Title
Towards an Optimal Cosmological Detection of Neutrino Mass with Bayesian Inference

Author
Bayer, Adrian Elazar

Publication Date
2023

Peer reviewed|Thesis/dissertation

Towards an Optimal Cosmological Detection of Neutrino Mass with Bayesian Inference

by

Adrian Elazar Bayer

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Physics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Uroš Seljak, Chair
Professor Martin White
Professor Liang Dai

Spring 2023

Towards an Optimal Cosmological Detection of Neutrino Mass with Bayesian Inference

Abstract

Towards an Optimal Cosmological Detection of Neutrino Mass with Bayesian Inference

by

Adrian Elazar Bayer

Doctor of Philosophy in Physics

University of California, Berkeley

Professor Uroš Seljak, Chair

High-precision measurements of large-scale cosmic structure are expected to revolutionize our understanding of fundamental physics, for example by the quantifying neutrino mass and elucidating the nature of dark energy. This dissertation tackles various of the challenges that must be faced in order to optimally extract information from cosmological surveys, taking a particular interest in neutrino mass.

Massive neutrinos suppress the growth of cosmic structure on small scales, where gravity is nonlinear. It is currently an urgent task to determine how to maximally retrieve information in the nonlinear regime, as a traditional power spectrum analysis is no longer optimal. We start by using simulations to investigate the amount of information regarding neutrino mass present in cosmic structure. We find that while there is in principle a lot of information, only a small fraction will be measurable by upcoming surveys such as DESI and LSST. This motivates the need to perform a combined analysis with other cosmological tracers, such as the cosmic microwave background, and galaxy peculiar velocities. We thus develop a Bayesian forward modeling framework to combine field-level inference with galaxy peculiar velocities as a means to optimally extract information.

Additionally, various numerical challenges arise due to the nonlinear effects of gravity. For example, computing the covariance matrices required for likelihood-based analyses becomes challenging, in particular due to the super-sample covariance effect. Moreover, the inference often involves non-trivial posterior surfaces which are plagued by volume effects such as the Look-Elsewhere Effect. This is particularly prevalent when searching for evidence of exotic new models. In the latter chapters of the dissertation we provide methods to solve both of these problems.

All put together, this dissertation provides the ingredients for the cosmology community to move closer to an optimal measurement of neutrino mass and other cosmological parameters.

To the reader,

for making it this far.

Contents

List of Figures

List of Tables

Acknowledgments

First of all, thank you to my family: Mum, Dad, and Hannah. Without them I would never have been able to start a PhD, never mind complete this dissertation. Living so far apart has certainly reinforced the unconditional love of family – I am abundantly grateful for all of their encouragement and support throughout my life. I also have to thank my grandparents: Oma, Opa, Nana, and Baba. I cherish the time I have spent with every one of them, but I have to give a special shout-out to Opa: ever since I was a young child he has always had something new to teach me, whether it be explaining how his camera works, or taking me through his reassembly of an old computer. While his school years were much disrupted by war, he has always maintained a fascination in understanding how everything works – perhaps I would never have set out on the path to a PhD in Physics were it not for his inspiration. Amazingly, it is his 100^{th} birthday on the day of my graduation ceremony – I suppose you could say it was written in the stars. Also, thank you to my Aunts and Uncles, Eva and Nigel, Caroline and Melvyn, and Esther and Bernhard for their love and support, and sometimes even interest in my work!

Now on to the academic acknowledgements. Firstly, a big thank you to my advisor, Uroš. If I had to describe my time working with Uroš in a single story, I would go all the way back to the first time I met him. I was visiting Berkeley after having received a PhD offer, sat in his office with another prospective student. After presenting us with a compelling pitch to come and work with him, he asked us about our thesis projects, and in turn told the other student that their research didn't make sense, suggesting various ways to move things forward. Needless to say, this has been a common phenomenon during my PhD, and I have come to greatly admire and appreciate Uroš's sharp intuition and direct approach. His ability to quickly offer insight on such a vast collection of topics ranging from physics to machine learning is quite amazing and I have benefited from it immensely. Scientific insight aside, I must also thank Uroš for his culinary expertise. There was once a week when the Journal Club Pizza Committee (of which I was a member) decided to order truffle-oil pizza. We later received an email from Uroš telling us not to order it again because the truffles were fake, and that he'd get us real truffles one day so we could taste the difference. Indeed, the following summer, we went to Slovenia and had a delightful time eating pasta with truffle sauce à la Uroš (washing it down with a glass of B-mode). From fake truffles to fake νs, completing my PhD with Uroš has been a terrific journey. Thanks to him I have learnt a lot about science, about being a successful researcher, and about myself – I will certainly carry these learnings with me for the rest of my life!

Next, I must thank Jia Liu for playing a key role in the last few years of my PhD. While locked down at home during the pandemic it became challenging to talk science and spawn new project ideas with my fellow Berkeley researchers. That was until November 2020 when the journal club organizers decided to randomly assign everyone to zoom breakout rooms to eat lunch together after journal club. I found myself in the same room as Jia, we chatted about our research on neutrinos, and have been working together ever since. After becoming a professor at the University of Tokyo, she flew me over to Japan for 3 months as soon as

the COVID restrictions were lifted. I learnt so much from that trip, both from interacting with her and with the wider group at IPMU (especially with Masahiro – a great drinking partner!). From our research collaborations to our karaoke collaborations, and from our zoom talks to our Kyoto walks, I will forever be grateful to Jia for all the advice and experiences she has bestowed upon me, both in terms of scientific research and life as whole!

A special thank you goes to Francisco Villaescusa-Navarro (a.k.a. Paco), a man who's worked on machine learning so much that he's learnt to become a machine, constantly writing papers. I had an extremely enjoyable and enlightening time collaborating with him on one of these papers, and I look forward to doing so more in the future as I move to his neck of the woods for a postdoc. Thank you to him and David for welcoming me into their group!

During my time at the Berkeley Center for Cosmological Physics I have come across many wonderful people. Firstly, I have to thank all the other great scientists I have closely worked with during graduate school: Arka, Chirag, Joe, Marius, Simo, Vanessa, and Yu. I learnt so much from every one of them! Secondly, thank you to all of Uroš's other padawans (in chronological order): Ben, Chirag, Byeonghee, Biwei, Jamie, and Jakob – together they helped ease my navigation through the PhD from start to finish. Also thank you to Chirag for sending me his dissertation template for inspiration [265]. Moreover, I have had the pleasure of discussing science with many amazing professors, including Martin and Liang – I am thankful for their involvement in my dissertation committee and have greatly valued all of my interactions with them. I'm also very grateful for all of the friendships I've formed, or strengthened, along the way – thank you to Vanessa for the after-work drinks and camping trips, to Stephen for being the first person I ever met in Berkeley, and to Boryana for all of our incredible walks and chats. Finally, I'm very appreciative of Joanne and Laura for all the great work that they do, and for always bringing plenty of positive energy!

To end the academic segment, I must thank all of my previous mentors, each of whom helped lead me to this point: Roberto Trotta, Henrique Araujo, Lindley Winslow, and Fay Dowker. Thinking all the way back to high school, thank you to my physics teacher, Dr Finkelstein, to my math teacher, Mr Kerr, and to the biggest physics nerd in the school, Jordana. Also, thank you to my particularly inspirational primary school teacher, Mr Levine.

In terms of personal life at Berkeley, there are too many people and positive moments to acknowledge: from kimchi lunches with Adam and Lizzy, to playing soccer[1] with the PhysCal squad; and from amazing adventures with my housemate, Omer, to continued connection with my fellow Imperial undergraduates, Victoria and Sara. While I can't list everyone, I want to specially thank those who helped me stay sane during the pandemic. Firstly, to Roy and Jade – while many people left Berkeley when the pandemic hit, we remained and enjoyed many wonderful dinners and road trips together (in line with COVID restrictions of course). Secondly, to my friends from undergrad – Yomna, Aidan, Syakira, Pablo, Momo, Anjali, and Kevin – while we spent parts of the pandemic in five different countries, I'm so glad we were brought closer together via our frequent group calls! Last but not least, thank you to Tsvi and Adam for their ongoing solid friendships since high school!

[1] Football.

Chapter 1

Introduction

From Neolithic Europe to Mesopotamia, understanding the mysteries of our Universe has been an ongoing quest dating back many thousands of years to the dawn of human civilization. More recently, over the last few decades, advances in technology have transformed this pursuit into the scientific field of cosmology. The purpose of this dissertation is to continue this journey, paving the way for humanity to continue uncovering the secrets of the Cosmos.

Having measured an abundance of cosmological data over the last few decades, a standard model of cosmology has emerged. This model has successfully explained various observations, ranging from the existence the cosmic microwave background (CMB), thought to have emerged shortly after the Big Bang 13.7 billion years ago, to the large-scale structure (LSS) observed in the distribution of galaxies today. It has also been able to explain the Universe's apparent accelerated expansion, and the abundances of light elements via Big Bang nucleosynthesis. The standard model, often refereed to as the ΛCDM model, shows exquisite agreement with all of these cosmological observations (and more) by fitting for only 6 physical parameters: the age of the Universe, the cold dark matter density, the baryon density, the scalar spectral index, the curvature fluctuation amplitude, and the optical depth of reionization.

According to this model, the Universe began with a Big Bang, in which spacetime expanded exponentially from a singularity for $\sim 10^{-32}$ seconds during an epoch known as inflation. The Universe then consisted of a hot plasma of interacting electrons, photons, and baryons. It then expanded, in turn cooling until eventually reaching a temperature of 3,000 K about 380,000 years after the Big Bang. At this temperature, photons could no longer scatter with electrons, allowing neutral atoms to form from the binding of protons and electrons. At this time, known as recombination, photons decoupled from baryons and electrons, allowing them to free-stream and in turn create the CMB which we observe today. Around this time, matter perturbations began to grow under gravity and evolve to give rise to the large-scale structure of galaxies in the Universe. More recently, around 4 billion years ago, the Universe entered a phase of accelerated expansion, thought to be driven by dark energy, as discovered by observations of supernovae. A visualisation of the Universe's history is shown in Fig. 1.1.

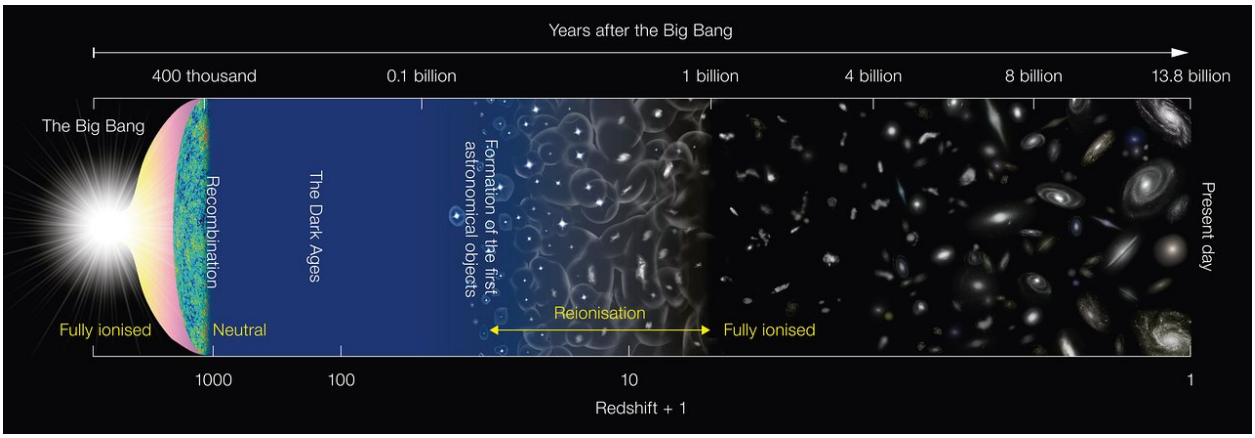

Figure 1.1: Cosmic history according to the ΛCDM model. Credit: NAOJ.

Many cosmological surveys measuring the CMB and LSS have come and gone. The most recent results come from the Stage III surveys such as Planck [292], SDSS BOSS [72], and DES [131], all of which have validated the ΛCDM model. However, while the model has been a great success in fitting the data, a physical understanding of the nature dark energy and dark matter still alludes us. Moreover, the dark energy density is unnaturally (order of magnitude 120) small compared to simple predictions from quantum field theory. There are also some tensions between different measurements of certain parameters. Firstly, the value of the Hubble parameter measured locally (from supernovae) is over 4 σ different to the value inferred by CMB (Planck) and LSS (BOSS) [305]. Secondly, there is a tension in the amplitude of fluctuations inferred from weak lensing compared to CMB experiments. Addressing these issues might involve various exotic modifications to the ΛCDM model.

To uncover the true cosmological model, Stage IV surveys such as DESI[1], PFS[2], Rubin Observatory LSST[3], Euclid[4], SPHEREx[5], SKA[6], Roman Space Telescope[7], and CMB-S4 [3] will provide unprecedented amounts of data spanning various cosmological probes. They will trace the 3d galaxy distribution, weak gravitational lensing, the Lyman-α forest, and the CMB together with CMB lensing. It is thus an imperative task to build the theory and analysis tools required to fully realize the potential of these vast datasets. The focus of this dissertation is to develop some such tools, with a particular interest in extending the ΛCDM model to include massive neutrinos.

[1] https://www.desi.lbl.gov
[2] https://pfs.ipmu.jp/index.html
[3] https://www.lsst.org
[4] https://www.euclid-ec.org
[5] https://www.jpl.nasa.gov/missions/spherex
[6] https://www.skatelescope.org
[7] https://wfirst.gsfc.nasa.gov/index.html

1.1 Cosmological Dynamics and Evolution

During the period of inflation, quantum fluctuations lead to slight fluctuations in the energy density of the Universe. These same fluctuations seeded the formation of structures in the late Universe, such as halos and galaxies. Much of modern cosmology relies on studying these fluctuations to learn about inflation and the subsequent evolution of the Universe.

The standard model of cosmology assumes that the Universe is homogeneous and isotropic. Under these assumptions, the the Friedmann equations follow from the General theory of Relativity, which lay out the relationship between the dynamical evolution of the Universe and its energy content and curvature as follows,

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho - \frac{K}{a^2} \tag{1.1}$$

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}(\rho + 3p), \tag{1.2}$$

$$\tag{1.3}$$

where a is the scale-factor, G is the gravitational constant, ρ is the energy density, p is the pressure, K is the curvature, and the dot denotes a derivative with respect to time $\cdot \equiv d/dt$. Note we assume units where the speed of light is unity.

Defining the Hubble parameter $H(t) = \dot{a}(t)/a(t)$, which quantifies the Universe's expansion rate, one can define the critical energy density as the energy density at which the curvature of the Universe vanishes as follows:

$$\rho_c(a) = \frac{3H^2(a)}{8\pi G}. \tag{1.4}$$

The cosmological parameters are defined as fractions of the energy relative to the critical density, $\Omega(a) \equiv \rho(a)/\rho_c(a)$.

Defining the density contrast δ via

$$\rho(a, \boldsymbol{x}) \equiv \bar{\rho}(a, \boldsymbol{x})[1 + \delta(a, \boldsymbol{x})], \tag{1.5}$$

Poisson's equation describes how density fluctuations are sourced by a peculiar gravitational potential ϕ as follows,

$$\nabla^2\phi(a, \boldsymbol{x}) = \frac{3}{2}\mathcal{H}^2(a)\Omega_m(a)\delta_m(a, \boldsymbol{x}), \tag{1.6}$$

where $\boldsymbol{x}$ is the comoving coordinate (related to the physical coordinate $\boldsymbol{r}$ by $\boldsymbol{r} = a\boldsymbol{x}$), $\Omega_m(a)$ denotes the energy associate with matter, and $\mathcal{H} \equiv aH$ is the conformal Hubble parameter.

Now let us consider the equation of motion of a particle at physical position $\boldsymbol{r}$ in a potential $\boldsymbol{\Phi}$. In the Newtonian limit, i.e. small distances ($r \ll H^{-1}$) and small velocities ($\dot{r} \ll 1$), the equation of motion is given by

$$\ddot{\boldsymbol{r}} = -\nabla_{\boldsymbol{r}}\boldsymbol{\Phi}, \tag{1.7}$$

where $\nabla_{\boldsymbol{r}}$ denotes the derivative with respect to the physical coordinate (unlike ∇ which denotes a derivative with respect to the comoving coordinate). It is instructive to work in comoving coordinates and conformal time, where conformal time τ is related to the cosmic time t by $\tau \equiv a(t)t$. Given the relation $\boldsymbol{r} = a\boldsymbol{x}$, taking a derivative with respect to time yields,

$$\dot{\boldsymbol{r}} = \mathcal{H}\boldsymbol{x} + \boldsymbol{x}', \tag{1.8}$$

where ' denotes a derivative with respect to conformal time. Taking a further derivative yields,

$$\ddot{\boldsymbol{r}} = \frac{1}{a}(\mathcal{H}'\boldsymbol{x} + \mathcal{H}\boldsymbol{x}' + \boldsymbol{x}'') = -\frac{1}{a}\nabla\boldsymbol{\Phi}. \tag{1.9}$$

The term proportional to $\mathcal{H}'x$ accounts for the expansion of spacetime itself, thus it is instructive to bring this term to the right hand side by defining the peculiar potential via

$$\phi = \boldsymbol{\Phi} + \frac{1}{2}\mathcal{H}'x^2. \tag{1.10}$$

This then yields the equation of motion,

$$\boldsymbol{x}'' + \mathcal{H}\boldsymbol{x}' = -\nabla\phi. \tag{1.11}$$

Thus Poisson's equation in terms of ϕ gives

$$-\nabla^2\phi = \nabla\cdot[\boldsymbol{x}'' + \mathcal{H}(\tau)\boldsymbol{x}'] = -\frac{3}{2}\mathcal{H}^2(\tau)\Omega_m(\tau)\delta_m(\tau,\boldsymbol{x}). \tag{1.12}$$

Solving this differential equation yields the dynamics and evolution of the cosmic matter fluid.[8]

1.2 Perturbation Theory

To solve Eqn. 1.12 analytically one can use perturbation theory. There are two main approaches to perturbation theory: Eulerian and Lagrangian. While the Eulerian approach considers the density and velocity fields in a fixed coordinate system, the Lagrangian approach follows the trajectories of individual particles. In this work we will only consider the Lagrangian prescription.

