determined by our choice of cosmology.

To convert velocities between frames (i.e. moving from V in the heliocentric frame to $V_{\rm con}$ in the new (converted) frame), we use

$$V_{\rm con} = V + V_{\rm apex}[\sin(b)\sin(b_{\rm apex}) + \cos(b)\cos(b_{\rm apex})\cos(l - l_{\rm apex})], \tag{B.3}$$

where b and l are the latitude and longitude coordinates of the source. In this expression, to convert the velocities to their values in the CMB frame we take values of $V_{\rm apex} = 371.0$ km/s, $l_{\rm apex} = 264.14$ deg, and $b_{\rm apex} = +48.26$ deg[463].

Appendix C

Testing uniform host-galaxy probability

In order to test our pipeline, we first carried out our GW bias calculations for a uniform host-galaxy probability (implemented by sending δ_l, and δ_h in Eq. (6.14) to infinity), so that the BBH hosts are chosen with equal probability irrespective of galaxy properties. The uniform selection sampling was only implemented for the 2MPZ catalog. Figure C.1 displays the calculated power spectra along with their shot noise for both the siren sample and galaxies within the second redshift bin. Note that if we subtract the shot-noise, the two power spectra statistically agree with each other, which is expected since the sirens catalog in this case is a uniform selection of the galaxy catalog and shares the same angular correlation. The resulting GW bias based on the uniform host-galaxy probability and the galaxy bias is shown in Figure C.2. These plots show that the galaxy bias and the GW bias are statistically the same, which is expected for the uniform host-galaxy probability.

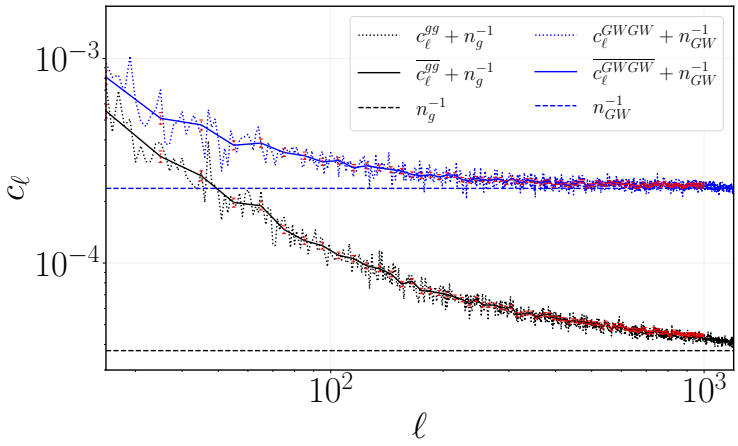


Figure C.1: Comparison of galaxy power spectrum (in black) and the mock siren catalog (in blue) based on uniform host-galaxy probability for 2MPZ within $0.1 < z < 0.2$. The dashed horizontal lines are the shot noise associated with each power spectrum. The red error bars are 1σ error for the power spectrum binned by $\Delta\ell = 10$.

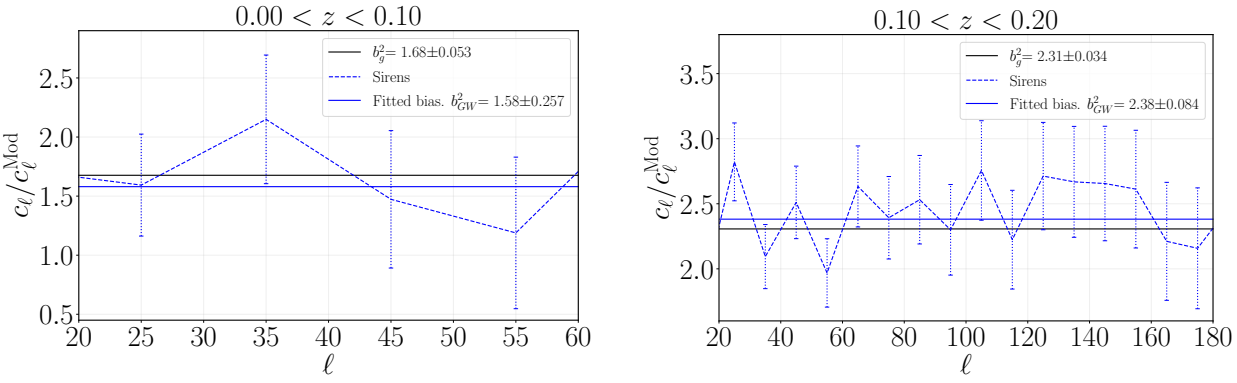


Figure C.2: Bias comparison for galaxy catalog and siren catalog based on uniform host-galaxy probability. The bars are 1σ error.

Appendix D

Determining the best choice for $\ell_{\rm min}$

As pointed out in Section 6.5.1, we find that the constant bias model is a valid approximation in some range for $\ell \in [\ell_{\rm min}, \ell_{\rm max}]$. To determine $\ell_{\rm min}$, we evaluate the bias for 50 different host-galaxy probabilities with varying $M_{\mathcal{K}}$ by fitting a constant bias. We repeat this process by changing $\ell_{\rm min}$ in $\Delta\ell = 10$ steps, as we had earlier computed the mean value of the power spectrum within the same $\Delta\ell = 10$ bins. We start with $\ell_{\rm min} = 20$ because it simplifies our treatment of mask convolution (simply dividing by $f_{\rm sky}$ rather than deconvolving the mask). To assess the goodness of fit, we compare the distribution of χ^2 for the 50 host-galaxy probabilities with different $M_{\mathcal{K}}$ to the expected distribution of χ^2 where the degree of freedom comes from the number of points between $\ell_{\rm min}$ and $\ell_{\rm max}$, minus one for the fitted bias parameter. The results are shown in Figure D.1 and Figure D.2 for 2MPZ and WISC, respectively. In addition, we performed a Kolmogorov–Smirnov test between the two distributions, to verify that the measured χ^2 followed the correct distribution. We required that the p value > 0.05 (i.e., that the measured χ^2 distribution for our fits and theoretical values match), and found $\ell_{\rm min} = 20$ to be a good choice for both redshift bins of 2MPZ, and $\ell_{\rm min} = [30, 50, 60]$ for the three redshift bins in WISC respectively in ascending order. In the case of WISC we found that $\ell_{\rm min} = 20$ resulted in a poor fit for all bins, so we raised $\ell_{\rm min}$ until the $p > 0.05$ criterion was satisfied. Although we did a careful examination of scale cuts here, the results for bias do not change significantly with an alternative choice.