In the Lagrangian picture, we consider the final position of a particle $\boldsymbol{x}$ as related to its initial positions $\boldsymbol{q}$ via the mapping $\boldsymbol{\Psi}(\boldsymbol{q})$ as follows,

$$\boldsymbol{x}(\tau) = \boldsymbol{q} + \boldsymbol{\Psi}(\tau,\boldsymbol{q}). \tag{1.13}$$

The continuity relation gives the density as,

$$\bar{\rho}_m(\tau)d^3\boldsymbol{q} = \rho_m(\tau,\boldsymbol{x})d^3\boldsymbol{x} = \bar{\rho}_m(\tau)[1+\delta_m(\tau,\boldsymbol{x})]d^3\boldsymbol{x}, \tag{1.14}$$

[8]This derivation was inspired by the lectures of Tobias Baldauf at the University of Cambridge.

where bars denote spatial averages. This can be related to the Jacobian J via

$$1 + \delta_m(\tau, \boldsymbol{x}) = \left| \frac{d^3\boldsymbol{q}}{d^3\boldsymbol{x}} \right| = \frac{1}{J(\tau, \boldsymbol{q})}, \tag{1.15}$$

thus, using Eqn. 1.13, we have

$$J(\tau, \boldsymbol{q}) = \det(\delta_{ij} + \boldsymbol{\Psi}_{i,j}(\tau, \boldsymbol{q})), \tag{1.16}$$

where indices after commas represent derivatives with respect to components of $\boldsymbol{q}$.

Eliminating δ_m for J in Eqn. 1.12, and replacing $\boldsymbol{x}$ with $\boldsymbol{q}$ and $\boldsymbol{\Psi}$, gives

$$J(\tau, \boldsymbol{q})[\delta_{ij} + \boldsymbol{\Psi}_{i,j}(\tau, \boldsymbol{q})]^{-1} \left[\boldsymbol{\Psi}''_{i,j}(\tau, \boldsymbol{q}) + \mathcal{H}(\tau)\boldsymbol{\Psi}'_{i,j}(\tau, \boldsymbol{q})\right] = \frac{3}{2}\mathcal{H}^2(\tau)\Omega_m(\tau)[J(\tau, \boldsymbol{q}) - 1]. \tag{1.17}$$

This equation describes the evolution of the displacement of any single matter particle starting at position $\boldsymbol{q}$. It can be solved perturbatively using

$$\boldsymbol{\Psi}(\tau, \boldsymbol{q}) = \boldsymbol{\Psi}^{(1)}(\tau, \boldsymbol{q}) + \boldsymbol{\Psi}^{(2)}(\tau, \boldsymbol{q}) + \dots. \tag{1.18}$$

The first order term $\boldsymbol{\Psi}^{(1)}$ is often referred to as the 1LPT term, and similarly for higher order terms. Once having solved for $\boldsymbol{\Psi}$ one can go on to solve for the particle positions and velocities. A popular approximation based on 1LPT is known as the Zel'dovich approximation (ZA), which extrapolates the linear theory solution into the nonlinear regime [398], but can lead to unphysical phenomena such as Zel'dovich pancakes.

1.3 N-body Simulations

Perturbation theory enables modeling the evolution analytically, but is only accurate on large scales, breaking down as evolution becomes nonlinear. To accurately model these nonlinear scales one must instead employ N-body simulations. In this approach one typically considers a cubic box, with periodic boundary conditions, filled with matter particles which evolve under gravity. Considering only the gravitational interactions of cold dark matter particles significantly outperforms perturbation theory on nonlinear scales, however, to obtain better accuracy on the smallest scales one must also account for neutrinos, baryons, photons, and non-Newtonian-gravitational interactions. In Chapter 2 we will discuss how to include the effects of massive neutrinos.

The calculation performed in an N-body simulation is to solve the Newtonian equations of motion for a set of N particles with masses m_i at positions $\boldsymbol{r}_i$, given by

$$\frac{d^2\boldsymbol{r}_i}{dt^2} = -G\sum_{j\neq i}^{N} m_j \frac{(\boldsymbol{r}_i - \boldsymbol{r}_j)}{|\boldsymbol{r}_i - \boldsymbol{r}_j|^3}. \tag{1.19}$$

To evolve the particles forward in time one typically employs leapfrog integration, which is symplectic and achieves second-order accuracy per evaluation. In terms of evaluating

the gravitational force term on the right hand side, there are various approaches of varying complexity. A brute force approach would require evaluating the gravitational force between all $N(N-1)/2$ pairs of particles, but this is prohibitively expensive. Instead, modern simulations use numerical schemes such as Particle Mesh (PM) [387, 154], and trees methods [49] to estimate the forces efficiently. The PM approach paints particles on a regular grid and then solves Poisson's equation in Fourier space by employing Fast Fourier Transforms (FFTs). FFTs are typically the bottleneck of a PM simulation. On the other hand, tree simulations are more computationally expensive, but produce more accurate results on small scales and late times.

In this work we will employ both PM and PM+Tree approaches. In particular we will consider FastPM [154], which modifies the classical PM approach by enforcing the Zel'dovich approximation at each step to converge in fewer time steps. We will also use the Quijote set of simulations [383] which were run using the PM+Tree code GADGET [346].

In both cases one typically sets the initial conditions of the simulation using LPT to speed up convergence relative to starting with linear theory. More details about setting initial conditions will be given in Chapter 2.

To illustrate the different levels of approximations, Fig. 1.2 shows a 2d projection of the (initial) linear field and the evolved field according to the Zel'dovich approximation, 2LPT, and 20 steps of FastPM.

Using the tools described above one can model the matter field. However, the large-scale structure we observe today consists of halos and galaxies arranged into a vast cosmic web of nodes, filaments, and voids. Modeling the full cosmic web from the initial linear field is one of the major challenges of modern cosmology. We will consider various methods to make the connection between dark matter and halos, including the Friends-of-Friends clustering algorithm, and effective field theory (discussed more in Chapter 5).

1.4 Massive Neutrinos

Of particular interest in this work are constraints on the sum of the neutrino masses $M_\nu \equiv \sum_\nu m_\nu$. The first evidence for neutrino mass came from oscillation experiments [166, 9, 25, 10, 22], which measured the difference in the squares of the masses of the three neutrino mass eigenstates. The best-fit results obtained from a joint analysis of oscillation experiments are $\Delta m^2_{21} \equiv m_2^2 - m_1^2 \simeq 7.55 \times 10^{-5}\text{eV}^2$ from solar neutrinos, and $|\Delta m^2_{31}| \equiv |m_3^2 - m_1^2| \simeq 2.50 \times 10^{-3}\text{eV}^2$ from atmospheric neutrinos [312]. Since atmospheric neutrino experiments only probe the magnitude of the mass difference, there are two possibilities for the neutrino mass hierarchy: $\Delta m^2_{31} > 0$, known as the normal hierarchy, or $\Delta m^2_{31} < 0$, known as the inverted hierarchy. This gives a lower bound on the sum of the neutrino masses of $M_\nu \gtrsim 0.06\text{eV}$ for the normal hierarchy, or $M_\nu \gtrsim 0.1\text{eV}$ for the inverted hierarchy. The current tightest upper bound on the effective electron neutrino mass from particle experiments is obtained by the

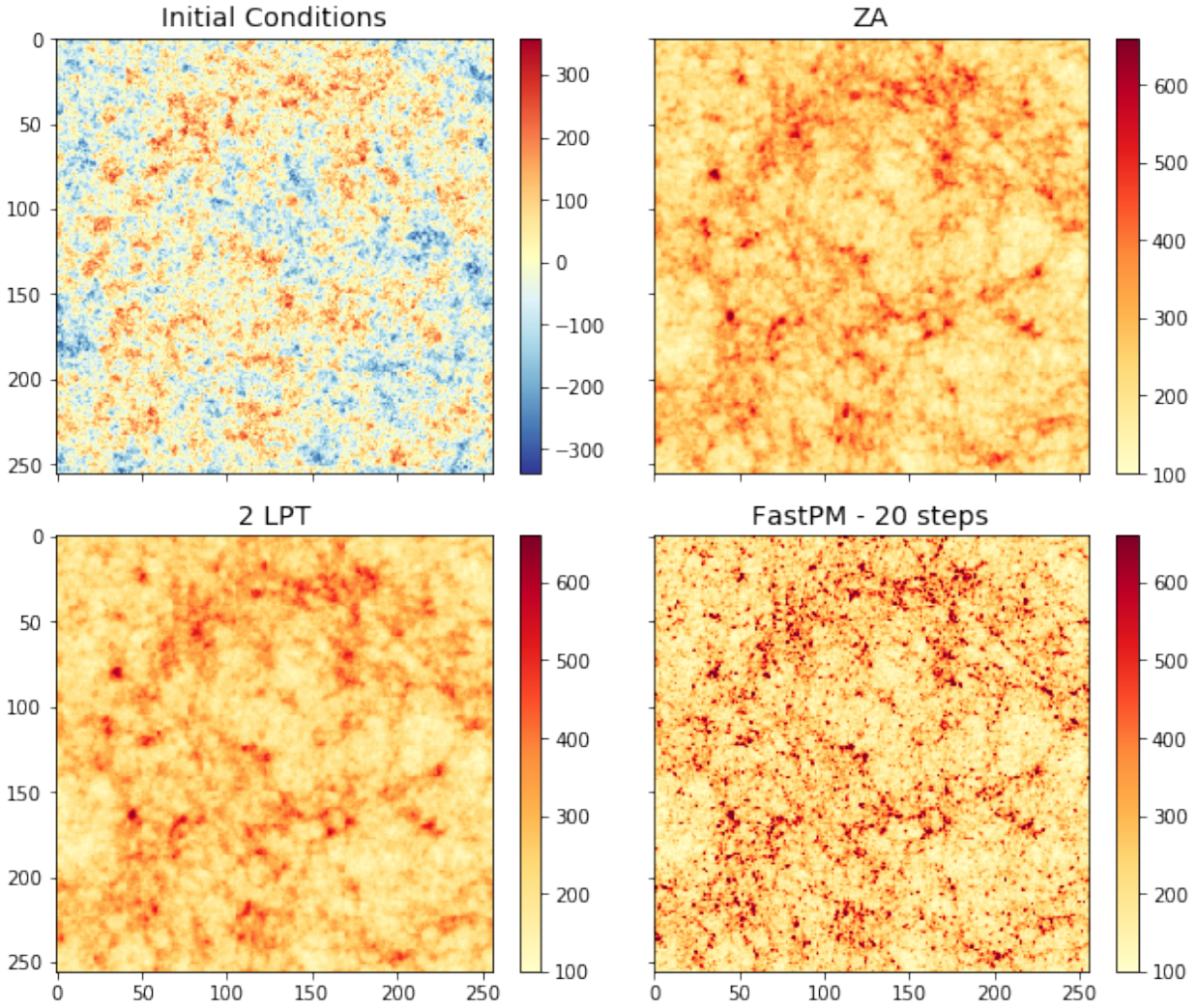

Figure 1.2: Top left: the (initial) linear field at $z = 9$, top right: Zel'dovich approximation at $z = 0$, bottom left: 2LPTat $z = 0$, and bottom right: 20 steps of FastPM at $z = 0$. Credit [265].

KATRIN β-decay experiment, $m_{\nu_e}^{\rm eff} \lesssim 1.1$eV [12] [9].

Neutrinos also play an important role in the Universe's history, as the presence of massive neutrinos both shifts the time of matter-radiation equality and suppresses the growth of structure on small scales. Measuring these effects enables determination of neutrino mass via cosmology, providing a complementary probe to particle physics [145, 199, 148, 237]. While the effects of neutrinos on linear (i.e. relatively large) scales are well understood

[9] Single β-decay experiments do not directly measure the neutrino mass sum, but rather the effective mass of electron neutrinos. In the quasi-degenerate regime where the eigenmasses $m_i > 0.2$ eV ($i = 1, 2, 3$), the three eigenmasses are the same to better than 3%, and hence $m_{\nu_e}^{\rm eff} \approx 1/3 M_\nu$.

theoretically, understanding the effects on nonlinear (i.e. relatively small) scales is an active field of research. There are numerous approaches to obtain theoretical predictions of the nonlinear effects of neutrinos, with varying computational efficiency [see, e.g. 311, 82, 83, 337, 375, 18, 68, 69, 119, 380, 382, 94, 95, 47, 26, 89, 368, 6, 150, 203, 330, 396, 48, 247, 130, 103, 104, 57]. One of the key challenges in simulating massive neutrinos is dealing with the shot noise they produce due to their high thermal velocities; we will present a method to quell this effect in Chapter 2.

For illustration, Fig. 1.3 shows the small-scale suppression in the matter power spectrum due to massive neutrinos according to linear theory (dotted) and a nonlinear approximation known as the halo model (solid). It can be seen that a larger neutrino mass leads to a larger suppression. Also, there is a noticeable difference between linear and nonlinear theory; this motivates the importance of modeling the nonlinear effects of massive neutrinos. Crucially, an N-body simulation will produce a more accurate result of nonlinear theory compared to the halo-model approximation, as will be shown in Chapter 2.

The current best constraints on M_ν arise by considering the CMB and combining it with other cosmological probes. Assuming a Λ cold dark matter (ΛCDM) cosmological model, the upper bound on the neutrino mass from the Planck 2018 CMB temperature and polarization data is $M_\nu < 0.26$eV (95% CL) [292]. When combined with baryonic acoustic oscillations (BAOs) a more stringent bound of $M_\nu < 0.13$eV (95% CL) is obtained. Further combining with CMB lensing gives $M_\nu < 0.12$eV (95% CL).

A major limiting factor of current cosmological constraints is that CMB experiments measure the combined quantity $A_s e^{-2\tau}$, where A_s is the amplitude of scalar perturbations and τ is the optical depth of reionization. Hence, accurate determination of τ is imperative to obtaining tight constraints when combining CMB with clustering/lensing [19, 244, 28, 395, 85]. Most upcoming ground-based CMB experiments, such as Simons Observatory and CMB-S4, will not observe scales larger than $\ell \sim 30$, and will therefore be unable to directly constrain τ [3]. Planck currently provides the best constraint of $\tau = 0.054 \pm 0.007$, with large improvements expected from the ongoing CLASS experiment [385] and the upcoming LiteBIRD [189] space mission. Furthermore, future radio 21cm and, e.g., near-infrared/optical galaxy observations will provide new information on the optical depth which would also help improve the constraints form the CMB [244, 85].

Before significant progress will be made in measuring τ, improved measurements of M_ν are expected from galaxy surveys such as DESI, LSST, and Euclid. These surveys will measure fluctuations on nonlinear scales with unprecedented precision. There is thus much motivation to explore other probes of neutrino mass, beyond the traditional 2-point clustering.[10]

[10]Much of the content of this subsection was taken verbatim from *Bayer A.E., et al.* (arXiv:2102.05049) ApJ 919 1 24 (2021) [62].

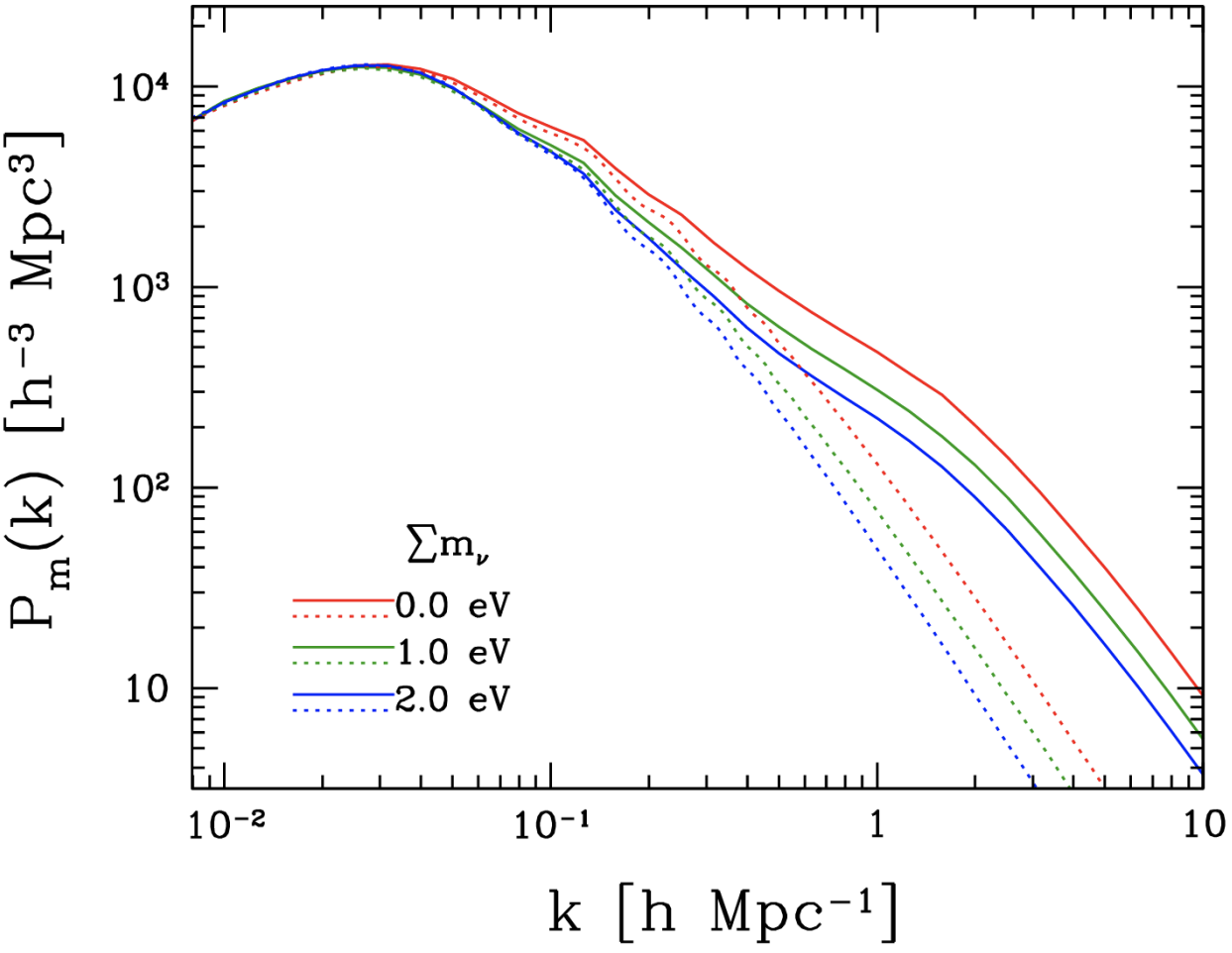

Figure 1.3: Suppression in the matter power spectrum due to massive neutrinos according to linear theory (dotted) and the halo-model approximation to nonlinear theory (solid). Credit [187].

1.5 Dissertation Outline

Having introduced the concepts of N-body simulations and the effects of neutrino mass on cosmology, we begin the dissertation by modeling massive neutrinos by incorporating them into a particle-mesh simulation, FastPM, in Chapter 2. By applying an innovative method we are able to bypass the problem of shot noise that typically plagues neutrino simulations, enabling modeling the effects of massive neutrinos at a more reasonable computational cost than traditional methods.

To fully realize the potential of measurements of LSS, we then use simulations to investigate the amount of information regarding neutrino mass in the 3d matter field, δ_m, in Chapter 3. We motivate how analysing statistics beyond the power spectrum contain a wealth of complementary information on nonlinear scales. In particular we consider the

power spectrum, halo mass function, and void size function. We find that these three statistics show very different degeneracies in the M_ν–σ_8 plane, meaning that tight constraints are achieved when combining all three. For the particular setup considered, we find that one can get constraints of 0.018eV using scales up to $k_{\rm max} = 0.5\,h/{\rm Mpc}$.

However, the 3d matter field is not observable in modern cosmological surveys. In Chapter 4 we then explore how much information regarding neutrino mass there will be in the observable galaxy and weak lensing fields. Additionally, instead of looking at specific types of higher-order statistics, we consider the total information in the field by looking at the complex phases. By fixing the linear physics between two N-body simulations, we find that there is actually very little information regarding neutrino mass in the galaxy and weak lensing fields up to scales of $k = 1\,h/{\rm Mpc}$. We do however find that there is some signal from velocities, which trace the matter field via $\delta_v \sim f\delta_m$, where f is the growth rate. This motivates using peculiar velocity information, together with accurate modeling of redshift-space distortions, to optimize constraints on neutrino mass. Moreover, combining large-scale structure with cosmic microwave background measurements can help break degeneracies, in turn unlocking further information.

To move towards obtaining optimal constraints, in Chapter 5 we then employ Bayesian forward modeling with field-level inference to extract all the information in cosmic fields. We develop a framework to perform a combined analysis of galaxy clustering information with galaxy peculiar velocity information to reconstruct the initial conditions of the Universe. We achieve this using differentiable forward modeling.

In Chapter 6 we then focus in on a particular challenge in doing inference for cosmological analyses, the super-sample covariance effect. The is a nonlinear effect on the covariance matrix which arises due to neglect of modes that are larger than the size of the simulation. The effect of these missing large modes can cause a significant difference in the covariance matrix on small, nonlinear, scales. We create a set of FastPM simulations to quantify this effect for the power spectrum, bispectrum, halo mass function, and void size function. These simulations can also be used to quantify the effect for other statistics in the future.

Finally, performing Bayesian inference in high dimensions, or for nonlinear models, can result in non-trivial volume effects which make it difficult to reliably perform inference. One such problem is known as the Look-Elsewhere Effect which occurs when scanning a large parameter space for evidence of a new model or a signal. In such cases it becomes difficult to quantify the statistical significance of a discovery, moreover, the posterior might be multimodal making it difficult to determine which peak is the most significant. In Chapter 7 we analyse the Look-Elsewhere Effect from a frequentist and Bayesian perspective, ultimately marrying the two to produce a unified method that quantifies the effect orders of magnitude faster than traditional methods. Then in Chapter 8 we present an even faster method to *self-calibrate* the effect by simply computing the significance from the heights and ranks of peaks in the likelihood. This approach has negligible computational cost as peaks in the likelihood are a byproduct of every peak-search analysis.

We then summarize and conclude in Chapter 9.

Chapter 2

A fast particle-mesh simulation of nonlinear cosmological structure formation with massive neutrinos

The contents of this chapter was originally published in [57],

> A fast particle-mesh simulation of non-linear cosmological structure formation with massive neutrinos
> *Bayer A.E., Banerjee A., Feng Y.* (arXiv:2007.13394) JCAP 01 (2021) 016

Quasi-N-body simulations, such as FastPM, provide a fast way to simulate cosmological structure formation, but have yet to adequately include the effects of massive neutrinos. In this chapter, we present a method to include neutrino particles in FastPM, enabling computation of the CDM and total matter power spectra to percent-level accuracy in the non-linear regime. The CDM-neutrino cross-power can also be computed at a sufficient accuracy to constrain cosmological observables. To avoid the shot noise that typically plagues neutrino particle simulations, we employ a quasi-random algorithm to sample the relevant Fermi-Dirac distribution when setting the initial neutrino thermal velocities. We additionally develop an effective distribution function to describe a set of non-degenerate neutrinos as a single particle to speed up non-degenerate simulations. The simulation is accurate for the full range of physical interest, $M_\nu \lesssim 0.6$eV, and applicable to redshifts $z \lesssim 2$. Such accuracy can be achieved by initializing particles with the two-fluid approximation transfer functions (using the REPS package). Convergence can be reached in ~ 25 steps, with a starting redshift of $z = 99$. Probing progressively smaller scales only requires an increase in the number of CDM particles being simulated, while the number of neutrino particles can remain fixed at a value less than or similar to the number of CDM particles. In turn, the percentage increase in runtime-per-step due to neutrino particles is between $\sim 5 - 20\%$ for runs with 1024^3 CDM particles, and decreases as the number of CDM particles is increased. The code has been made publicly available, providing an invaluable resource to produce fast predictions for cosmological surveys and studying reconstruction.

2.1 Introduction

Understanding the nature and generation mechanism of neutrino mass is a challenge that unites particle physics with cosmology. Neutrino oscillation experiments were the first to provide evidence for neutrino mass [166, 9, 25, 10, 22] by measuring the difference in the squares of the masses of the three neutrino mass eigenstates. The best fit results obtained from a joint analysis of oscillation experiments are $\Delta m^2_{21} \equiv m_2^2 - m_1^2 \simeq 7.55 \times 10^{-5}\text{eV}^2$ from solar neutrinos, and $|\Delta m^2_{31}| \equiv |m_3^2 - m_1^2| \simeq 2.50 \times 10^{-3}\text{eV}^2$ from atmospheric neutrinos [312]. Because atmospheric neutrino experiments are only sensitive to the magnitude of the mass difference, there are two possibilities for the neutrino mass hierarchy: $\Delta m^2_{31} > 0$, known as the normal hierarchy, or $\Delta m^2_{31} < 0$, known as the inverted hierarchy. This leads to a lower bound on the sum of the neutrino masses, $M_\nu \equiv \sum_\nu m_\nu$, of $M_\nu \gtrsim 57\text{meV}$ for the normal hierarchy, or $M_\nu \gtrsim 96\text{meV}$ for the inverted hierarchy. An upper bound on neutrino mass, given by β-decay experiments, is $M_\nu \lesssim 1.1\text{eV}$ [12]. While current particle physics experiments provide bounds, they are unable to determine either M_ν, or the absolute mass scale of each eigenstate.

By virtue of the high number density of neutrinos in the universe, cosmology provides a complementary probe to particle physics when studying various properties of neutrinos. Numerous cosmological observables can be used to study neutrino mass, with one example being the cosmic microwave background (CMB) [292, 111, 176, 212, 299, 333], including secondary effects such as the thermal Sunyaev-Zeldovich (tSZ) effect [193, 252] and kinetic Sunyaev-Zeldovich (kSZ) effect [269, 309]. Another example is large-scale structure, which includes galaxy lensing, cosmic shear, and baryon acoustic oscillations (BAO) [5, 363, 66]. A further example is the Lyman-alpha forest [328, 279, 278, 394]. Assuming a ΛCDM cosmological model, the upper bound on the neutrino mass from the Planck 2018 CMB temperature and polarization data is $M_\nu < 0.26\text{eV}$ (95% CL) [292]. Combining with BAO gives a more stringent bound of $M_\nu < 0.13\text{eV}$ (95% CL), and further adding lensing gives $M_\nu < 0.12\text{eV}$ (95% CL). Allowing more flexibility in the cosmological model, such as letting the spectral index run and considering a varying dark energy equation of state, can increase this upper bound to $M_\nu < 0.52\text{eV}$ (95% CL) [369]. In all cases, the current upper bound on neutrino mass from cosmology is stronger than the bound from particle physics, and the cosmological bound is expected to improve with upcoming surveys.

This work focuses on the effects of neutrinos on cosmological structure formation [145, 199, 148, 237]. Upcoming galaxy surveys, such as DESI [113, 161], LSST [114], Euclid [231, 21], eBOSS [136], WFIRST [345], and SKA [170, 379, 402], are predicted to give precise measurements of neutrino mass. For example, DESI and LSST forecast constraints of order $\sim 0.02\text{eV}$, thus the minimal neutrino mass should be detectable at the $\sim 3\sigma$ level. Similar levels of accuracy are expected from CMB experiments when combined with BAO measurements from DESI.

Neutrinos affect structure formation because their low masses cause them to behave as relativistic particles in the early universe, gradually becoming non-relativistic as the universe expands. This means that neutrinos possess high thermal velocities during the epoch of

structure formation in the late universe, distinguishing them from the relatively slow cold dark matter (CDM). As a result, massive neutrinos do not cluster on small scales, leading to a relative suppression in the growth of matter perturbations compared to cosmologies with massless neutrinos. A useful way of quantifying the extent of small-scale suppression is to consider the ratio of the matter power spectrum between a cosmology with massive neutrinos, P_m, and a cosmology with massless neutrinos, $P_m^{M_\nu=0}$. To linear order, this is given by

$$\frac{P_m}{P_m^{M_\nu=0}} \approx 1 - 8f_\nu, \tag{2.1}$$

where $f_\nu \equiv \Omega_{\nu,0}/\Omega_{m,0} \lesssim 0.05$ is the ratio of neutrino to total matter density at $z = 0$ [199]. Hence, the relative suppression is proportional to the total neutrino energy density, which itself depends on M_ν as follows

$$\Omega_{\nu,0} = \frac{M_\nu}{93.14h^2\text{eV}}. \tag{2.2}$$

Thus the matter power spectrum is sensitive to the sum of the neutrino masses. Additionally, the profile of the matter power spectrum is, in principle, sensitive to the individual mass of each eigenstate [236]; however, measuring the individual masses may not be achievable in the foreseeable future [27].

While the effects of neutrinos on linear (i.e. relatively large) scales are well understood theoretically, understanding the effects on non-linear (i.e. relatively small) scales is an active field of research. For a fixed volume, there are many more independent modes on small scales than on large scales. Consequently, theoretical understanding of small scales would greatly increase the information that can be extracted from experimental surveys, in turn increasing the precision of their results. There is thus much motivation for simulations capable of modeling the effects of massive neutrinos on small scales.

In recent years, many techniques have been developed to model structure formation with massive neutrinos, and they can mostly be separated into two methodologies. The first is to use a fluid description for the neutrinos, and coupling this to the non-linear CDM gravitational potential [82, 26, 368, 18, 337, 203, 330, 311, 130]. Most of these approaches use linear theory, or perturbative approaches, to close the Boltzmann hierarchy in the absence of a known equation of state. These methods do not capture the full non-linear evolution of the neutrino field, leading to a reduction in accuracy at late times and on small scales, especially when the neutrino masses are much larger than the minimum total mass ($M_\nu \sim 57$meV). The second methodology is to include neutrinos as an extra set of particles in the simulation [375, 68, 83, 380, 119, 94, 95, 89, 396, 150, 382, 6, 383]. This approach is fully non-linear, unlike the fluid approximation above. Here the neutrino velocities are typically assigned by randomly sampling from the Fermi-Dirac distribution. This, however, can be problematic because the large neutrino thermal velocities cause a significant proportion of neutrino particles to traverse the simulation box multiple times, in turn erasing clustering on small scales and leading to shot noise. The amount of shot noise in the neutrino power spectrum is inversely proportional to the number of neutrino particles, thus a simple approach to avoid this problem is by using a large number of particles. This is the approach taken in [150], but is

extremely computationally expensive. A substantially faster approach is to use fluid-particle hybrid methods. One example of a hybrid method is to use tracer particles to estimate the higher order moments of the Boltzmann hierarchy, requiring fewer particles [47]. However, this method is still relatively expensive because it requires both neutrino particles and a non-linear neutrino fluid on a grid, which itself requires hydrodynamic techniques. A more efficient hybrid method [69] treats fast neutrinos using a linear fluid approximation, while slow neutrinos are treated as particles after some user-defined redshift threshold. This minimizes computational cost, while still ensuring that the power induced by neutrino particle clustering is larger than the shot noise.

Recent work [48] provides a superior means to evade shot noise at all redshifts, by sampling the Fermi-Dirac distribution in a low entropy, quasi-random manner[1]. This has been shown to reduce the shot noise by a factor of $> 10^7$, enabling more accurate study of all scales. It is this method that we will focus on in this work. A follow up study [84], in the context of hybrid simulations, showed that this method can induce spurious correlations between neutrinos and CDM on small scales. However, we note that this study did modify the methodology of [48] in a manner that could have exaggerated their results, by using a different sampling scheme for the Fermi-Dirac distribution and by initializing neutrino particles at a late redshift of $z = 4$ (as typical for hybrid methods). We will show that such spurious effects can be avoided with simple considerations in our particle-only implementation.

While one could study neutrinos using a full N-body or hydrodynamic simulation, there is increased interest in quasi-N-body methods, such as FastPM [154] and COLA [356], as they produce significantly faster simulations of structure formation. Some attempts have been made to include neutrinos in COLA by using fitting formulae for the growth factors [392], but this approach does not provide the required accuracy in the non-linear regime for upcoming surveys. We focus on FastPM, which implements a particle-mesh (PM) approach and enforces the correct linear evolution by using modified kick and drift factors. Moreover, while quasi-N-body methods often fail at very small scales, typically $\lesssim 1\mathrm{Mpc}/h$, this has recently been addressed within FastPM by traversing the gravitational potential using gradient descent techniques to increase the small scale force resolution [128]. The purpose of this work is to add neutrino particles to FastPM by employing the methods of [48] to enable study of the non-linear regime. This will allow the inclusion of neutrino mass as one of the cosmological parameters in forward models, providing a fast way to interpret galaxy survey data. Furthermore, after applying FastPM's inbuilt halo-finder, one can reconstruct the initial conditions of the universe[2] from galaxy positions and luminosities using the methods of [329, 153, 266, 196]. Because massive neutrinos modify the information content of the CDM and total matter fields, this work will enable the study of massive neutrinos in the context of reconstruction.

[1] We use the term quasi-random to distinguish from pseudo-random, see e.g. [343] for definitions of the two. Using a quasi-random sampling scheme ensures that the entropy of the underlying physical system is not increased by sampling.

[2] One example is BAO reconstruction, which was first considered in [149] and applied in [276].

It is worth mentioning another recently explored method to study massive neutrinos that uses neither the fluid approximation nor neutrino particles. This approach seeks to add the effects of massive neutrinos to the results of simulations without massive neutrinos, either by employing cosmological-rescaling algorithms [400], convolutional neural networks [168], or by perturbing the particles' final positions using a carefully designed gauge transformation [283]. We note that these methods ultimately rely on having full neutrino simulations available to compare against and tweak their input parameters, whereas our approach explicitly includes neutrinos using physical parameters as input.

The organization of this paper is as follows. In §2.2 we outline the methodology employed to simulate massive neutrinos. We discuss the Fermi-Dirac sampling scheme, the setting of initial perturbations, and comment on changes made to FastPM's evolution algorithm. We then present the results of our simulation in §2.3, comparing with full N-body simulations such as Quijote [383]. Comments on the runtime are given in §2.4. Finally, we conclude in §2.5, outlining ideas for future work and applications to surveys.

2.2 Methodology

This section outlines the approach used to include massive neutrinos within FastPM[3]. We refer the reader to [154] for a comprehensive review of FastPM.

Initializing massive neutrino particles

To model the effects of massive neutrinos we include an additional species of particle in the simulation. To set the initial thermal velocities of the neutrinos one must sample the Fermi-Dirac distribution. The sampling is usually performed randomly, leading to shot noise dominating small scales. In this work we develop the methods of [48], which have been shown to reduce this shot noise by a factor of more than 10^7 by sampling the Fermi-Dirac distribution in a quasi-random manner. This enables accurate study of small scales and late times. We illustrate the initial configuration of particles in Figure 2.1, and will describe the features of this setup throughout the remainder of this subsection.

A total of N_n neutrino particles are initialized on a grid containing N_{sites} uniformly spaced sites. Each site comprises of N_n/N_{sites} particles, with each particle having a different initial thermal velocity. Note this is different to CDM, which has N_c grid sites and a single particle at each site. The neutrino grid used in our applications is coarser than the CDM grid, and we will show later that one can achieve accurate results using far fewer neutrino particles than CDM particles.

In dimensionless units, the Fermi-Dirac distribution function is given by

$$f(q) = \frac{1}{e^q + 1}, \tag{2.3}$$

[3]The code can be found at https://github.com/fastpm/fastpm. In the code, massive neutrinos are labelled as NCDM (not-cold dark matter), following the CLASS [73] convention.