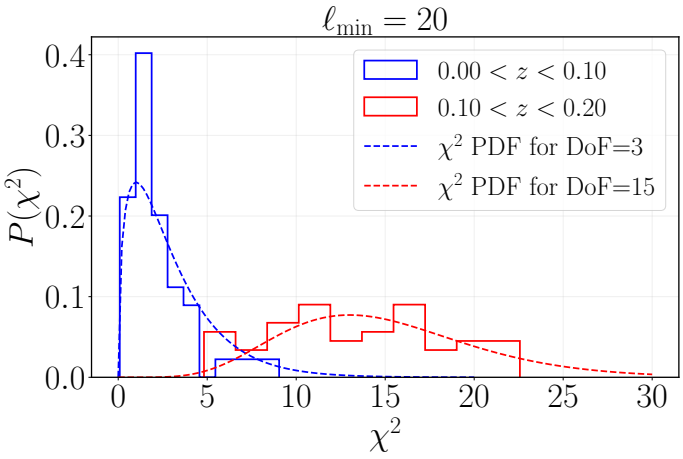


Figure D.1: The distribution of χ^2 for 50 host-galaxy probabilities with different values of $M_{\mathcal{K}}$ for 2MPZ (solid lines) vs. the theoretical distribution of χ^2 with the given number of degrees of freedom (dashed lines), using $\ell_{\rm min} = 20$ as indicated in the title.

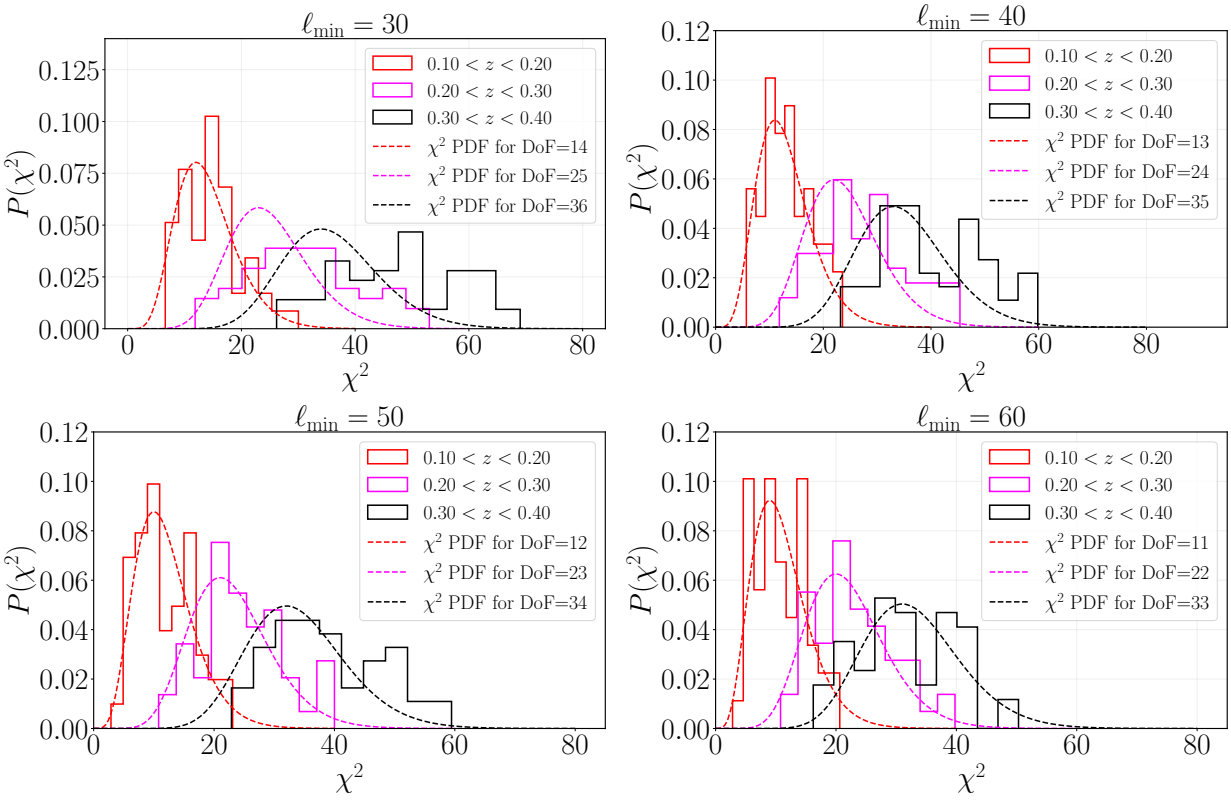


Figure D.2: Like Fig. D.1, but for WISC.

Appendix E

Bias with a different photometric redshift convolution

In Section 6.5.2, we describe the convolution of the photometric redshift distribution with the redshift error. Here, we provide more details on the sensitivity of our results on the convolution scheme. We performed this analysis only for 2MPZ, which should also be more sensitive to changes in the redshift distribution, since the mean redshift is lower. First, as we discussed in Section 6.5.2, we allow the mean of the Gaussian to deviate from the mean of the photometric distribution within the bin through a free parameter, $z_{\rm shift}$, which also depends on $M_{\mathcal{K}}$, δ_l and δ_h. We further check whether the variance of the scatter itself, that is, σ, also varies with $M_{\mathcal{K}}$ (setting $\delta_l = 4$ and $\delta_h = 0.5$) in addition to the dependence on the redshift. The results of our investigation are shown in Figure E.1. The red dashed line corresponds to the whole 2MPZ catalog. The black dashed line represents the fitting formula given in Eq. (6.19). Similarly to the analysis performed for $z_{\rm shift}$, we find the best fit to σ from $z \simeq 0.11$ onward using both a constant and a linear extrapolation, since the actual data (dashed lines) fluctuate. Although we used two different fitting schemes here, the difference in their impact in the final bias calculation turned out to be insignificant. So in what follows, we discuss only the constant extrapolation version in comparison to the other models for σ.

To see the effect of different convolution schemes, we evaluated the bias in six different cases (including the original error distribution properties derived based on entire 2MPZ catalog):

- The original redshift dependent $\sigma(z)$ given in Eq. (6.19) and $z_{\rm shift} = 0$. The results for GW bias are shown by the pink dashed curves in Figure E.2.
- The original result for $\sigma(z)$ with redshift dependence given in Eq. (6.19) but we take

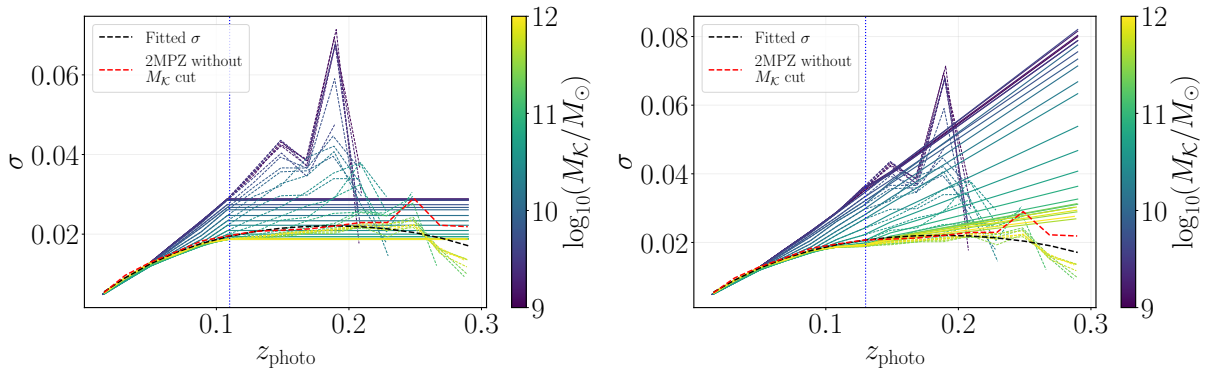


Figure E.1: The spread of spectroscopic redshift distribution around a particular photometric redshift, i.e. σ. This quantity enters Eq. (6.18). The dashed lines represent the actual data. Since after $z \simeq 0.11$, there are not enough hosts (in the siren catalog), the data fluctuates, and we approximate them using (I) either a constant (left panel), (II) or a linear extrapolation (right panel). The red dashed line corresponds to the whole 2MPZ catalog. The black dashed line represents the fitted formula given in Eq. (6.19).

the fitted $M_{\mathcal{K}}$-dependent $z_{\rm shift}$ as displayed in Figure 6.8. The resulting b_{GW} is shown by the black solid curves in Figure E.2.

- A redshift and $M_{\mathcal{K}}$ dependent $\sigma(z, M_{\mathcal{K}})$ as shown in Figure E.1 in addition to $M_{\mathcal{K}}$ dependent $z_{\rm shift}$ as shown in Figure 6.8. The results for b_{GW} are shown by the blue dashed curves in Figure E.2.

- A redshift and $M_{\mathcal{K}}$ dependent $\sigma(z, M_{\mathcal{K}})$ shown in Figure E.1 but setting $z_{\rm shift} = 0$. The results for b_{GW} are shown by the red dashed curves in Figure E.2.

- Setting $\sigma = 0.015$, and $z_{\rm shift} = 0$. The results are shown by the gold dashed curves in Figure E.2

- Setting $\sigma = 0.030$, and $z_{\rm shift} = 0$. The results are shown by the green dashed curves in Figure E.2

Putting all together, these results show the bias calculation is somewhat affected by the details of the redshift convolution scheme, in particular the second redshift bin. In the first redshift bin, $z_{\rm shift}$ is the parameter that has a greater impact on b_{GW} in the low $M_{\mathcal{K}}$ range. In the second redshift bin, σ is more impactful and induces differences from the original error distribution assumption. This is to be expected, since for $z < 0.1$, $\sigma(z, M_{\mathcal{K}})$ agrees well with $\sigma(z)$ used for the galaxy catalogs, whereas for $z > 0.1$, we see much more variation with $M_{\mathcal{K}}$. Given that it is difficult to measure $\sigma(z, M_{\mathcal{K}})$ precisely for $z > 0.1$, we opt to use $\sigma(z)$ fitted to the galaxy catalog, without any variation with $M_{\mathcal{K}}$.

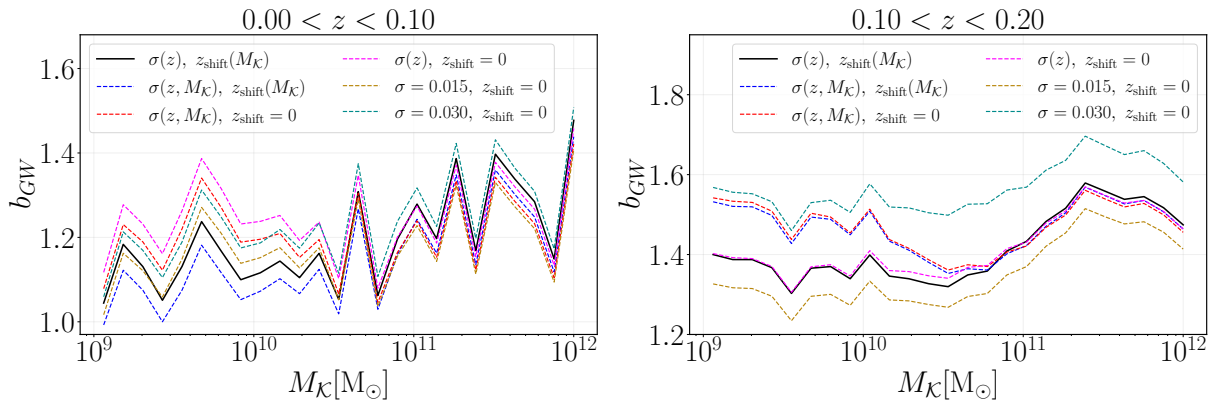


Figure E.2: Bias for 2MPZ catalog with six different photometric redshift convolution schemes. The solid black line represents the result for assuming the original $\sigma(z)$ but considering $z_{\rm shift}(M_{\mathcal{K}})$ used in the analysis of Section 6.6. Refer to the text for an explanation of other cases. The fluctuating behaviour is due to the statistical scatter in the results. We omit the corresponding error bars for the sake of clarity. The error bars can be seen by comparing this plot to the corresponding plot in the main text, i.e. Figure 6.10.