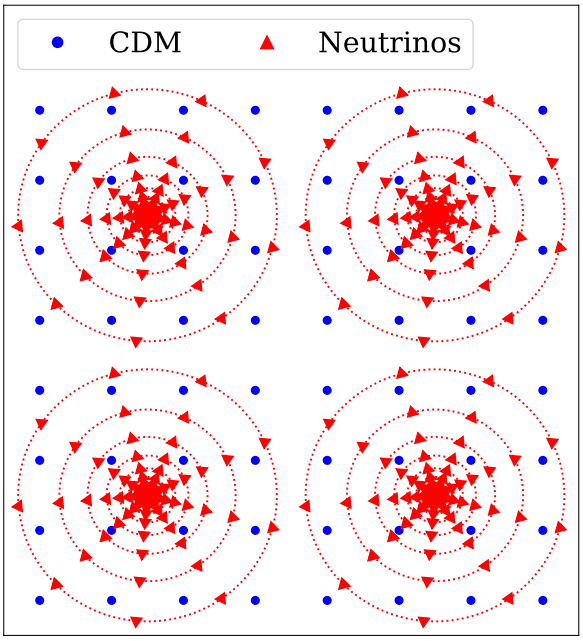

Figure 2.1: 2D illustration of the initial configuration for a neutrino grid 4 times coarser than the CDM grid. Each CDM particle is represented by a blue dot, and each neutrino particle is represented by a red triangle, with velocity in the direction of the triangle. Neutrino particles are initialized on spherical shells of radius proportional to the particle's speed. Each spherical shell is discretized according to a Fibonacci grid (not to scale).

such that the number density of particles in an infinitesimal volume d^3q centered at q is given by

$$dn = f(q)d^3q = q^2 f(q) dq d\Omega. \tag{2.4}$$

In the final step we split the expression into the terms relevant for the velocity magnitude, $q^2 f(q)dq$, and direction (or solid angle), $d\Omega$. To assign the velocity of each particle at a particular site, the magnitude and direction are assigned separately. To assign the magnitude, the magnitude distribution $q^2 f(q)$ is discretely sampled at $N_{\rm shell}$ shells, such that the boundaries of the $i^{\rm th}$ shell are $(q^{(i)}_{\rm min}, q^{(i)}_{\rm max})$, where $1 \leq i \leq N_{\rm shell}$. The left-most boundary is $q^{(1)}_{\rm min} = 0$, i.e. zero velocity, and the right-most boundary $q^{(N_{\rm shell})}_{\rm max} \equiv q_{\rm max}$ is a numerical cutoff such that $q^2_{\rm max} f(q_{\rm max})$ is negligible. The velocity magnitude of particles in the $i^{\rm th}$ shell is given by

$$q^{(i)}_{\rm shell} = \sqrt{\frac{\int_{q^{(i)}_{\rm min}}^{q^{(i)}_{\rm max}} q^4 f(q) dq}{\int_{q^{(i)}_{\rm min}}^{q^{(i)}_{\rm max}} q^2 f(q) dq}}, \tag{2.5}$$

and the mass of particles in the i^{th} shell is given by

$$m_{\text{shell}}^{(i)} = \frac{\int_{q_{\text{min}}^{(i)}}^{q_{\text{max}}^{(i)}} q^2 f(q) dq}{\int_0^{q_{\text{max}}} q^2 f(q) dq}. \tag{2.6}$$

The shell boundaries are chosen according to

$$\frac{1}{N_{\text{shell}}} = \frac{\int_{q_{\text{min}}^{(i)}}^{q_{\text{max}}^{(i)}} g(q) dq}{\int_0^{q_{\text{max}}} g(q) dq}, \tag{2.7}$$

for some arbitrary kernel $g(q)$. This is designed so that each shell has an equal area under $g(q)$, with the choice of $g(q)$ depending on the application. A natural choice would be to use the velocity magnitude distribution $g(q) = q^2 f(q)$, which splits the distribution into shells of equal phase-space volume. However, it was shown in [48] that better results are achieved by using $g(q) = qf(q)$, as this more finely samples the low-velocity tail of the distribution. This better resolves slow neutrino particles, which are most relevant for clustering. We thus employ this choice of kernel.

To study non-degenerate neutrinos one would have to separately sample the Fermi-Dirac distribution for each mass eigenstate and apply the above method multiple times. This would require more neutrino particles in the simulation and thus longer runtimes. To avoid this we use a single effective distribution to approximately describe all eigenstates. Imposing mass conservation, the appropriate distribution is given by

$$\tilde{f}(q) = \sum_{j=1}^{N_\nu} \alpha_j^4 f(\alpha_j q), \tag{2.8}$$

where $\alpha_j \equiv m_j/m_1$ is the mass of eigenstate j divided by the mass of eigenstate 1 (chosen to be the heaviest eigenstate), and N_ν is the number of eigenstates. This expression is exact when there are no cosmological perturbations. The derivation is included in the appendix for this chapter (Section 2.6).

The velocity directions are chosen according to a Fibonacci grid [186, 351, 171] in which each shell isotropically emits particles in $2N_{\text{fib}} + 1$ directions, for integer N_{fib}. We choose a Fibonacci prescription instead of the HEALPix [172] implementation used in [48] as it gives more freedom when selecting the number of directions, and thus the number of neutrino particles. Hence, accounting for the discretization of the magnitude and direction, each neutrino site consists of $N_{\text{shell}} \times (2N_{\text{fib}}+1)$ particles. This gives the total number of neutrinos in the simulation,

$$N_n = N_{\text{sites}} \times N_{\text{shell}} \times (2N_{\text{fib}} + 1). \tag{2.9}$$

Because FastPM employs a Kick-Drift-Kick (KDK) algorithm [300], as opposed to Drift-Kick-Drift, CDM particles initialized close to the neutrino sites will feel a gravitational attraction from the neutrinos and move towards the neutrino sites during the first kick of

the simulation. The same is true for neutrino particles. This will produce spikes in the simulated power spectra at scales corresponding to the neutrino grid spacing, in a similar manner to the findings of [84]. To prevent such numerical artifacts, while keeping the more numerically stable KDK scheme, we take two precautions. Firstly, the neutrino grid is staggered with respect to the CDM grid to separate the two species. Secondly, the neutrino particles are initialized on spherical shells of radius proportional to their thermal velocity magnitude. The radii are chosen such that shells from different sites do not overlap. Because neutrinos are orders of magnitude faster than CDM, this amounts to adding an infinitesimal drift step before the start of the KDK evolution. This drift prevents large overdensities at the neutrino grid sites at the start of the simulation, in turn suppressing the spurious coupling caused by particles getting drawn into neutrino grid sites. We note that the effectiveness of these precautions is sensitive to the mass per neutrino particle. The more massive a neutrino particle, the stronger its gravitational pull on nearby particles. Thus for cosmologies with larger M_ν, a larger value of N_n is required to quell this effect. We will discuss the appropriate choice of N_n in §2.3.

Perturbations

Having initialized the thermal velocities of the neutrino particles in the previous section, we must next include the effects of gravitational perturbations on the initial positions and velocities of all particles in the simulation. To do this, one would typically input the true $z = 0$ linear power spectrum from a Boltzmann solver such as CLASS [73] or CAMB [239]. The simulation would then use a modified linear growth factor to backscale the power spectrum to the starting redshift of the simulation, and in turn set up the initial perturbations. Because N-body simulations make various approximations, such as Newtonian dynamics, the growth factor used for backscaling is modified to contain the same physics as the simulation's forward model. This is done to ensure that the results of the simulation on linear scales matches the true linear physics at $z = 0$ [158]. In the case of massive neutrino simulations, the forward model additionally includes both radiation and neutrinos, which must thus be accounted for when backscaling. This is a non-trivial procedure due to the scale-dependent growth introduced by massive neutrinos. We therefore perform backscaling using REPS [399], which applies the two-fluid approximation to compute the transfer functions of CDM and neutrinos. This is then used to obtain the power spectra and growth rates at the starting redshift of the simulation, and in turn compute the initial perturbations. Moreover, while analytical forms for the 2LPT CDM and neutrino growth factors for massive-neutrino cosmologies have recently been presented in [30], there is currently no framework to apply this to generate non-linear initial conditions for simulations. We thus use the Zel'dovich approximation when setting the initial perturbations, which requires starting the simulation at early times when non-linear effects are small ($z \gtrsim 99$).

In accordance with 'scenario 4' of the REPS paper [399], we treat neutrino particles as non-relativistic in the forward model. Their mass is thus fixed throughout the evolution, and the total matter cosmological parameter used to source the gravitational potential (for

example in Poisson's equation and the growth ODE of §2.2) is computed as

$$\Omega_m(a) = (\Omega_{c,0} + \Omega_{\nu,0})a^{-3}, \tag{2.10}$$

where $\Omega_{c,0}$ and $\Omega_{\nu,0}$ are respectively the CDM and neutrino cosmological parameters at $z = 0$. The initial perturbations computed using REPS are designed to account for this non-relativistic approximation and produce percent-level accuracy in the simulated power spectra at late times. An advantage of REPS over traditional backscaling is that REPS is optimized to give agreement over a range of late redshifts, whereas traditional backscaling optimizes for a single redshift. For consistency in this work, we will also use REPS when initializing runs without massive neutrinos to enable comparison.

To compute the transfer functions, REPS uses a Boltzmann solver, thus its output depends on the parameters used for the Boltzmann solver. In this work we modified the neutrino precision parameters in accordance with appendix B of [130] to improve the accuracy of the transfer functions at small scales. In hindsight this was unnecessary as it causes little difference in the output of FastPM, so we plan to use the default neutrino precision settings in future work.

An alternative approach to backscaling would be to directly input the true linear power spectrum at the starting redshift together with the velocity transfer function. This method has been applied to small volume simulations [69], but would require a more realistic forward model for accurate general implementation.

Evolution

As outlined in section 2.4 of [154], FastPM employs modified kick and drift factors to speed up convergence. This ensures the Zeldovich approximation is accurately followed at each timestep, using the Zeldovich equation of motion $x(a) = q + D(a)s$. We solve for the first order growth factor $D(a)$ using the following ODE [285],

$$D''(a) + \left(2 + \frac{E'(a)}{E(a)}\right) D'(a) = \frac{3}{2}\Omega_m(a)D(a), \tag{2.11}$$

where $D' \equiv dD/d\ln a$, $E(a) \equiv H(a)/H_0$ is the normalized Hubble parameter, and $\Omega_m(a)$ is given in equation 2.10. In this work, the background comprises of radiation (γ), CDM (c), neutrinos (ν), and a cosmological constant (Λ), giving the appropriate Hubble parameter,

$$E(a) = \left[\Omega_{\gamma,0}a^{-4} + \Omega_{c,0}a^{-3} + \Omega_\nu(a)E^2(a) + \Omega_\Lambda\right]^{1/2}. \tag{2.12}$$

The neutrino component is given by

$$\Omega_\nu(a)E^2(a) = \frac{15}{\pi^4}\Gamma_\nu^4\frac{\Omega_{\gamma,0}}{a^4}\sum_{j=1}^{N_\nu}\mathcal{F}\left(\frac{m_j a}{k_B T_{\nu,0}}\right), \tag{2.13}$$

where $\Gamma_\nu \equiv T_{\nu,0}/T_{\gamma,0}$ is the neutrino-to-photon temperature ratio today, m_j is the mass of neutrino species j, and

$$\mathcal{F}(y) \equiv \int_0^\infty dx \, \frac{x^2\sqrt{x^2+y^2}}{1+e^x} \tag{2.14}$$

is an integral arising from the Fermi-Dirac distribution while noting that neutrinos freeze out while relativistic [399]. Note that the neutrino component of the background is treated exactly, with relativistic effects being accounted for. On the other hand, neutrinos are treated as matter-like in the source term of equation 2.11, as motivated in §2.2.

A further point of note is that equation 2.11 assumes the large-scale limit, i.e. scales larger than the neutrino free-streaming scale. This neglects scale-dependent effects by treating neutrinos as non-relativistic particles, analogously to CDM. Because the growth factor is just used by FastPM to speed up convergence, this limit is appropriate as it ensures accelerated convergence on large scales, while letting small scales converge naturally. We set the initial conditions to solve the ODE by assuming matter domination, giving

$$D(a_{\rm ini}) = a_{\rm ini}, \tag{2.15}$$
$$D'(a_{\rm ini}) = a_{\rm ini}. \tag{2.16}$$

We use $z_{\rm ini} = 159$ to enable the simulation to begin at any time after this.

For users of FastPM, we note that FastPM previously assumed a ΛCDM background and thus employed the results of [190, 286] to compute the growth factor, and [81] to approximate the growth rate. This is unsuitable for neutrino simulations and has thus been replaced with the above.

2.3 Results

We consider a 1Gpc/h box with CDM and neutrino grid-numbers given by $N_c^{1/3} = 512$ and $N_{\rm sites}^{1/3} = 128$ respectively. The resolution of the force mesh is always chosen as $N_{\rm mesh}^{1/3} = 2N_c^{1/3}$. The cosmological parameters are set as follows: $h = 0.6711$, $\Omega_m = 0.3175$, $T_\gamma = 2.7255$K, $N_{\rm eff} = 3.046$, $\Omega_k = 0$, $A_s = 2.4 \times 10^{-9}$, and $n_s = 0.9624$. We begin by considering 3 neutrinos of total mass $M_\nu = 0.2$eV. Simulations are started at $z = 99$, at which time non-linear effects are small, as required for an accurate Zeldovich approximation. In order to achieve accurate results with a small number of timesteps, we first take 5 steps in $\log a$ until $z = 19$, which is a sufficiently early time before non-linear neutrino effects come into play. We then take a further $N_{\rm steps}^{\rm lin}$ steps, spaced linearly in a, until $z = 0$. Throughout this section we consider a single run of FastPM; averaging over many realizations would reduce variance, but is unnecessary for the purposes of this work.

We first study the case of 3 degenerate neutrinos. Figure 2.2 shows the ratio of the total matter (CDM+neutrinos) power spectrum between a cosmology with and without massive neutrinos. Different combinations of $N_{\rm steps}^{\rm lin}$ and $N_{\rm fib}$ are considered. For comparison, we

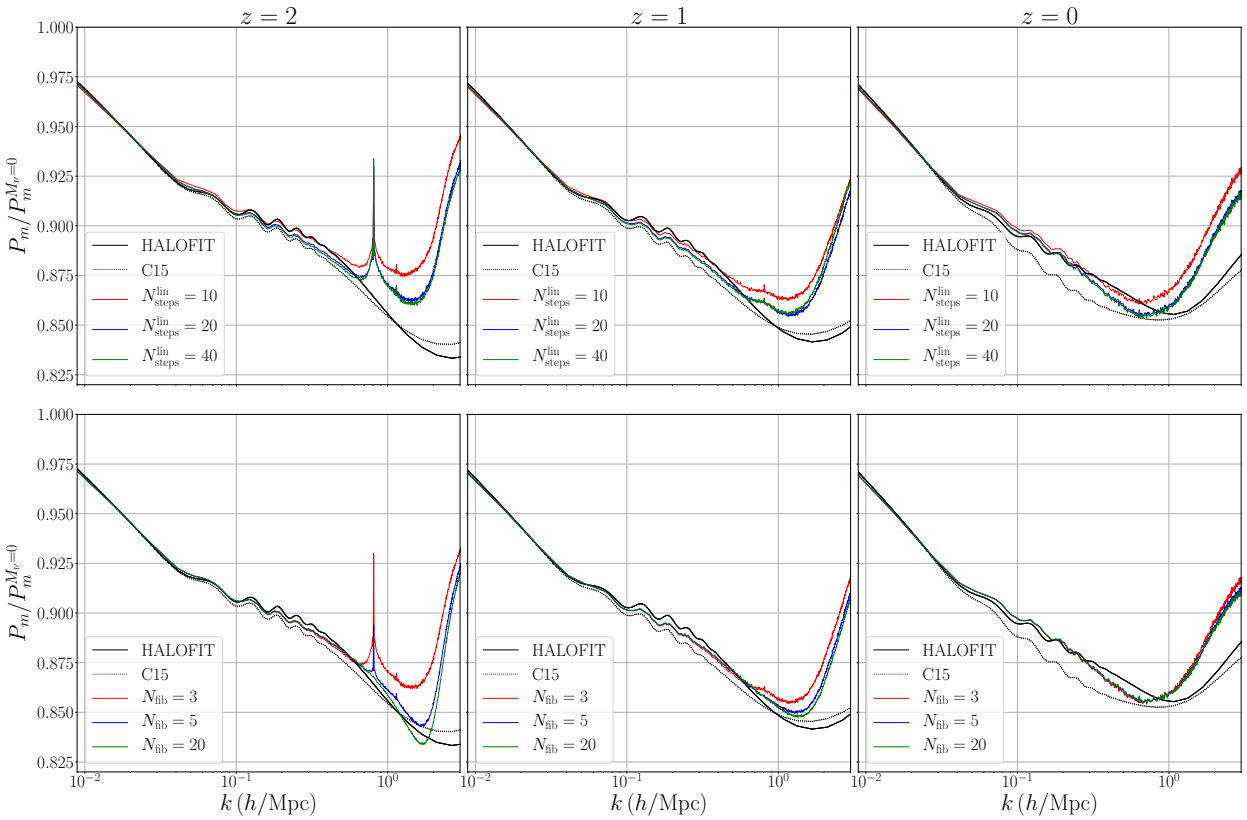

Figure 2.2: Ratio of massive to massless power spectrum at $z = 2, 1, 0$ (left to right) for 3 degenerate $M_\nu = 0.2$eV neutrinos. The top row shows the variation with $N_{\rm steps}^{\rm lin}$ for fixed $N_{\rm fib} = 3$. The bottom row shows the variation with $N_{\rm fib}$ for fixed $N_{\rm steps}^{\rm lin} = 20$. Theoretical predictions based on HALOFIT (solid black) and C15 [95] (dotted black) are also shown.

also plot the theoretical predictions obtained from HALOFIT [342, 354, 68], as well as the modification of [95] which will henceforth be referred to as C15. While these are not exact theoretical predictions they provide a useful diagnostic. Each column of Figure 2.2 represent a different redshift, $z = 2, 1, 0$ from left to right. The top row considers the variation of $N_{\rm steps}^{\rm lin}$ while holding $N_{\rm fib} = 3$ fixed. It can be seen that all choices of steps produce accurate results on large scales, and that the result is suitably converged on small scales by $N_{\rm steps}^{\rm lin} = 20$. One important point to note is the occurrence of a spike at $z = 2, 1$ at $k \sim 0.8h/\mathrm{Mpc}$. This spike corresponds to the spacing of the neutrino grid, and arises due particles being gravitationally attracted to the neutrino grid sites at the start of the simulation, as discussed in §2.2. This numerical artifact can be removed by distributing the neutrino mass over a larger number of particles, for example by increasing $N_{\rm fib}$. To show this, the bottom row of Figure 2.2 considers the variation of $N_{\rm fib}$ while holding $N_{\rm steps}^{\rm lin} = 20$ fixed. It can be seen that $N_{\rm fib} = 5$ and $N_{\rm fib} = 20$ lead to accurate results at $z = 1$ and $z = 2$ respectively. Thus if one wishes to study these earlier redshifts, one must use the appropriate $N_{\rm fib}$. Table 2.1 summarizes some

Name	$N_c^{1/3}$	$N_{\rm mesh}^{1/3}$	$N_{\rm sites}^{1/3}$	$N_{\rm shell}$	$N_{\rm fib}$	$N_{\rm steps}^{\rm log}$	$N_{\rm steps}^{\rm lin}$	Runtime Increase
NC512_NF3	512	1024	128	10	3	5	20	25%
NC512_NF20	512	1024	128	10	20	5	20	115%
NC1024_NF3	1024	2048	128	10	3	5	20	6%
NC1024_NF20	1024	2048	128	10	20	5	20	20%

Table 2.1: A summary of parameters used for some of the runs considered in this paper. Also included is the percentage increase in runtime due to massive neutrinos (discussed in §2.4). In all cases the force mesh is two times finer than the CDM grid. The two differences between the runs are the values of N_c and $N_{\rm fib}$. Increasing N_c enables studying smaller scales, while increasing $N_{\rm fib}$ enables studying higher redshift. As discussed in the text, $N_{\rm fib} = 3$ is suitable to study only $z = 0$, but $N_{\rm fib} = 20$ is required for $z = 2$.