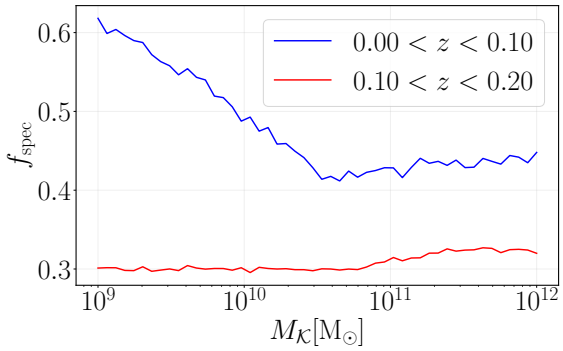


Figure E.3: Fraction of GW hosts with measured spectroscopic redshift, $f_{\rm spec}$, versus $M_{\mathcal{K}}$, for mock siren catalogs in 2MPZ.

Note that the photometric redshift convolution affects only the part of the catalog that does not have spectroscopic redshift information. The fraction of GW hosts that have spectroscopic redshifts in each redshift bin of 2MPZ is shown in Figure E.3.

Appendix F

3D plots of the combined effect of δ_l, δ_h and $M_{\mathcal{K}}$ on GW bias

Here, in addition to Figures 6.13, we offer a 3D visualization of the bias parameter for 2MPZ and WISC, respectively, across various selections of $M_{\mathcal{K}}$, δ_l, and δ_h in Eq. (6.14), covering a large volume of the parameter space. A 3D rotating GIF file is also linked in the captions of Figure F.1, and Figure F.2 to allow the viewer to better see these plots.

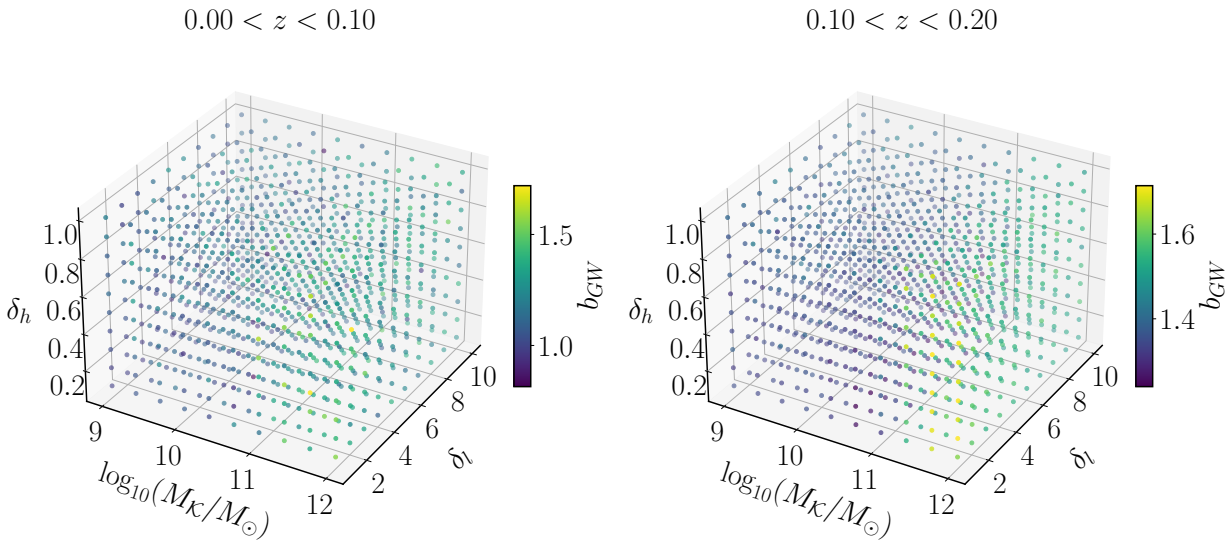


Figure F.1: Bias for varying $M_{\mathcal{K}}$, δ_l, and δ_h in Eq. (6.14) for 2MPZ catalog. The left and right panels are for the first and second redshift bins, respectively, i.e. $0 < z < 0.1$, and $0.1 < z < 0.2$. The rotating 3D GIF files are hyperlinked for $0 < z < 0.1$, and for $0.1 < z < 0.2$.