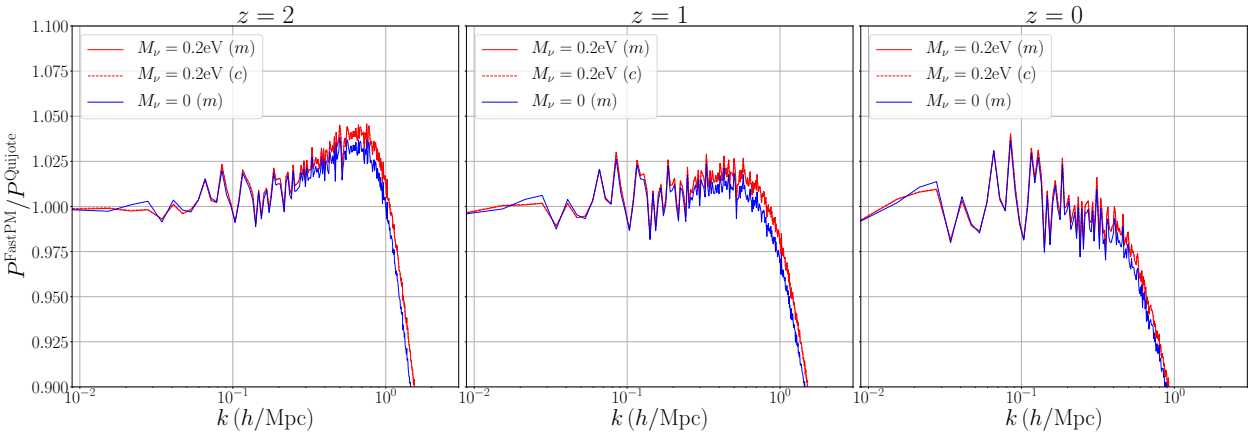

Figure 2.3: Comparison of the FastPM and Quijote power spectra at $z = 2, 1, 0$ (left to right) for a cosmology with degenerate massive neutrinos with $M_\nu = 0.2$eV. Specifically, we consider the NC512_NF20 FastPM run and the "M_ν^{++}" Quijote run. The solid red line compares the total matter power spectrum (m), while the dashed red line compares the CDM power spectrum (c) – note it is difficult to distinguish the two by eye. Also included is a massless neutrino cosmology with matched σ_8 (blue).

typical choices of parameters for runs with massive neutrinos. So far we have illustrated that NC512_NF3 is suitable for $z = 0$ simulations, while NC512_NF20 should be used when one is interested in redshifts up to $z = 2$. In the remainder of this section we will consider the NC512_NF20 run in order to study $z \leq 2$, unless stated otherwise.

For a more careful analysis, Figure 2.3 compares the matter and CDM power spectra from FastPM with Quijote [383], a full N-body simulation. We consider the NC512_NF20 FastPM run and the "M_ν^{++}" Quijote run. For reference we also plot the matter power spectrum for a massless neutrino cosmology with matched σ_8. Firstly, it can be seen that both P_c and P_m show equally good agreement in the massive neutrino case – the dashed red line overlaps

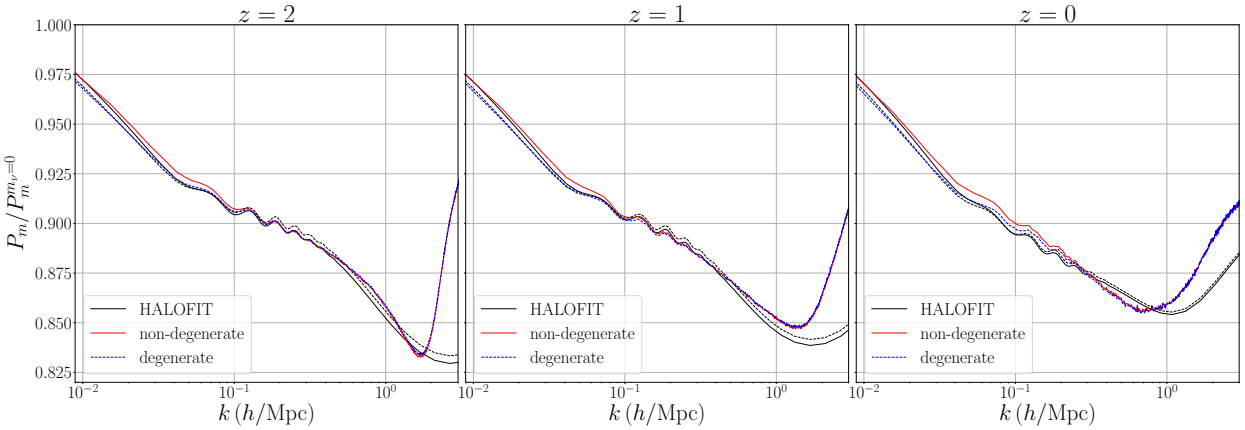

Figure 2.4: Ratio of massive to massless power spectrum at $z = 2, 1, 0$ (left to right) for 3 non-degenerate (solid red) and degenerate (dashed blue) neutrinos with $M_\nu = 0.2$eV. The non-degenerate masses are $0.12, 0.06, 0.02\,$eV. Theoretical predictions based on HALOFIT are shown for both the non-degenerate (solid black) and degenerate (dashed black) cases.

the solid red line – hence FastPM computes both power spectra with equivalent accuracy. Secondly, the difference between FastPM and Quijote is comparable in both the massive (red) and massless (blue) neutrino case, suggesting that any discrepancy with Quijote is not due to the inclusion of massive neutrino particles. There is generally good agreement on large scales and an apparent under-prediction of the power on small scales. The reason for this is that while FastPM uses a particle-mesh approach to compute the forces throughout the simulation, Quijote employs tree methods at low redshift. This leads to Quijote producing more power on small scales, explaining the rapid drop in $P^{\mathrm{FastPM}}/P^{\mathrm{Quijote}}$ at large k — we note that this is not due to the shot noise present in Quijote. It can also be seen that there is a slight bump on intermediate scales at $z = 2$, which is less prominent at lower redshift. We found that the bump grows when using a finer force mesh or initial-condition mesh in FastPM. We thus believe the bump is due to our use of a finer force mesh than that used by Quijote. This has the effect of increasing the power on small scales, but is eventually dominated by Quijote's tree force calculation on small scales and late times, therefore it is only significant at $z = 2$. The exact nature of the bump is also dependent on the parameters used in Quijote that define the redshift and scale at which the particle-mesh to tree transition occurs.

Next, we investigate the performance of our approximation for non-degenerate neutrinos given in equation 2.8. Figure 2.4 compares the massive to massless power spectrum ratio for the case of 3 neutrinos of masses $0.12, 0.06, 0.02$eV. The degenerate case is also included for reference. The agreement of the non-degenerate simulation with the theoretical lines is good on large scales and worsens on intermediate scales. This is likely because REPS assumes degeneracy when computing the initial conditions, causing a relative gain in power. Even so, the non-degenerate results are of suitable accuracy for studying such mass schemes in

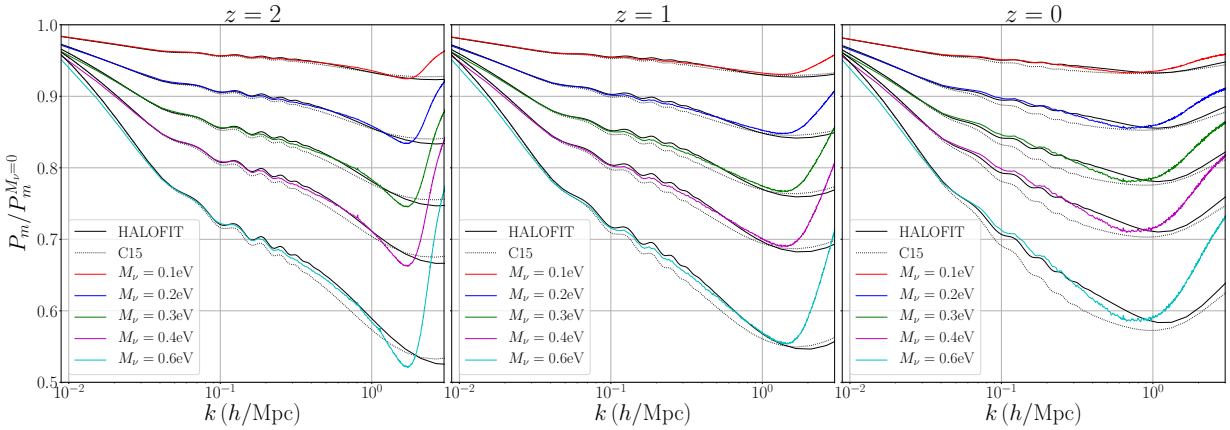

Figure 2.5: Ratio of massive to massless power spectrum at $z = 2, 1, 0$ (left to right) for a variety of M_ν, using $N_{\mathrm{steps}}^{\mathrm{lin}} = 20$ and $N_{\mathrm{fib}} = 20$. Note that 10 additional steps were taken at early z for $M_\nu = 0.6$eV, as discussed in the text. The theoretical predictions of HALOFIT (solid black) and C15 [95] (dotted black) are also shown.

the context of future surveys.

To investigate the accuracy of FastPM for different choices of neutrino mass, Figure 2.5 shows the ratio of the matter power spectrum between a massive and massless neutrino cosmology for a variety of choices of M_ν. It can be seen that there is good agreement for the full range of physical interest ($M_\nu \lesssim 0.6$eV). Increasing M_ν beyond 0.2eV leads to a small spike at $z = 2$ caused by the neutrino grid, as discussed in §2.2. This is an expected result of the increase in mass per neutrino particle and can be alleviated by a small increase in N_{fib}, or alternatively by increasing the number of steps at early redshift to prevent particles getting drawn into the neutrino grid sites. For $M_\nu = 0.6$eV, which is the upper bound of physical interest, the data in Figure 2.5 was generated using an extra 10 steps in $\log a$ between $z = 99$ and 79 to avoid the occurrence of a lager spike. While interest in cosmologies with $M_\nu = 0.6$eV is limited, it is useful to know that accurate results can be achieved with an additional 10 steps compared to lower mass runs.

FastPM is also capable of computing the CDM-neutrino cross-power spectrum, as required for observables such as galaxy-galaxy lensing. Figure 2.6 compares the FastPM cross-power to the linear cross-power computed by CLASS. There is good agreement on large scales, and the agreement worsens as k increases due to non-linear effects that are not simulated by CLASS. There is negligible dependence on N_{fib} at large scales. We note that the cross-power is always weighted by a factor of f_ν in cosmological observables, thus one can tolerate larger error on the cross-power and still produce accurate observable predictions. We also found that the spikes discussed in §2.2 do occur in the cross-power at $z = 2$, but are negligible for $z \leq 1$. This effect can be reduced at $z = 2$ by using a finer neutrino grid, which is relatively inexpensive for large N_c simulations (as we will discuss in §2.4).

While not directly observable, the neutrino power spectrum serves as a useful diag-

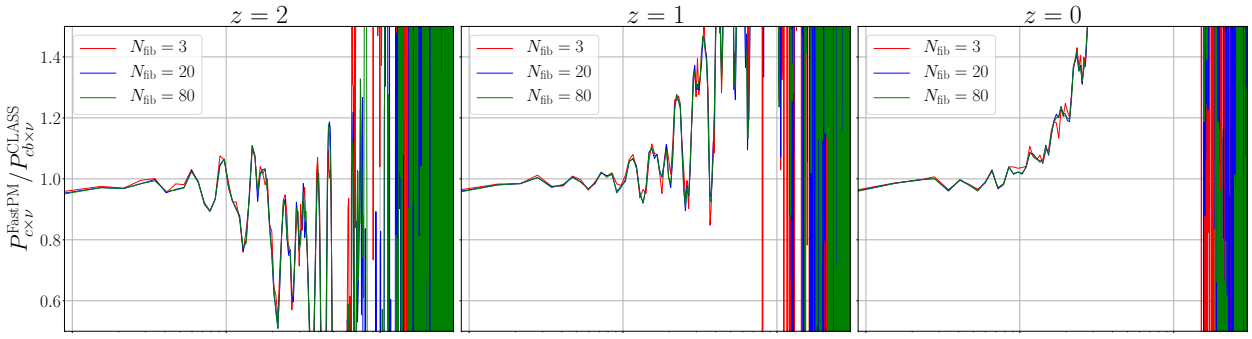

Figure 2.6: Comparison of the FastPM CDM-neutrino cross-power spectrum with the linear cross-power from CLASS for $M_\nu = 0.2$eV using a variety of $N_{\rm fib}$. Note that, unlike FastPM, CLASS includes baryonic effects.

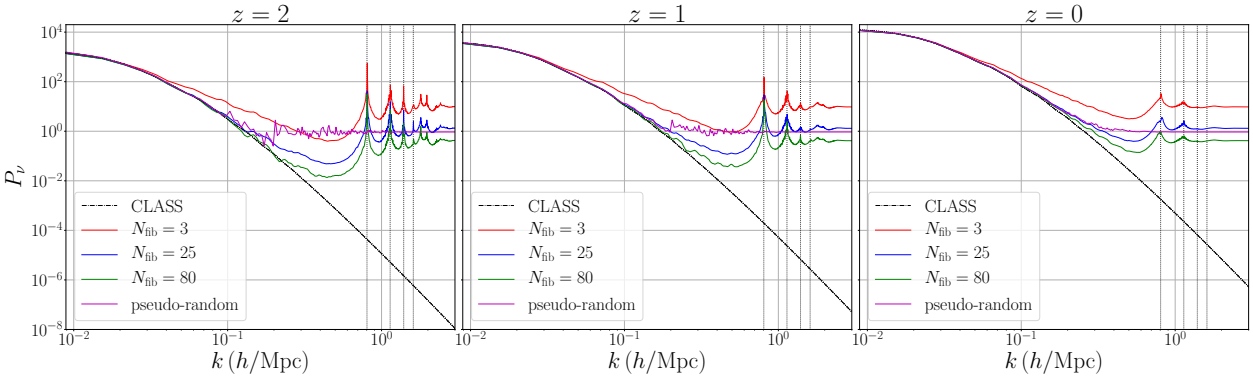

Figure 2.7: The FastPM neutrino power spectrum computed using a variety of $N_{\rm fib}$ for $M_\nu = 0.2$eV. The linear result from CLASS is also included (black). For reference, a FastPM simulation initialized with pseudo-random neutrino thermal velocities is shown (magenta), and can be seen to produce shot noise. The total number of neutrinos, N_n, for the pseudo-random run is approximately the same as for the $N_{\rm fib} = 25$ quasi-random run, hence the similar power as $k \to \infty$. Also shown are dotted vertical lines representing the four smallest wavenumbers associated with the neutrino grid: $k_n, \sqrt{2}k_n, \sqrt{3}k_n, 2k_n$.

nostic for the quasi-random sampling scheme. Figure 2.7 shows good agreement between the neutrino power spectrum computed by FastPM and CLASS on large scales. There is more sensitivity to $N_{\rm fib}$ compared to the cross-power, with larger $N_{\rm fib}$ required to ensure convergence at progressively smaller scales. It can be seen that quasi-random sampling produces noisy P_ν on small scales, even with $N_{\rm fib} = 80$. To enable comparison with the noise produced by a typical pseudo-random sampling scheme we perform a FastPM simulation using pseudo-randomly sampled neutrino thermal velocities. As expected, a pseudo-random scheme produces shot noise, which in Figure 2.7 is manifested by the flattening of the

power for $k \gtrsim 2 \times 10^{-1} h/\text{Mpc}$. For the pseudo-random example, a grid of $N_n^{\text{pseudo}} = 1024^3$ neutrino particles was used, with one neutrino per grid site. Since shot noise is known to scale with the total number of neutrino particles, we compare with a quasi-random scheme using $N_{\text{sites}} = 128^3$, $N_{\text{shell}} = 10$ and $N_{\text{fib}} = 25$, such that, using equation 2.9, $N_n/N_n^{\text{pseudo}} = 128^3 \times 10 \times (2 \cdot 25 + 1)/1024^3 = 0.996 \approx 1$. As expected, the $N_{\text{fib}} = 25$ and pseudo-random schemes approximately have the same power as $k \to \infty$ (there is a slight difference because the number of neutrino particles is not exactly matched). Firstly, it can be seen that the quasi-random scheme enables study of smaller scales compared to the pseudo-random scheme, and, ignoring spikes, has lower small-scale noise. Secondly, the difference in noise between the two sampling approaches becomes larger at earlier redshifts. Thirdly, the pseudo-random power fluctuates around the more stable quasi-random power — this can most clearly be seen at $z = 2$ for $k \sim 10^{-1} h/\text{Mpc}$. In fact, it was shown in [48] that such fluctuations are caused by early time artifacts produced by pseudo-random sampling, and also leaves a signature on P_m at scales as large as $k \sim 10^{-2} h/\text{Mpc}$. Thus quasi-random sampling not only helps avoid shot noise on small scales, but also reduces noise at larger scales.