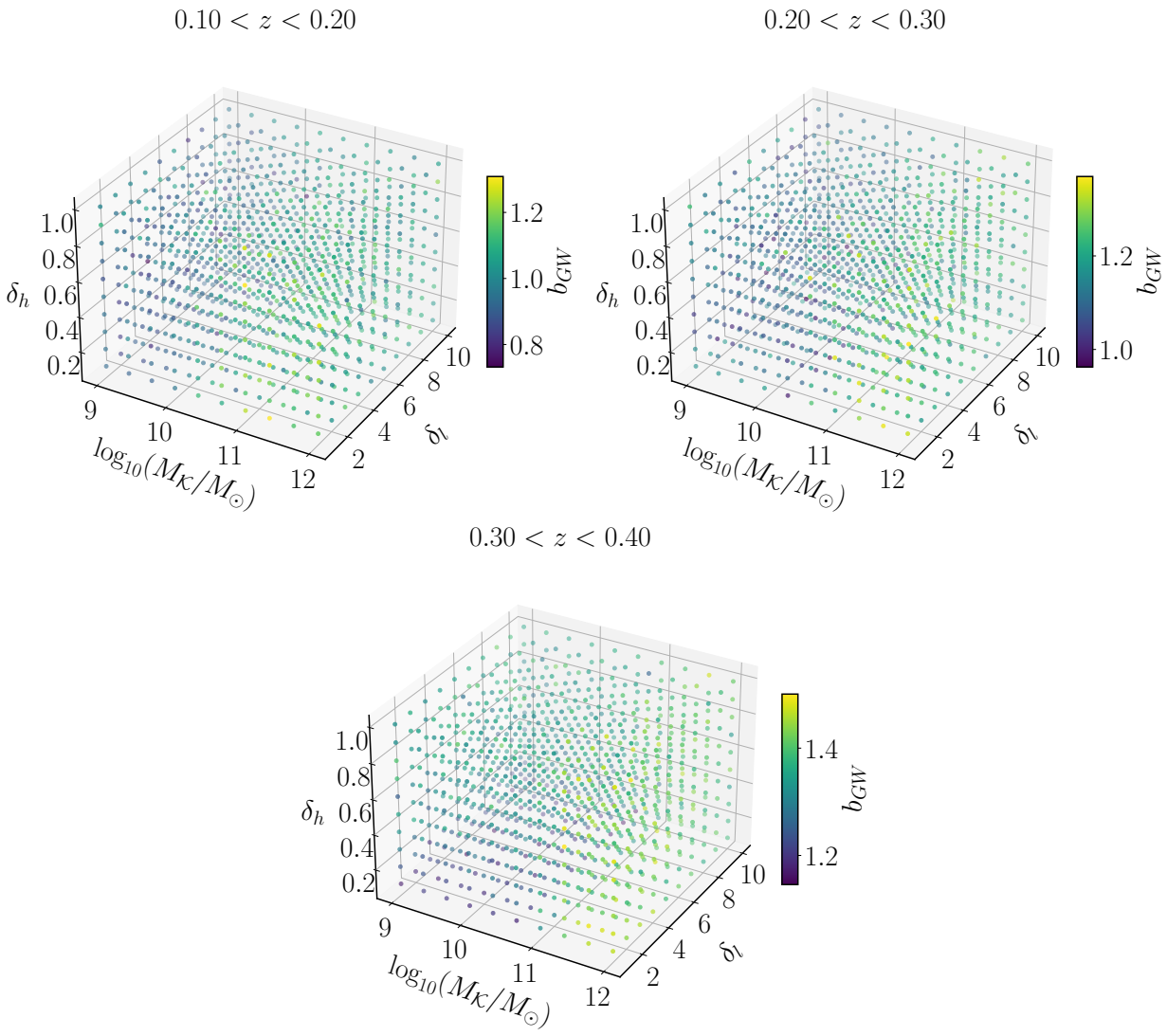


Figure F.2: Bias for varying $M_{\mathcal{K}}$, δ_l, and δ_h in Eq. (6.14) for WISC catalog. The rotating 3D GIF files are hyperlinked for $0.1 < z < 0.2$, for $0.2 < z < 0.3$, and for $0.3 < z < 0.4$.

Appendix G

1D distribution of galaxies as a function of Mass, SFR, and Metallicity

In the main text, we showed the 3D, and 2D galaxy distributions in Figures 7.3, and 7.4. Here, Figure G.1 illustrates the histogram of stellar mass, metallicity, and SFR distribution of galaxies in both redshift bins $0.07 \leq z < 0.135$ and $0.135 \leq z < 0.2$.

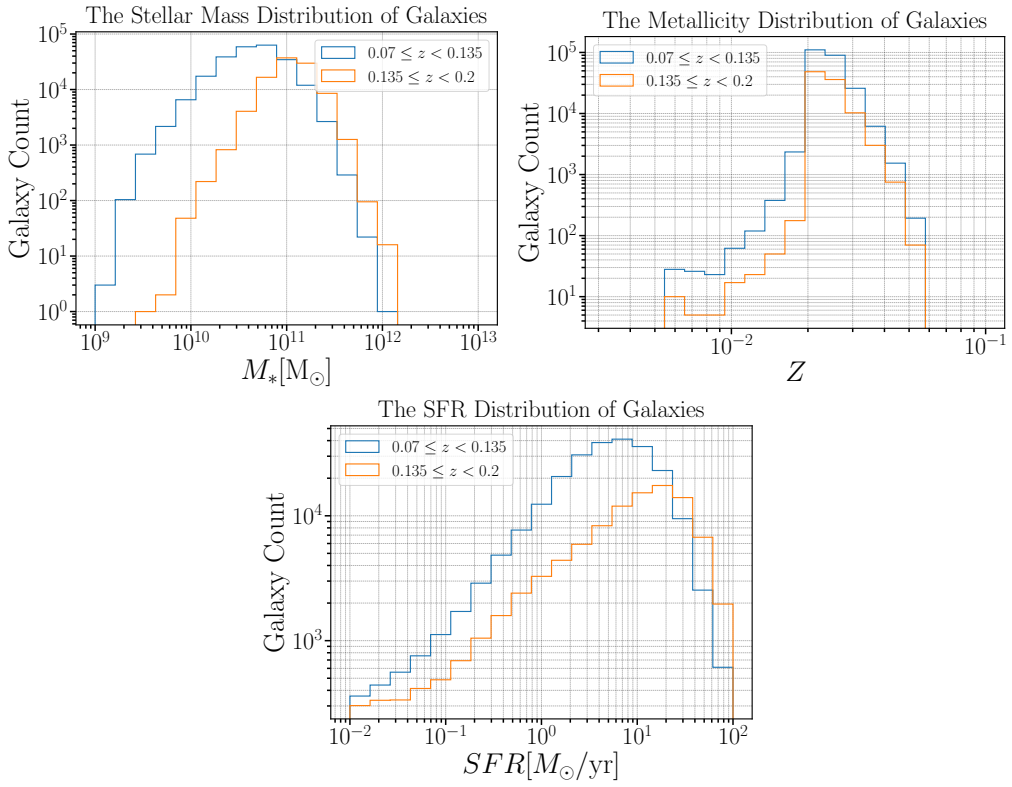


Figure G.1: The distribution of galaxies as a function of stellar mass, SFR and metallicity in two redshift bins $0.07 \leq z < 0.135$ and $0.135 \leq z < 0.2$.

Appendix H

3D power spectrum of the SDSS DR7 galaxy catalog-validation

As a sanity check, we first apply all cuts considered in [381] to reproduce a similar galaxy power spectrum. Here are two more additional cuts assumed in Ross et al. :

- Considering galaxies with $0.07 < z < 0.2$ (in order to sample at lower redshifts in comparison with BOSS)
- Adding some colour cuts which make the sample more homogeneous (as a function of z) and enhance the clustering amplitude, which causes an increase in the mass of the halo hosting the galaxy. For more detail see ([381]).

Considering all the mentioned cuts in Ross et al. and applying them to the galaxy data and random (which are dispersed inside the window with equal surface density and outside the mask) catalogs. Figure H.1 shows our reproduced power spectrum compared to the one calculated in Ross et al. paper and the standard deviation between the two, validating our work.