One apparent drawback of the quasi-random scheme is the introduction of spikes due to spurious correlations between CDM and neutrino particles. Spurious correlations were first noted by [84] in the context of hybrid simulations. However, this study did modify the methodology of [48] by using a different sampling scheme for the Fermi-Dirac distribution and by initializing neutrino particles at a late redshift of $z = 4$ (as typical for hybrid methods), making it difficult to present a direct comparison. To understand the nature of the spurious peaks observed in FastPM, we consider the Fourier transform of the initial neutrino grid. For a box of side length $1\text{Gpc}/h$ and $N_{\text{sites}}^{1/3} = 128$ neutrino grid sites per side, the fundamental wavenumber is given by

$$k_n = 2\pi \frac{128}{1\text{Gpc}/h} \approx 0.8 h/\text{Mpc}. \tag{2.17}$$

This corresponds to the spacing between two adjacent neutrino grid sites. Because the grid is 3-dimensional, the next three smallest wavenumbers are $\sqrt{2}k_n$, $\sqrt{3}k_n$, and $2k_n$. It can be seen in Figure 2.7 that the spikes in P_ν exactly align with these wavenumbers. This verifies the explanation in §2.2 that particles are drawn to the large overdensities at neutrino grid sites at the start of the simulation, in turn leaving a numerical artifact at late times. Hence, in our FastPM implementation, the spikes are a result of the coarse neutrino grid and can be removed by using a denser grid. That said, P_ν has little effect on the small scale behaviour of cosmological observables as it is always weighted by a factor of f_ν^2, and is itself small. Thus it typically suffices to use a coarse neutrino grid with a large enough number of neutrino particles per grid site, as shown for P_m in the discussion surrounding Figure 2.2.

Finally, we consider using a finer CDM grid with $N_c^{1/3} = 1024$. As a reference we use an $N_c^{1/3} = 1024$ Gadget [346] simulation for degenerate neutrinos with $M_\nu = 0.12\text{eV}$. The left of Figure 2.8 shows a fixed-amplitude comparison between the Gadget simulation and $N_c^{1/3} = 512$ & 1024 FastPM simulations, considering a variety of step numbers. We use $N_{\text{fib}} = 3$ as we only compare $z = 0$. In terms of Table 2.1, we consider `NC512_NF3` and

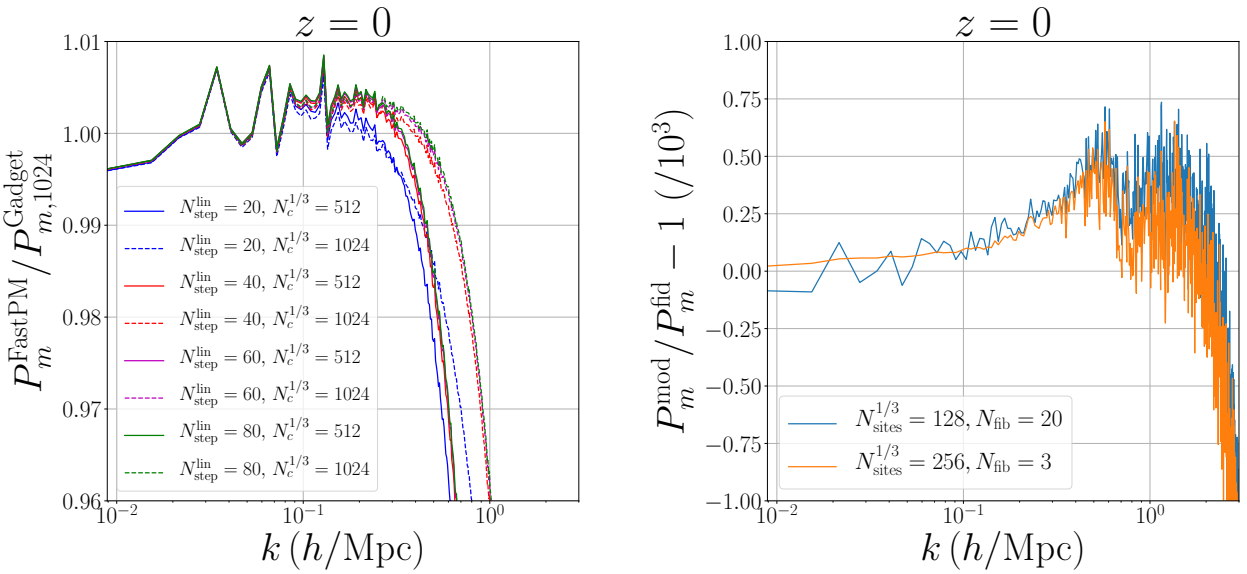

Figure 2.8: (Left) Comparison of the FastPM matter power spectrum at $z = 0$ for a run with degenerate massive neutrinos, $M_\nu = 0.12$eV, with an $N_c^{1/3} = 1024$ Gadget simulation. The step size and CDM grid are varied. (Right) Ratio of P_m for FastPM runs with modified (mod) $N_{\rm sites}^{1/3}$ and $N_{\rm fib}$ compared to the fiducial (fid) values of 128 and 3 from `NC1024_NF3` of Table 2.1. Note that the ratio has been shifted by -1 and scaled by 10^3, thus the vertical range represents a ratio range of 1 ± 10^{-3}, i.e. $\pm 0.1\%$.

`NC1024_NF3`, while varying $N_{\rm steps}^{\rm lin}$. It can be seen that there is sub-percent agreement on large scales, and that using a two-times finer grid leads to agreement at approximately two-times higher k, as expected. Moreover, increasing $N_{\rm steps}^{\rm lin}$ to 40 extends the accuracy to slightly smaller scales, but the difference compared to $N_{\rm steps}^{\rm lin} = 20$ is small.

It is important to note that one does not need to increase N_n with N_c to obtain accurate results at small scales. This is illustrated on the right of Figure 2.8 where we consider modifications of `NC1024_NF3` to increase N_n. It can be seen that increasing either $N_{\rm sites}^{1/3}$ or $N_{\rm fib}$ beyond the fiducial values of 128 and 3 causes a negligible ($< 0.1\%$) change in the $z = 0$ power spectrum. This is a key finding in terms of studying smaller scales, as it shows one need only increase the number of CDM particles while keeping the number of neutrino particles fixed. This is aided by the fact that small scales are almost entirely dominated by the CDM evolution and the background cosmology. While the results presented here are for a run with $N_c^{1/3} = 1024$, $M_\nu = 0.12$eV, and $z = 0$, we find similar results for larger N_c, M_ν, and z – in all cases there is a sub-percent change in the power spectrum when increasing N_n. Thus the relative increase in runtime-per-step caused by the inclusion of massive neutrinos will decrease as N_c increases, enabling the study of small scales with only a small increase in runtime-per-step. We will now give a more thorough account of the runtime.

2.4 Runtime

Firstly, we note that FastPM previously considered only CDM and Λ. To include massive neutrinos, this work has added functionality to simulate radiation in the background evolution by including photons and massless neutrinos. The runtime increase caused by this is negligible, thus it is the inclusion of neutrino particles, required to simulate massive neutrinos, that must be considered when studying runtime. To enable comparison in the following discussion, we compute the percentage difference in runtime-per-step between simulations of massive and massless neutrinos by using the same non-neutrino parameter values (including the number of timesteps) and number of CPUs in both cases.

Simulations were performed using the Cori supercomputer at the National Energy Research Scientific Computing Center (NERSC). We first consider $N_{\rm fib} = 3$ runs, shown in §2.3 to give accurate results at $z = 0$. For the `NC512_NF3` run of Table 2.1, there are approximately equal numbers of CDM and neutrino particles: using equation 2.9 gives $N_n/N_c = 10(2 \times 3 + 1)/4^3 = 1.1$. Such runs can be performed on a single Cori Haswell node in $\sim$ 715s, whereas the corresponding massless neutrino run takes $\sim$ 565s. Thus for this configuration there is a 25% increase in runtime. We find that doubling both $N_c^{1/3}$ and $N_{\rm sites}^{1/3}$ requires 8 nodes and also shows a 25% increase in runtime. However, as discussed at the end of the previous section, one does not need to increase N_n as one increase N_c – accurate results can be achieved by increasing $N_c^{1/3}$ to 1024, while keeping $N_{\rm sites}^{1/3} = 128$ fixed (`NC1024_NF3`). Such a run requires only 4 nodes, and has an increase in runtime of 6%, as massive and massless runs take $\sim$1387s and 1307s respectively. As expected the change in runtime is sensitive to the ratio of N_n to N_c, hence simulations with progressively larger N_c and fixed N_n have a smaller relative increase in runtime-per-step. This means that even runs with $N_{\rm fib} = 20$, required for accurate results at $z = 2$, only have an increase in runtime of 20% for $N_c^{1/3} = 1024$ (`NC1024_NF20`). Using a lower $N_c^{1/3}$ of 512 with $N_{\rm fib} = 20$ (`NC512_NF20`) does lead to a larger runtime increase of 115% because in this case $N_n/N_c \approx 6$. However, this large runtime increase is not problematic as $N_c^{1/3} = 512$ runs without massive neutrinos are relatively inexpensive anyway. The key results of this paragraph are reported in Table 2.1.

Massive neutrino runs typically require more timesteps than runs without massive neutrinos: while FastPM can achieve high accuracy for cosmologies without massive neutrinos in a couple of steps [154], a massive neutrino simulation requires $\sim$ 25 steps. This is because of the need to start simulations at an earlier redshift and to carefully capture the interplay between CDM and neutrinos, as documented in §2.2. Thus the increase in total runtime for massive neutrino simulations is dominated by the need for additional steps. Note that we have ignored the effects of I/O and setting initial conditions; these scale with the total number of particles and will thus also lead to increases in runtime for large runs, but are typically subdominant for $N_c^{1/3} \lesssim 1024$.

2.5 Conclusions

This work has presented a fast and scalable particle-only method to study the effects of massive neutrinos on cosmological structure formation. This is enabled by three key ingredients. Firstly, we sample the Fermi-Dirac distribution in a low entropy, quasi-random manner when setting the neutrino initial conditions. This reduces the noise that typically plagues pseudo-random neutrino particle simulations, with the reduction becoming more significant at earlier redshift. Secondly, while massive neutrinos introduce scale dependence, we use REPS [399] to set the initial perturbations via the two-fluid approximation transfer functions. This allows us to treat neutrinos as non-relativistic particles throughout the evolution and achieve accurate results at low redshift. Finally, we incorporate the above methodology into FastPM [154] to enable fast evolution. Altogether, the simulation produces accurate results for the matter, CDM, CDM-neutrino, and neutrino power spectra across the full range of neutrino masses permitted by current experimental constraints ($M_\nu \lesssim 0.6$eV) at $z \lesssim 2$. Furthermore, the increase in runtime-per-step due to massive neutrinos is small, as the required number of neutrino particles is typically less than or similar to the number of CDM particles. Together with the fact a run requires ~ 25 steps, FastPM is considerably faster than alternative schemes based on full N-body simulations.

We have also addressed the problem of small scale spurious correlations caused by the quasi-random sampling method of [48], found by [84] for hybrid simulations. We have argued that, in our setup, spurious correlations are caused by the neutrino grid being coarser than the CDM grid, leading to nearby particles being attracted to the neutrino grid sites at the start of the simulation. Such spurious correlations in P_m and P_c can be adequately reduced by applying an infinitesimal drift step for neutrinos at the start of the simulation, and using a sufficiently large number of neutrino particles. The cross-power $P_{c\times\nu}$ is similarly free of numerical artifacts arising from the sampling scheme — this is true at $z = 2$ provided a sufficiently fine neutrino grid is used. Any remaining artifacts in P_ν are rendered subdominant by two effects: first, the contribution of the P_ν term to any observable is weighed by a factor of f_ν^2, and second, P_ν itself is extremely damped on small scales compared to P_c.

There are many avenues for future work. Our technique provides a quick way to predict the clustering of both CDM and total matter down to $\sim 1\text{Mpc}/h$ in the presence of massive neutrinos. Combined with FastPM's inbuilt halo finder [154], analysis pipelines for fitting cosmological parameters can be built by interfacing with `nbodykit` [185]. This will enable the prediction of galaxy-clustering and weak-lensing measurements for surveys such as DES [5, 363]; one could implement an emulator-like approach [233, 225, 216, 141, 401, 388, 218] to study the effects of massive neutrinos on clustering. Moreover, recent work in effective field theory applied to BOSS [16] has suggested that combining the full-shape BOSS data with Planck [292] can reduce the upper limit of the sum of neutrino mass to $M_\nu < 0.16$eV (95% CL) [127, 204]. One could test these results by performing a re-analysis of BOSS that considers small scale neutrino effects. Finally, because massive neutrinos modify the information content of the CDM and total matter fields, one can use the techniques of [329, 153, 266, 196] to study the effect of massive neutrinos on reconstruction.

Acknowledgements

We thank Uroš Seljak for fruitful discussion on the project. We also thank Yin Li and Patrick McDonald for insightful discussion on the methodology. We acknowledge the use of `nbodykit` [185] for computing the power spectra presented in this work. This research made use of the Cori supercomputer at the National Energy Research Scientific Computing Center (NERSC), a U.S. Department of Energy Office of Science User Facility operated under Contract No. DE-AC02-05CH11231.

2.6 Appendix A: Effective distribution for non-degenerate neutrinos

We seek to describe a set of non-degenerate neutrinos with masses $\{m_j\}_{j=1}^{N_\nu}$ by a single effective particle, for use in the sampling scheme described in §2.2. Because equations 2.5, 2.6, and 2.7 are fractions of moments of the Fermi-Dirac distribution $f(q)$, we need only find the distribution function for the effective particle $\tilde{f}(q)$ up to a constant factor and arbitrary transformation of the argument. Working in dimensionless units, as in equations 2.3 and 2.4, the number of particles of eigenstate j in an infinitesimal volume of size d^3q_j is

$$dn_j = f(q_j)d^3q_j. \tag{2.18}$$

The j dependence arises due to the implicit dependence of q_j on the non-degenerate mass m_j. Using the non-relativistic dispersion relation, the scaling of q_j is given by $q_j \sim m_j$. We thus change variables to $q \equiv q_j/\alpha_j$ with $\alpha_j \equiv m_j/\mu$, for some constant with units of mass μ, giving the number of particles of eigenstate j in the *common* infinitesimal volume d^3q,

$$dn_j = f(\alpha_j q)\alpha_j^3 d^3q. \tag{2.19}$$

The number of effective particles in d^3q, denoted $d\tilde{n}$, is defined such that

$$d\tilde{n} \equiv \tilde{f}(q)d^3q, \tag{2.20}$$

and the effective particle mass is denoted $\tilde{m}$. Enforcing mass conservation in each infinitesimal volume d^3q gives

$$\tilde{m}d\tilde{n} = \sum_j m_j dn_j \tag{2.21}$$

$$\tilde{m}\tilde{f}(q)d^3q = \sum_j m_j f(\alpha_j q)\alpha_j^3 d^3q. \tag{2.22}$$

Rearranging gives

$$\tilde{f}(q) = \frac{\mu}{\tilde{m}}\sum_j \alpha_j^4 f(\alpha_j q) \propto \sum_j \alpha_j^4 f(\alpha_j q), \tag{2.23}$$

which is the result stated in equation 2.8, having dropped the constant factor which is unneeded for the sampling algorithm. The choice of μ to define $\alpha_j = m_j/\mu$ is arbitrary, but we choose $\mu = m_1$, the mass of the heaviest eigenstate, for two reasons. Firstly, as long as the mass ratios α_j are close to 1, a good sampling for the heaviest eigenstate also implies a good sampling for the other mass eigenstates. If the mass ratio of an eigenstate is much smaller than 1, then there will be no significant clustering for this light eigenstate, and the sampling scheme is irrelevant. Moreover, this choice ensures that $\tilde{f}$ will equal the correct Fermi-Dirac distribution in the degenerate limit, because $\alpha_j \to 1 \; \forall \; j$.

Chapter 3

Detecting neutrino mass by combining matter clustering, halos, and voids

The contents of this chapter was originally published in [62],

> Detecting neutrino mass by combining matter clustering, halos, and voids
> *Bayer A.E., et al.* (arXiv:2102.05049) ApJ 919 1 24 (2021)

In this chapter, we quantify the information content of the non-linear matter power spectrum, the halo mass function, and the void size function, using the *Quijote* N-body simulations. We find that these three statistics exhibit very different degeneracies amongst the cosmological parameters, and thus the combination of all three probes enables the breaking of degeneracies, in turn yielding remarkably tight constraints. We perform a Fisher analysis using the full covariance matrix, including all auto- and cross-correlations, finding that this increases the information content for neutrino mass compared to a correlation-free analysis. The multiplicative improvement of the constraints on the cosmological parameters obtained by combining all three probes compared to using the power spectrum alone are: 137, 5, 8, 20, 10, and 43, for Ω_m, Ω_b, h, n_s, σ_8, and M_ν, respectively. The marginalized error on the sum of the neutrino masses is $\sigma(M_\nu) = 0.018\,\mathrm{eV}$ for a cosmological volume of $1\,(h^{-1}\mathrm{Gpc})^3$, using $k_{\mathrm{max}} = 0.5\,h\mathrm{Mpc}^{-1}$, and without CMB priors. We note that this error is an underestimate insomuch as we do not consider super-sample covariance, baryonic effects, and realistic survey noises and systematics. On the other hand, it is an overestimate insomuch as our cuts and binning are suboptimal due to restrictions imposed by the simulation resolution. Given upcoming galaxy surveys will observe volumes spanning $\sim 100\,(h^{-1}\mathrm{Gpc})^3$, this presents a promising new avenue to measure neutrino mass without being restricted by the need for accurate knowledge of the optical depth, which is required for CMB-based measurements. Furthermore, the improved constraints on other cosmological parameters, notably Ω_m, may also be competitive with CMB-based measurements.

3.1 Introduction

High-precision measurements of large-scale structure from upcoming cosmological surveys, such as DESI[1], Euclid[2], PFS[3], Roman Space Telescope[4], Vera Rubin Observatory[5], SKA[6], and SPHEREx[7], are expected to revolutionize our understanding of fundamental physics, for example, by measuring neutrino mass. To fully realize the potential of these surveys, an urgent task is to determine the key observables that can maximize the scientific return. For Gaussian density fields, the answer is well known: the power spectrum, or equivalently, the correlation function, is the statistic that completely characterizes the field. Therefore, on large scales and at high redshift, where the density fluctuation in the Universe resembles a Gaussian field, the power spectrum encapsulates all the information.

However, at low redshift and on small scales, non-linear gravitational evolution moves information from the power spectrum into higher-order moments. It is currently ill-understood which observable(s) will allow retrieval of the maximum information in the non-linear regime. For instance, it has been shown that for non-Gaussian fields, all clustering information may not be embedded in the infinite N-point statistics [92, 91]. Since the number of modes increases rapidly by going to small scales, it is expected that the amount of information will also increase by considering observables in the mildly to fully non-linear regime. While the amount of information, at least for some parameters, may saturate in the power spectrum [306, 383] [see however 74], many authors have shown that other statistics contain complementary information [see, e.g. 353, 323, 64, 213, 315, 245, 248, 211, 332, 255, 178, 177, 129, 366, 20, 175, 188, 46, 257].