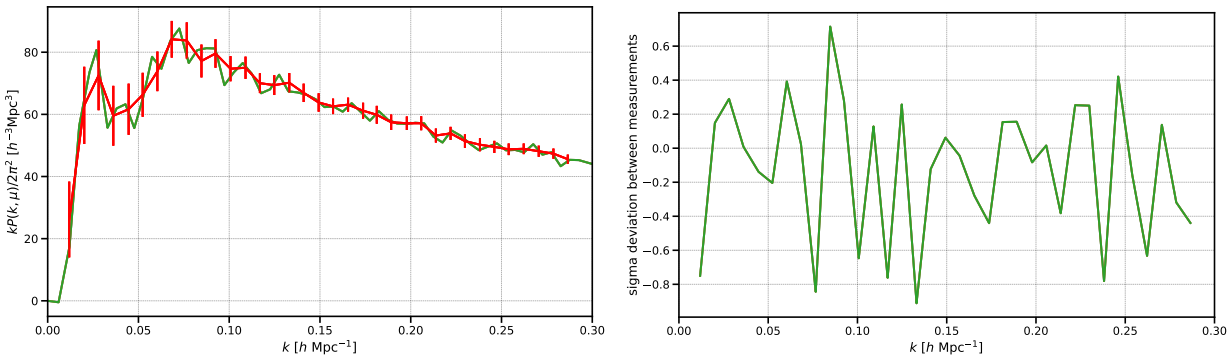


Figure H.1: On the right, we are plotting the deviation of our galaxy power spectrum from the Ross et al measurement, showing that we reproduce their measurement reasonably well. The left panel shows the ratio of the error bars that we calculate to the ones from Ross et al., so that we can argue that our covariance measurement is also accurate.

Appendix I

Testing uniform M_*-Z-SFR host-galaxy probability

As shown in Figure I.1, galaxies and sirens have different number densities in the two redshift bins, so that the FKP weights will be different – leading to different effective redshifts and thus different effective biases – resulting in a few percent offset between galaxy and siren power spectra in the most extreme cases. To provide the cleanest comparison between the two, we always use the galaxy FKP weights for the sirens as well. As shown in Figure I.2 for the case of having different FKP weights (red lines), one can see a 5% offset (particularly between $0.135 \leq z < 0.2$) which will be removed by considering the galaxy FKP weight for sirens as well (green lines). Note that Figure I.2 is in the case of random selection. The random selection was performed as a sanity check, which in this case, we expect the ratio of GW and galaxy biases to fluctuate around 1 since we randomly chose the host galaxies (sirens) out of the galaxy catalog so that sirens' distribution is exactly following the distribution of galaxies. Table I.1 illustrates the number of galaxies and sirens in two redshift bins. The average for the ratio of two power spectra (we call it average bias squared), χ^2, and the standard deviation (std) for two different amplitudes are also shown in Table I.2.

Redshift	Number of galaxies	Number of sirens
$0.07 < z \leq 0.135$	237084	15375
$0.135 < z \leq 0.2$	98734	38282

Table I.1: Number of galaxies and sirens in each redshift bin. This is calculated from Eq (7.12), multiplied by Amplitude=900 or in other words setting $T_{\rm obs} = 900$ years.

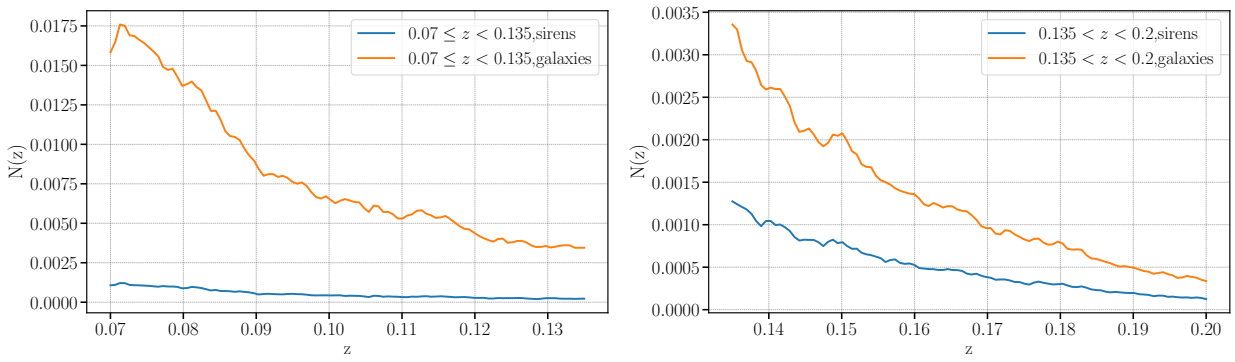


Figure I.1: Number density of galaxies and sirens in first redshift bin $0.07 \leq z < 0.135$ and second redshift bin $0.135 \leq z < 0.2$. Note that, amplitude is considered to be 900 and sirens are randomly selected from galaxy catalog.

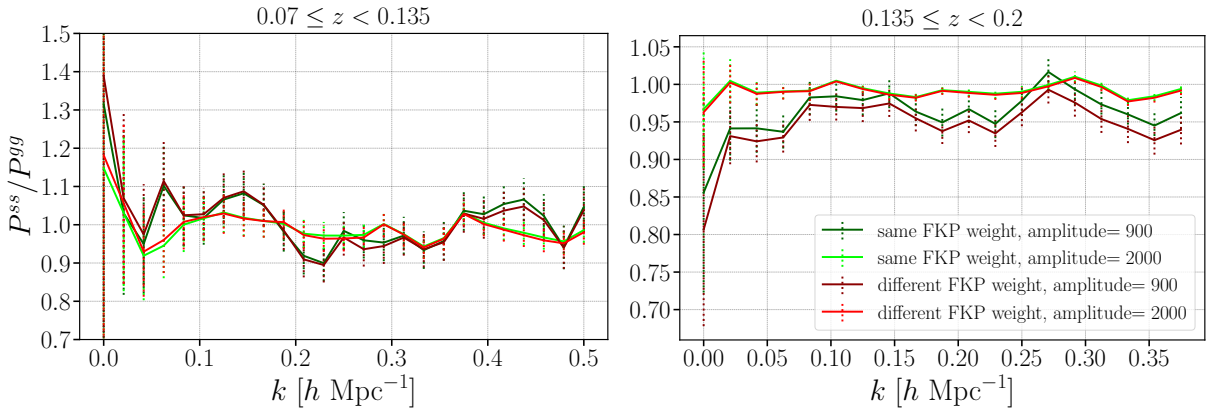


Figure I.2: Ratio of sirens and galaxies power spectra for (I) FKP weights for sirens and galaxies are their corresponding weights (different FKP weight) and (II) both sirens and galaxies weights are set by those of galaxies (same FKP weight). Note that both cases are considered using two different amplitudes of 900 and 2000. P^{ss}, and P^{gg} respectively show the sirens and galaxies power spectrum, while their ratio represents the ratio of their biases.

Redshift	amplitude	Averaged bias squared	χ^2	std
$0.07 < z \leq 0.135$	900	1.01	21.95	0.09
$0.07 < z \leq 0.135$	2000	0.99	17.02	0.07
$0.135 < z \leq 0.2$	900	0.96	56.1	0.03
$0.135 < z \leq 0.2$	2000	0.99	59.7	0.01

Table I.2: Calculated averaged ratio of sirens and galaxies biases, the χ^2, and the standard deviation (std) for two different Amplitude= 900, and Amplitude= 2000 corresponding to Figure I.2.

Appendix J

GW bias vs BH masses

J.1 Modeling BH masses in a binary system

Modeling the Mass distribution of stellar BHs at formation time (source frame): The initial mass distribution of stellar BHs is dictated by the initial mass distribution of the star population (*i.e.* the Initial Mass Function (IMF)). Therefore, approximating that the analytic behaviour of the IMF gets reflected in the initial mass distribution of stellar BHs, only with shifted mass ranges instead, we can implement a similar IMF model with only a few parameters. In particular, we consider that the mass distribution of BHs at formation time mostly follows the same broken power-law behaviour as the stellar population given by Kroupa IMF [420], but in the viable BH mass ranges,

$$P_{Kr}(m) \propto \begin{cases} 0 & m < m_{min}, \\ m^{-\alpha_K} & m \geq m_{min}. \end{cases} \tag{J.1}$$

Here, m refers to the mass of the BH after formation. For stellar populations, α_K varies across different mass ranges, and in higher mass ranges it is estimated to be about 2.3 ± 0.7[420]. However, in the case of BHs, since the minimum mass, m_{min}, corresponds to the upper bound on the mass of neutron stars, we can set $P_{Kr}(m)$ to zero below that value and above that value, the progenitor stars would be in the higher mass range of the IMF. In this work, we assume m_{min} to be greater than $5M_{\odot}$. In principle, $P_{Kr}(m)$ can be redshift dependent as well. In that case, it should be convoluted in the integration for the merger distribution discussed in the next section. In this work we neglect such an effect and assume it is redshift independent.

Furthermore, it is important to consider that stellar mass BHs possess not only a minimum mass threshold, but also an upper limit on their mass at the time of their formation,

which is attributed to the pair-instability supernova (PISN)[340–342]. The prevailing theories of stellar evolution suggest the existence of an upper *mass gap* for stellar BHs, resulting from the excessive production of electron-positron pairs in more massive stars, followed by a subsequent decrease in internal pressure that triggers a partial collapse of the star, leading to an intensified thermonuclear explosion [464]. Ultimately, the star is completely blown apart, leaving no compact remnants in its wake.

The current theoretical estimate for the maximum stellar BH masses ranges between 40 and 50 $M_\odot$ which so far has been mostly consistent with the population of the detected binary BHs. However, theoretically there may be pathways to occasionally get stellar mass BHs in the PISN mass gap. In fact, the most massive stellar BH discovered to date is the primary of GW190521, with one of the BHs having a mass of about $85^{+21}_{-14}M_\odot$ and the other $142^{+28}_{-16}M_\odot$ [465]. A possible explanation for this is that they might have been formed by the mergers of smaller BHs, which is expected to be a rare and subdominant phenomenon.

The start of the mass gap in the stellar mass distribution due to the PISN process is expected to be around $45M_\odot$ with some dependency on the stellar metallicity. The PISN mass scale may also be affected by nuclear reaction rates but since this parameter will not evolve with redshift, it will not be relevant to our analysis [340]. In this work, we model the redshift dependence of the PISN mass[1] as

$$m_{\rm PISN}(Z) = m_{\rm PISN}(\tilde{Z}) - \alpha_Z \log_{10}(Z/\tilde{Z}), \tag{J.2}$$

where Z is the metallicity mass fraction, α_Z is a free parameter and $\tilde{Z} = 10^{-4}$ for $m_{\rm PISN}(\tilde{Z}) = 45{\rm M}_\odot$ [340]. Statistically, galaxies at higher redshifts have lower metallicity. Therefore, implementing Eq. (J.2) leads to larger PISN mass scales for higher redshifts compared to lower redshifts. In other words, when modeling the BH mass distribution in the source frame there is implicit redshift dependence due to the PISN mass gap. We impose this mass gap upper bound in the source frame, using the following mass window function,

$$\mathcal{W}_s(m,z) = \begin{cases} 1 & m_{min} \leq m \leq m_{PISN}(Z), \\ 0 & \text{otherwise.} \end{cases} \tag{J.3}$$

We have set the allowed region of masses in the source frame to be from the minimum stellar BH mass value up to the PISN mass. Finally, we can write the overall mass distributions of the stellar mass BHs in the source frame or at formation redshift, z_f, to be estimated by

$$\mathcal{P}_s(m,z_f) = \mathcal{W}_s(m,z_f) P_{Kr}(m). \tag{J.4}$$

[1] See Refs. [341, 342] for more details.