In this paper we quantify the information embedded in the non-linear matter power spectrum, the halo mass function (HMF), and the void size function (VSF). We apply the Fisher formalism using a subset of the *Quijote* simulations [383], comprising of 23,000 N-body simulations for 16 different cosmologies spanning six cosmological parameters: Ω_m, Ω_b, h, n_s, σ_8, and M_ν. We study the information that these probes contain individually and when combined together, showing how the combination of these three statistics breaks degeneracies amongst the cosmological parameters, in turn setting very tight constraints. We consider the effects of both the auto-correlation for each probe and the cross-correlation between different probes when computing the total information content. A simpler, theoretical, treatment combining cluster and void abundances has been studied by [310].

By adding probes such as the halo and void abundances, we demonstrate that it is possible to break the strong degeneracy between M_ν and σ_8 usually seen in 2-point clustering constraints [see, e.g. 382]. In turn, this gives tight constraints on neutrino mass, and in fact

[1] https://www.desi.lbl.gov
[2] https://www.euclid-ec.org
[3] https://pfs.ipmu.jp/index.html
[4] https://wfirst.gsfc.nasa.gov/index.html
[5] https://www.lsst.org
[6] https://www.skatelescope.org
[7] https://www.jpl.nasa.gov/missions/spherex

all cosmological parameters, potentially without the need for including CMB priors. In addition to improved constraints, having multiple independent probes of neutrino masses will allow for more robust controls of systematics.

The paper is organized as follows. We first review the *Quijote* simulations in Section 3.2. The Fisher formalism used to quantify the information content on the different observables is described in Section 3.3. We explain how the matter power spectrum, halo mass function, and void size function are obtained in Section 3.4. We show the results of our analysis in Section 3.5. Finally, we conclude in Section 3.6.

3.2 Simulations

We quantify the information content of different cosmological observables using the Fisher matrix formalism. We model the observables using the *Quijote* simulations [383], a set of 23,000 N-body simulations that at a given redshift contain about 8 trillion (8×10^{12}) particles over a total combined volume of 44,100 $(h^{-1}\mathrm{Gpc})^3$. Each simulation considers a box of size 1 $(h^{-1}\mathrm{Gpc})^3$. The simulation subset used in this work spans a total of 16 different cosmological models that have been designed to evaluate the two ingredients required to compute the Fisher matrix: (1) the covariance matrix of the observables and (2) the derivatives of the observables with respect to the cosmological parameters. Despite their larger computational cost than analytic approaches (e.g. perturbation theory or the halo model), numerical simulations are more accurate into the fully non-linear regime and rely on fewer assumptions and approximations.

We consider six cosmological parameters: Ω_m, Ω_b, h, n_s, σ_8, and M_ν. The set of cosmological parameters is shown in Table 3.1. To evaluate the covariance matrix, we use the 15,000 simulations of the fiducial cosmology. We compute the derivatives by considering simulations where only one cosmological parameter is varied, with all others fixed. We use 1,000 simulations (500 pairs) for each derivative, with the exception of neutrino mass, where we use 1,500 (see below).

The initial conditions (ICs) were generated in all cases at $z = 127$ using second-order Lagrangian perturbation theory (2LPT) for simulations with massless neutrinos, by rescaling the $z = 0$ matter power spectrum using the scale-independent growth factor from linear theory. Because the 2LPT formalism has not yet been developed to account for massive neutrinos, the ICs for massive neutrino cosmologies adopt the Zel'dovich approximation with scale-dependent growth factors and rates, following [399]. For this reason there is also a 'Fiducial (ZA)' class of simulations, which is identical to the fiducial simulations but with Zel'dovich ICs to match the M_ν simulations [see 383, for further details]; this enables accurate computation of derivatives with respect to M_ν. Note that in the full *Quijote* simulations there are two sets of Ω_b cosmologies; we use the Ω_b^{++} and Ω_b^{--} set too obtain smoother derivatives.

All simulations follow the evolution of 512^3 dark matter particles down to $z = 0$. The simulations with massive neutrinos also contain 512^3 neutrino particles. The gravitational

Quijote Simulations								
Name	Ω_m	Ω_b	h	n_s	σ_8	M_ν(eV)	ICs	Realizations
Fiducial	0.3175	0.049	0.6711	0.9624	0.834	0.0	2LPT	15,000
Fiducial ZA	0.3175	0.049	0.6711	0.9624	0.834	0.0	Zel'dovich	500
Ω_m^+	0.3275	0.049	0.6711	0.9624	0.834	0.0	2LPT	500
Ω_m^-	0.3075	0.049	0.6711	0.9624	0.834	0.0	2LPT	500
Ω_b^{++}	0.3175	0.051	0.6711	0.9624	0.834	0.0	2LPT	500
Ω_b^{--}	0.3175	0.047	0.6711	0.9624	0.834	0.0	2LPT	500
h^+	0.3175	0.049	0.6911	0.9624	0.834	0.0	2LPT	500
h^-	0.3175	0.049	0.6511	0.9624	0.834	0.0	2LPT	500
n_s^+	0.3175	0.049	0.6711	0.9824	0.834	0.0	2LPT	500
n_s^-	0.3175	0.049	0.6711	0.9424	0.834	0.0	2LPT	500
σ_8^+	0.3175	0.049	0.6711	0.9624	0.849	0.0	2LPT	500
σ_8^-	0.3175	0.049	0.6711	0.9624	0.819	0.0	2LPT	500
M_ν^+	0.3175	0.049	0.6711	0.9624	0.834	0.1	Zel'dovich	500
M_ν^{++}	0.3175	0.049	0.6711	0.9624	0.834	0.2	Zel'dovich	500
M_ν^{+++}	0.3175	0.049	0.6711	0.9624	0.834	0.4	Zel'dovich	500

Table 3.1: Characteristics of the subset of the *Quijote* simulations used in this work. The fiducial cosmology contains 15,000 simulations, that are used to compute the covariance matrix. In the other cosmological models, one parameter is varied at a time, and these simulations are used to compute the numerical derivatives. The initial conditions of all simulations were generated at $z = 127$ using 2LPT, except for the simulations with massive neutrinos and a copy of the fiducial cosmology, where the Zel'dovich approximation is used (see main text for further details). All realizations follow the evolution of 512^3 CDM (+ 512^3 Neutrino) particles in a box of size 1 h^{-1}Gpc down to $z = 0$, with a gravitational softening length 50 h^{-1}kpc. For massive neutrino simulations, we assume three degenerate neutrino masses.

force tree for neutrinos is turned on at $z = 9$. The gravitational softening for both dark matter and neutrinos is 50 h^{-1}kpc (1/40 of the mean interparticle distance). In this work, we consider redshift $z = 0$ only.

3.3 Fisher information

We use the Fisher matrix formalism [358, 191, 192, 373] to calculate the information embedded in the non-linear matter power spectrum, the halo mass function and the void size

function, individually and when combined. The Fisher matrix is defined as

$$F_{ij} = -\left\langle \frac{\partial^2 \log \mathcal{L}}{\partial\theta_i \partial\theta_j} \right\rangle \ , \tag{3.1}$$

where $\mathcal{L}$ is the likelihood and $\vec{\theta}$ is the vector representing the parameters of the model [159]. Under the assumption that the region around the maximum of the likelihood can be approximated as a multivariate normal distribution, one can write the Fisher matrix as

$$\begin{aligned} F_{ij} &= \frac{1}{2}\left[\frac{\partial\vec{O}}{\partial\theta_i}C^{-1}\frac{\partial\vec{O}^T}{\partial\theta_j} + \frac{\partial\vec{O}}{\partial\theta_j}C^{-1}\frac{\partial\vec{O}^T}{\partial\theta_i}\right] \\ &\quad + \frac{1}{2}\mathrm{Tr}\left[C^{-1}\frac{\partial C}{\partial\theta_i}C^{-1}\frac{\partial C}{\partial\theta_j}\right] \ , \end{aligned} \tag{3.2}$$

where $\vec{O}$ is the vector with the values of the observables and C is the covariance matrix. In order to avoid underestimating the errors, we follow [90] and neglect the dependence of the covariance on the cosmological parameters, by setting the last term of Eq. 3.2 to zero. This is necessary when assuming a Gaussian likelihood. Note that we use Greek (Latin) characters to index observables (parameters).

In this work, the observables and parameters are given by

$$\begin{aligned} \vec{O} &= \{P_m(k_1), ..., P_m(k_A), \mathcal{H}(M_1), ..., \mathcal{H}(M_B), \\ &\qquad\qquad \mathcal{V}(R_1), ..., \mathcal{V}(R_D)\}, \\ \vec{\theta} &= \{\Omega_m, \Omega_b, h, n_s, \sigma_8, M_\nu\} \end{aligned}$$

respectively, where $P_m(k)$ is the matter power spectrum at wavenumber k, $\mathcal{H}(M)$ is the halo mass function at mass M, and $\mathcal{V}(R)$ is the void size function at radius R. Note there are a total of A, B, and D bins for the matter power spectrum, the halo mass function, and the void size function respectively, giving a total dimensionality of $A + B + D$.

We quantify the information content by considering the marginalized error on the cosmological parameters,

$$\sigma(\theta_i) \equiv \sqrt{(F^{-1})_{ii}} \ , \tag{3.3}$$

which is a lower bound.

Covariance matrix

We estimate the covariance matrix using the $N_{\rm cov} = 15{,}000$ simulations of the fiducial cosmology as

$$C_{\alpha\beta} = \langle (O_\alpha - \langle O_\alpha \rangle)\,(O_\beta - \langle O_\beta \rangle) \rangle \ , \tag{3.4}$$

where $\langle\rangle$ denotes the mean over simulations. This is the largest number of simulations used for covariance estimation to date. We have verified that our combined results are converged even with half of the simulations. We show the results of our convergence tests in the appendix for this chapter (Section 3.7).

Derivatives

For the cosmological parameters Ω_m, Ω_b, h, n_s, and σ_8, we approximate the derivatives using a central difference scheme centered on the fiducial cosmology,

$$\frac{\partial \vec{O}}{\partial \theta_i} \simeq \frac{\vec{O}(\theta_i + \delta\theta_i) - \vec{O}(\theta_i - \delta\theta_i)}{2\theta_i}. \tag{3.5}$$

Note that only the value of the i^{th} cosmological parameter is perturbed about its fiducial value, θ_i, while the values of all other parameters are held fixed. The error of this approximation is $\mathcal{O}(\delta\theta_i^2)$.

For neutrinos we cannot use Eq. 3.5 because the fiducial model has massless neutrinos, so $\vec{O}(\theta_i - \delta\theta_i)$ would correspond to a cosmology with negative neutrino mass. We thus compute the derivatives for neutrinos using a second-order forward difference scheme,

$$\frac{\partial \vec{O}}{\partial M_\nu} \simeq \frac{-3\vec{O}(M_\nu + 2\delta M_\nu) + 4\vec{O}(M_\nu + \delta M_\nu) - 3\vec{O}(M_\nu)}{2\delta M_\nu}, \tag{3.6}$$

which has error $\mathcal{O}(\delta M_\nu^2)$. We exclusively use the M_ν^{++} and M_ν^{+++} cosmologies in Eq. 3.6 throughout this work.

We use a total of $N_{\text{der}} = 1{,}000$ (500+500) simulations to compute derivatives when using Eq. 3.5, and 1,500 when using Eq. 3.6. In the appendix for this chapter (Section 3.7) we show that our results are robust and converged with this number of simulations. We also give evidence of robustness with respect to the choice of finite difference scheme for M_ν.

3.4 Cosmological probes

In this section we outline the cosmological observables considered in this work: the matter power spectrum, the halo mass function, and the void size function.

Matter power spectrum

The first observable we study is the matter power spectrum. For each realization, the density field is computed by depositing particle masses to a regular grid using the cloud-in-cell mass assignment scheme. In simulations with massive neutrinos we consider both CDM and neutrino particles when constructing the density field. The density contrast field, $\delta(\vec{x}) = \rho(\vec{x})/\bar{\rho} - 1$, is then Fourier transformed and the power spectrum is computing by averaging $|\delta(\vec{k})|^2$ over spherical bins in $|k|$. The size of each bin is equal to the fundamental frequency, $2\pi/L$, where $L = 1\,h^{-1}$Gpc is the simulation box size.

A grid with 1024^3 cells is used, which is large enough to avoid aliasing effects on the scales of interest for this work. In our analysis we consider wavenumbers up to $k_{\text{max}} = 0.5\ h\text{Mpc}^{-1}$, using 79 bins. This choice of k_{max} is based on the fact that the clustering of the simulations is converged at this scale for this mass resolution [see 383]. We will however show that using

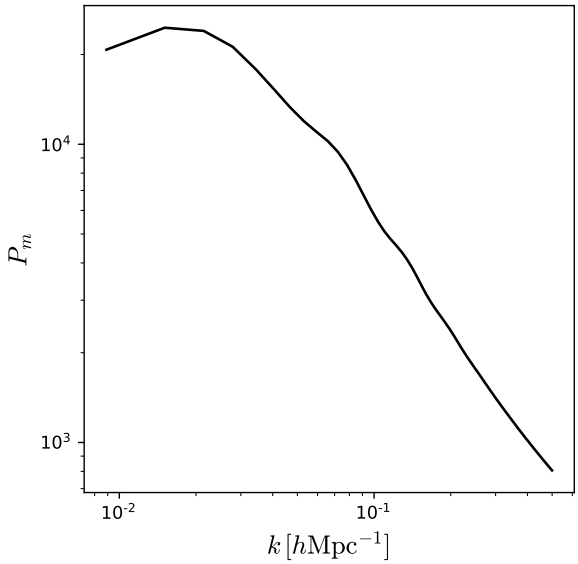

Figure 3.1: The matter power spectrum for the fiducial cosmology.

a larger $k_{\max}$ would likely lead to even tighter constraints than the ones we report. We show the power spectrum for the fiducial cosmology in Fig. 3.1.

Halo mass function

The second observable we consider is the halo mass function (HMF). Dark matter halos are identified using the Friends-of-Friends algorithm [135], with a linking length $b = 0.2$. The halo finder considers only the dark matter distribution, as the contribution of neutrinos to the total mass of a halo is expected to be negligible [376, 377, 202, 250].

The halo mass function is defined as the comoving number density of halos per unit of (log) halo mass, $dn/d\ln M$. The mass of a halo is estimated as

$$M = Nm_p, \tag{3.7}$$

where N is the number of dark matter particles in the halo and m_p is the mass of a single dark matter particle. Note that in the *Quijote* simulations, there are only dark matter and neutrino particles, i.e. dark matter particles represent the CDM+baryon fluid. The mass of a dark matter particle is thus normalized according to Ω_{cb}, such that

$$m_p = \frac{V\rho_c}{N_p}\Omega_{cb} = \frac{V\rho_c}{N_p}\left(\Omega_m - \frac{M_\nu}{93.14h^2}\right), \tag{3.8}$$

where $V = L^3$ is the simulation volume, N_p is the total number of dark matter particles in the simulation, and ρ_c is the Universe's critical energy density at $z = 0$. Thus $m_p = m_p(\Omega_m, M_\nu)$

is a cosmology dependent quantity, which induces noise when computing the derivatives of the HMF with respect to Ω_m or M_ν in a fixed *mass* bin. This is because it is the *number* of dark matter particles that is the fundamental constituent of the halo mass: a halo with a given number of particles will lie in the same *number* bin for all cosmologies, whereas it may lie in a different *mass* bin depending on the value of m_p. This noise can thus be avoided by instead working with bins of fixed particle *number* by considering the derivative of the comoving number density of halos per unit (log) number of particles, $dn/d\ln N$. One can then transform these derivatives in bins of fixed N to derivatives in bins of fixed M to obtain the derivatives of the halo mass function.

Using the shorthand $\mathcal{H}$ to denote the halo mass function, we now derive this transformation. In practice, one measures the halo mass function for a fixed cosmology, thus working in logarithmic bins gives

$$\mathcal{H} := \frac{dn}{d\ln M} = \frac{dn}{d\ln N}, \tag{3.9}$$

where it is understood that the derivative is taken with fixed cosmological parameters, $\vec{\theta}$. Explicitly, one can think of the halo mass function as a function of the cosmological parameters and halo mass, $\mathcal{H}(\vec{\theta}, M)$, or the cosmological parameters and number of particles, $\mathcal{H}(\vec{\theta}, N)$. Thus the derivative of the HMF with respect to one of the cosmological parameters, θ, while holding all other cosmological parameters, $\vec{\phi}$, fixed can be written as

$$\left(\frac{\partial\mathcal{H}}{\partial\theta}\right)_{\vec{\phi}} = \left(\frac{\partial\mathcal{H}}{\partial\theta}\right)_{M,\vec{\phi}} + \left(\frac{\partial\mathcal{H}}{\partial\ln M}\right)_{\vec{\theta}}\left(\frac{\partial\ln M}{\partial\theta}\right)_{\vec{\phi}}, \tag{3.10}$$

or

$$\left(\frac{\partial\mathcal{H}}{\partial\theta}\right)_{\vec{\phi}} = \left(\frac{\partial\mathcal{H}}{\partial\theta}\right)_{N,\vec{\phi}} + \left(\frac{\partial\mathcal{H}}{\partial\ln N}\right)_{\vec{\theta}}\left(\frac{\partial\ln N}{\partial\theta}\right)_{\vec{\phi}}. \tag{3.11}$$

Equating these two equations and rearranging gives

$$\begin{aligned}\left(\frac{\partial\mathcal{H}}{\partial\theta}\right)_{M,\vec{\phi}} &= \left(\frac{\partial\mathcal{H}}{\partial\theta}\right)_{N,\vec{\phi}} + \left(\frac{\partial\mathcal{H}}{\partial\ln N}\right)_{\vec{\theta}}\left[\frac{\partial\ln N}{\partial\theta} - \frac{\partial\ln M}{\partial\theta}\right]_{\vec{\phi}} \\ &= \left(\frac{\partial\mathcal{H}}{\partial\theta}\right)_{N,\vec{\phi}} - \left(\frac{\partial\mathcal{H}}{\partial\ln N}\right)_{\vec{\theta}}\left(\frac{\partial\ln m_p}{\partial\theta}\right)_{\vec{\phi}},\end{aligned} \tag{3.12}$$

where Eq. 3.7 was used in the final step.