Modeling mass distribution of the BH at merger time: Since individual compact object need some time (so-called delay-time) to find their companion and form a binary system, the mass distribution of BHs merging at a specific redshift, z_m, includes contributions from BHs which were formed at earlier times ($z > z_m$) (see Figure 1 in Ref. [341]). Therefore incorporating for the delay-time probability, $P_D(t_d)$ (see Eq. (7.9)), we can write the observer frame mass distribution (at the onset of coalescence) as [2]:

$$\begin{aligned}\mathcal{P}_m(m, z_m) &= \mathcal{N} \int_{z_m}^{\infty} P_D(t_d) \frac{dt_f}{dz_f} \mathcal{P}_s(m, z_f) dz_f \\ &= \mathcal{W}(m, z_m) P_{Kr}(m). \end{aligned} \tag{J.5}$$

where $t_d = t_m - t_f$ and $\mathcal{N}$ is a normalization factor so that $\int \mathcal{P}_m(m, z_m) dm = 1$. Note, the factor $\frac{dt_f}{dz_f}$ is due to the fact that for the delay time distribution we had obtained in Eq (7.9), was expressed as a function of time difference between the formation and merger time. In the last line we factorized $\mathcal{P}_{Kr}(m)$ since it is assumed to be redshift independent (given by Eq. (J.1)) and defined $\mathcal{W}(m, z_m)$ as the accumulated window function for allowed BH masses in a binary system at the merging redshift:

$$\mathcal{W}(m, z_m) \equiv \mathcal{N} \int_{z_m}^{\infty} P_D(t_d) \frac{dt_f}{dz_f} \mathcal{W}_s(m, z_f) dz_f. \tag{J.6}$$

Note that, the sources with higher redshift and consequently lower metallicity will have a larger PISN mass scale value. In the case of having metallicity information from observation, $m_{\rm PISN}$ (in Eq. (J.2)), which is a function of the metallicity of galaxies, can be mostly assumed as a constant. This means that there will not be redshift evolution for BH mass distributions of one galaxy. Therefore, Eq.(J.6) can be written as

$$\mathcal{W}(m) = \mathcal{N} \mathcal{W}_s(m) \int_{z_m}^{\infty} P_{t_d}(z_m, z') dz', \tag{J.7}$$

leading to

$$\mathcal{P}(m) = W(m) \mathcal{P}_s(m). \tag{J.8}$$

Modeling the mass distribution of the secondary mass in the Binary: While so far we have modelled the probability mass distribution of individual BHs that are merging at z_m, every BBH consists of two companion masses. Therefore it involves sampling a primary BH (m_1) out of the distribution and a secondary BH (m_2) out of it as well. To do this, We choose the conditional probability for m_2 to include a power-law with a positive index (β_{m_2}), minimum and maximum cuts as $[m_{min}, m_1]$. Note that, although a single m_2 in a binary system is required to be sampled from a positive power-law distribution, the global distribution of m_2 in all binaries, more or less, follows the the negative power-law of m_1.

[2] See Ref. [342] for more details.

J.2 Results for GW bias vs BH masses

Here, we study the relationship between GW bias and BH masses. We follow the prescription described in the previous Section J.1, selecting this set of parameters, $m_{\rm min} = 5M_\odot$, $\alpha_K = 2.3$, $\alpha_Z = 1.5$, and $\beta_{m_2} = 1.0$. . Alongside the two redshift bins ($0.07 < z < 0.135, 0.135 < z < 0.20$), we also bin the sirens based on their BH masses. Here, we chose two mass bins as $4.3 \leq m_{\rm ch} < 6$ $[M_\odot]$, and $6 \leq m_{\rm ch} < 34$ $[M_\odot]$, where $m_{\rm ch}$ is the *chirp mass* given by

$$m_{\rm ch} = \frac{(m_1 m_2)^{3/5}}{(m_1 + m_2)^{1/5}}\,, \tag{J.9}$$

where m_1, and m_2 are the two masses in the binary system drawn from the distributions explained in Section J.1. We choose this BH mass binning to have roughly the same number of sources in each BH mass bin. By doing so, we can roughly find the dependence of GW bias on the masses in the binary system.

Figure J.1 shows the GW bias parameter within two redshift ranges and two BH mass bins. Our results do not show any significant correlation between GW bias and the masses of BHs. This is because the PISN mass scale is not strongly dependent on the host galaxy mass. In the range that the mass of the majority of galaxies in the two catalogs are, i.e. $M_* \in \{10^8, 10^{13}\}M_\odot$, the PISN mass scale does not change significantly. In addition, the PISN mass scale is only a cut-off in the mass distribution of BHs; most of the BHs in the mock catalog have masses below this mass scale since the probability drops rapidly by the power law function given in Equation (J.1). Therefore, the stellar mass of the host galaxy, through our prescription here, has a minor effect on the mass distribution of the majority of the BHs in the mock catalog. In addition, we have not found any strong indication of correlation between GW bias as a function of BH masses and SFR or metallicity of the host galaxies. However, our method is limited to having the current observed metallicity of the galaxy. A model that can include the full evolution of metallicity of the galaxies might show different results.

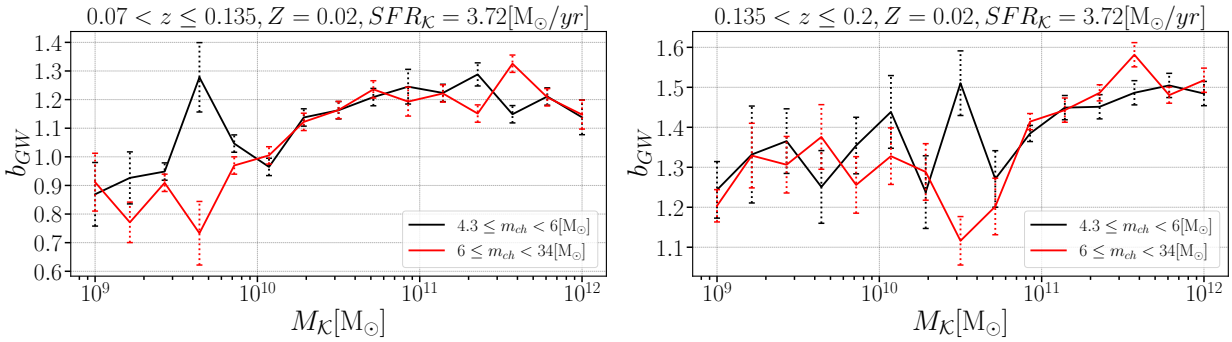


Figure J.1: GW bias parameter for varying $M_{\mathcal{K}}$ using M_*-Z-SFR model and considering different redshift and BH mass ranges. In this case $\delta_l = \infty$, $\delta_h = 0.5$, $\epsilon_l = \infty$, $\epsilon_h = 0.5$, $\zeta_l = 0.5$, $\zeta_h = 0.25$, $Z_{\mathcal{K}} = 0.02$, and $\mathrm{SFR}_{\mathcal{K}} = 3.72$ [$\mathrm{M}_\odot$/yr].