The cosmology dependence of m_p takes effect in the final term of Eq. 3.12. There is only a difference between the fixed N and fixed M derivative of the HMF when m_p depends on θ, i.e., when $\theta \in \{\Omega_m, M_\nu\}$. Using Eq. 3.8, one finds that

$$\frac{\partial\ln m_p}{\partial\Omega_m} = \frac{1}{\Omega_m}, \tag{3.13}$$

$$\frac{\partial\ln m_p}{\partial M_\nu} = \frac{1}{\Omega_{cb}93.14h^2}, \tag{3.14}$$

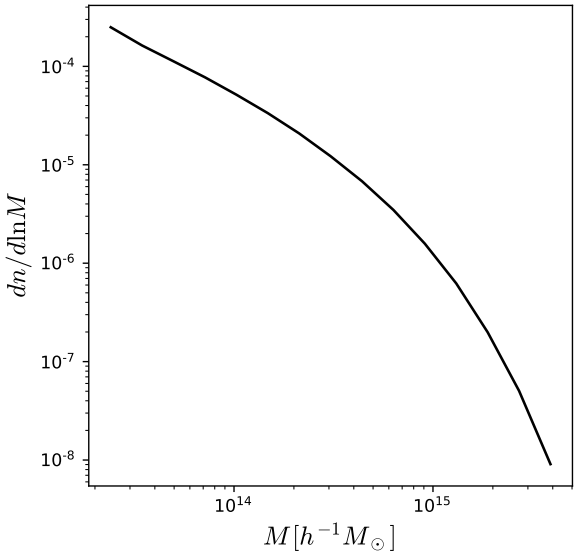

Figure 3.2: The halo mass function for the fiducial cosmology.

where it is understood that all cosmological parameters apart from the one in the derivative are held fixed at their fiducial values.

Thus our procedure to compute derivatives of the HMF using Eq. 3.12 is as follows. We first bin the number of halos according to the number of dark matter particles they contain. We then compute the derivatives for each fixed-N bin using the equations from Section 3.3, yielding the first term on the right-hand side of Eq. 3.12. This will be sufficient for all cosmological parameters except for Ω_m and M_ν, as these require a correction term to transform to fixed-M bins due to the variation of m_p. The $\partial\mathcal{H}/\partial \ln N$ term can be computed via spline interpolation or by using finite difference methods between the bins of the halo mass function of the fiducial cosmology. We have confirmed the stability of both approaches. Finally, the derivative of $\ln m_p$ with respect to θ is computed using Eqs. 3.13 and 3.14 evaluated at the fiducial values.

We consider halos with a number of dark matter particles between 30 and 7,000, using 15 logarithmically spaced bins. The corresponding halo mass range is approximately 2.0×10^{13} to 4.6×10^{15} $h^{-1}M_\odot$. As with the matter power spectrum, this choice of binning and cuts is made to ensure convergence of the derivatives based on the resolution and number of the simulations available. Hence, using more bins and/or a larger mass range would likely lead to stronger constraints than we report. We show the HMF for the fiducial cosmology in Fig. 3.2.

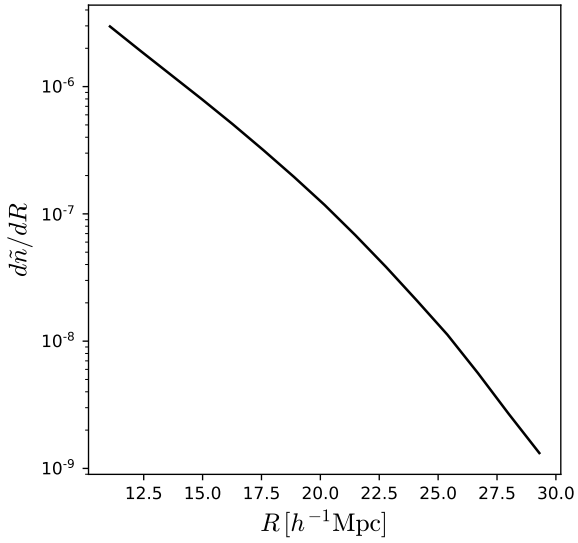

Figure 3.3: The void size function for the fiducial cosmology.

Void size function

We identify voids in the underlying matter field using a spherical void finding algorithm developed by [47], which we now outline. We use a grid of resolution 768^3 to look for voids — this is slightly finer than the CDM grid resolution of 512^3 to enable detection of small voids. The density contrast field is then smoothed with a top-hat filter over a large-scale, $R = 53.4\,h^{-1}$Mpc, which is a multiple of the grid spacing and is chosen to be bigger than the size of the largest void. Next, minima that are smaller than the threshold $\delta_{\rm th} = -0.7$ in the smoothed field are considered as voids with radius R, unless they overlap with existing voids. This procedure is then performed iteratively while decrementing R by the grid spacing. In this work we use a threshold of $\delta_{\rm th} = -0.7$, but have checked that results are similar for $\delta_{\rm th} = -0.5$.

The void size function (VSF) is then computed as the comoving number density of voids per unit of radius, denoted $d\tilde{n}/dR$. Unlike the halo mass function, the VSF is not prone to the changes in particle mass, since the void finder operates directly in the same unit as the VSF. The range of void sizes is limited by our resolution and the size of our simulated volume. Having found the voids, we apply radius cuts of $R_{\rm min} = 10.4$ and $R_{\rm max} = 29.9$ h^{-1}Mpc, corresponding to 15 bins linear in R. As with the matter power spectrum and the halo mass function, this choice of binning and cuts is made to ensure convergence of the derivatives based on the resolution and number of the simulations available. Hence, using more bins and/or a larger range of void sizes may lead to stronger constraints than we report. We show the VSF for the fiducial cosmology in Fig. 3.3.

Investigation of the void size function, and void abundances, is a rich field that has shown

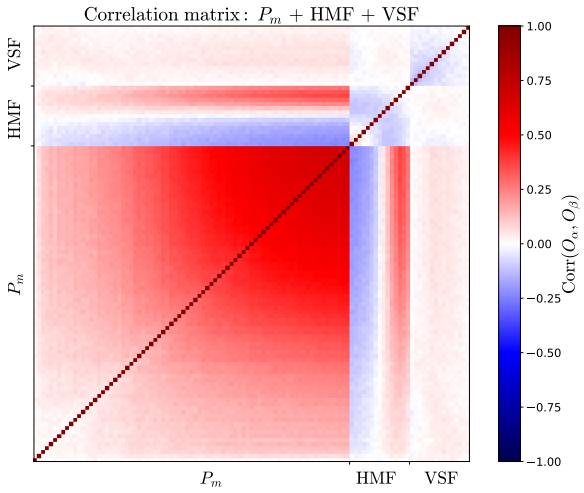

Figure 3.4: Correlation matrix for the matter power spectrum (P_m, with 72 linear bins and $k_{\rm max} = 0.5\ h{\rm Mpc}^{-1}$), the halo mass function (HMF, 15 log bins between 2.0×10^{13} and $4.6 \times 10^{16}\ h^{-1}M_\odot$), and the void size function (VSF, 15 linear bins between 10.4 and 29.9 h^{-1}Mpc), from bottom left to top right. Bin values increase from left to right for each probe. While the HMF shows clear off-block correlation with P_m, the VSF is somewhat independent from both P_m and the HMF.

promising theoretical work to match mocks [see, e.g. 294, 80, 349, 208, 288, 277, 310, 118, 374].

3.5 Results

In this section we present the main results of this work.

Full covariance of the probes

In Fig. 3.4 we show the correlation matrix, defined as $\text{Corr}(O_\alpha, O_\beta) := C_{\alpha\beta}/\sqrt{C_{\alpha\alpha}C_{\beta\beta}}$, where $C_{\alpha\beta}$ is the covariance matrix (Eq. 3.4). First we discuss the correlations for each individual probe (auto-correlations). For the matter power spectrum (bottom-left region of Fig. 3.4), we observe some well-known structures: the covariance is almost diagonal on large scales, while mode-coupling induces significant off-diagonal correlations on small scales. For the halo mass function (central region of Fig. 3.4), the covariance matrix is almost diagonal, with some small correlations between the different mass bins; the correlations are negative for heavy halos, but are positive for the lightest halos considered in this work. The covariance of the void size function (top-right region of Fig. 3.4) is also almost diagonal, with the abundance of different void sizes slightly anti-correlated with nearby bins due to conservation of volume.

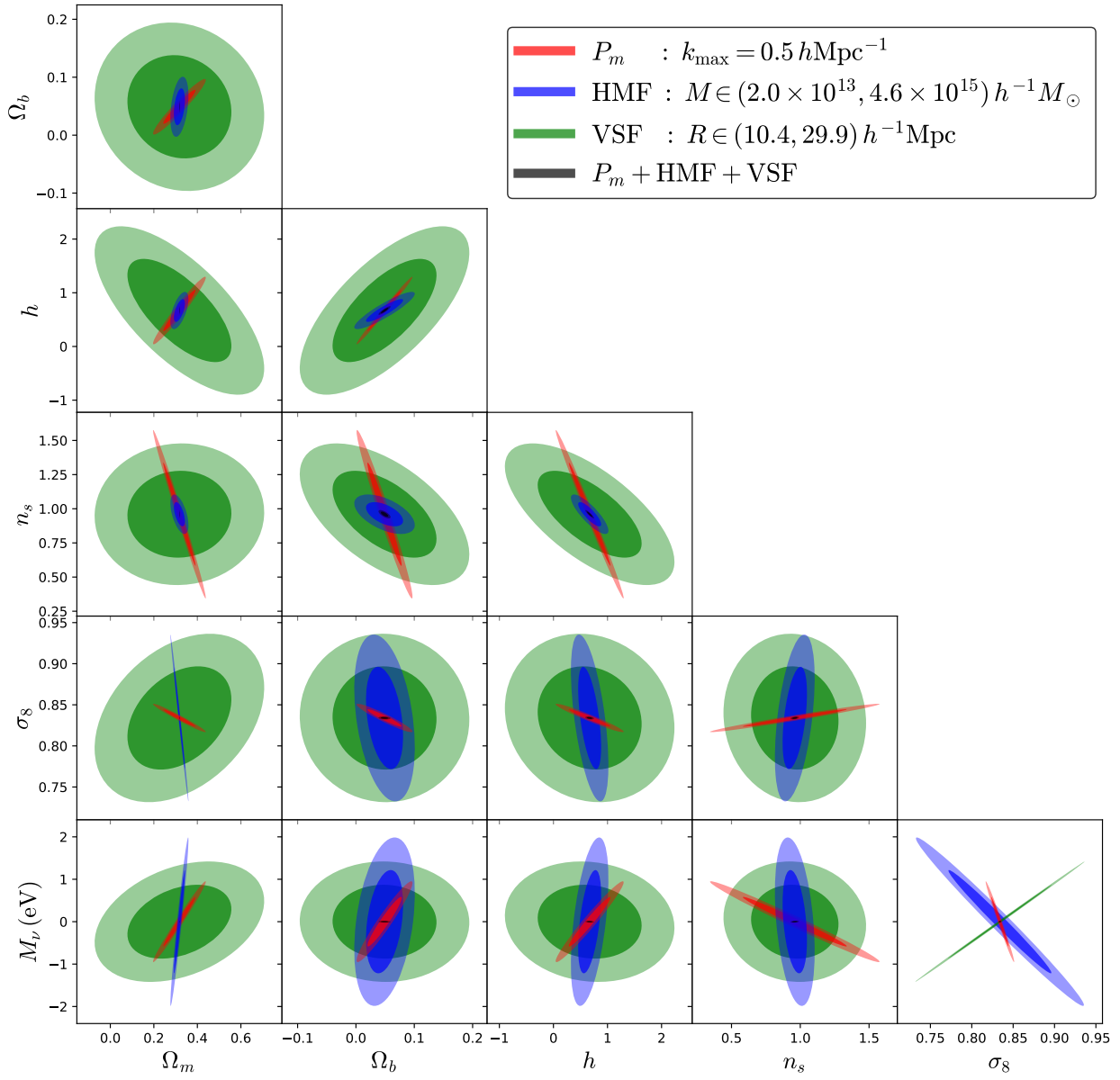

Figure 3.5: 68% (darker shades) and 95% (lighter shades) confidence contours for the cosmological parameters for the non-linear matter power spectrum (P_m, red), the halo mass function (HMF, blue), and the void size function (VSF, green). Due to the often different degeneracies of each probe, we obtain significantly tighter constraints when combining the three probes (black). We note that some contours extend into unphysical regions ($\Omega_b < 0, h < 0, M_\nu < 0$): this is just a result of the Gaussian approximation associated with a Fisher analysis.

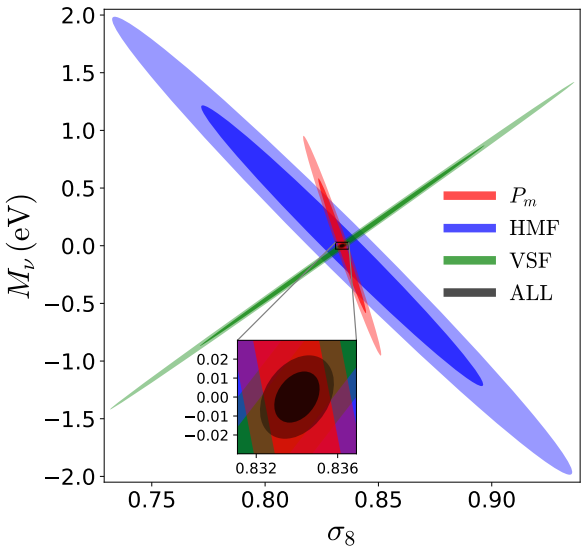

Figure 3.6: The M_ν–σ_8 plane from Fig. 3.5. We inset a zoom-in of the contour obtained by combining all three probes. The marginalized error on M_ν from P_m alone is 0.77eV, while the error after combining all three probes is 0.018eV, corresponding to a factor ~ 43 improvement.

Next, we consider the correlations between different probes (cross-correlations). The halo mass function shows an interesting correlation pattern with the matter power spectrum: the abundance of the more (less) massive halos shows a $\sim 20\%$ correlation (anti-correlation) with small scales of the matter power spectrum. Similar trends are seen between halos and large scales of the matter power spectrum, albeit at a weaker level. On the other hand, voids can be seen to be somewhat independent of both the matter power spectrum and halos, as their cross-correlation is $\lesssim 5\%$ for all scales and masses.

As discussed in Section 3.3, we combine the covariance matrix with the numerically computed derivatives to calculate the Fisher matrix. The numerical derivatives and related numerical convergence tests are shown in the appendix for this chapter (Section 3.7).

Cosmological constraints

We show the two-dimensional (2D) 68% and 95% confidence intervals obtained from our Fisher analysis for each individual probe, and the combination of all probes, in Fig. 3.5. The constraints on the parameters are not generally tight when considering any of three probes alone, because we adopt a conservative survey volume of 1 $(h^{-1}\text{Gpc})^3$, which is significantly smaller than what is achievable by DESI, $\sim 10^2\,(h^{-1}\text{Gpc})^3$.

The three probes show different degeneracies and are sensitive to each parameter at different levels. For example, the halo mass function provides a relatively tight constraint on Ω_m when compared to the other two probes, as the halo mass function depends non-linearly

on and is highly sensitive to Ω_m [see, e.g. 179]. The void size function provides weaker constraints than the other two probes on almost all parameters, except for n_s compared to P_m. Naively, this is not surprising, considering the relatively smaller range of scales being probed by the void size function compared to the matter power spectrum. More information could probably be retrieved by using other void-related observables, such as the void-matter correlation function.

Because the degeneracies between parameters are often very different for each probe, it is expected that combining the probes will break the degeneracies and in turn yield significantly tighter constraints on the cosmological parameters than the individual probes do. Indeed, the black ellipses in Fig. 3.5 show the tight constraints obtained by combining the three probes. We emphasize that these constraints account for all the correlations between the different observables, i.e. by using the full covariance matrix of Fig. 3.4.

The benefit of combining the three probes is particularly well demonstrated in the M_ν–σ_8 plane. Because the combined constraints are too small to be visible in Fig. 3.5, we zoom in on this plane in Fig. 3.6. We find that, despite not being as powerful tools as P_m in constraining M_ν, the HMF and VSF both show degeneracies in different directions from that of P_m, which guarantees that constraints on the neutrino masses will be largely reduced by combining the three probes. In turn this helps break the well-known M_ν–σ_8 degeneracy for the matter power spectrum. We note that the area of these confidence contours, particularly for the HMF, can potentially be reduced by increasing the bin boundaries and/or by fine-tuning the binning schemes. Our choice of binning is restricted by our simulation resolution. We leave these investigations to future works.

For a direct comparison to the usual constraints expected from the matter power spectrum, we show the 1D marginalized errors (Eq. 3.3) from different combinations of the probes with P_m in Fig. 3.7. We study how the errors vary with the cutoff scale $k_{\rm max}$. Combining P_m with either the HMF, VSF, or both, can achieve a significant level of improvement on all 6 parameters. The combination with the HMF is typically more beneficial than the combination with the VSF. The only exception is for M_ν, where the VSF is the better probe to combine with P_m.

While the constraints from P_m alone saturate at around $k_{\rm max} = 0.2\,h{\rm Mpc}^{-1}$ for all parameters, the combined constraints for M_ν (and Ω_m) continue to improve beyond $k_{\rm max} = 0.5\,h{\rm Mpc}^{-1}$. This can be explained by the breaking of degeneracies when combining probes. It was shown in Fig. 5 of [383] that increasing $k_{\rm max}$ beyond $0.2\,h{\rm Mpc}^{-1}$ leads to a squeezing along the semi-minor axes (i.e, the most constraining direction) for the P_m ellipses. While this squeezing has little effect on the marginalized error on M_ν from P_m alone, its effects are manifest when combined with other probes with misaligned contours, resulting in significant tightening of constraints. Even though the numerical resolution of the *Quijote* simulations prevent us from confidently investigating beyond $k_{\rm max} = 0.5\,h{\rm Mpc}^{-1}$, our results hint that even tighter constraints could be achieved by including smaller scales.

In Table 3.2 we list the errors for $k_{\rm max} = 0.5h{\rm Mpc}^{-1}$ using different probe combinations. We list the constraints obtained by combining all three probes while (1) only using the diagonals of the covariance matrix (diag), (2) only considering auto-covariance (auto), and

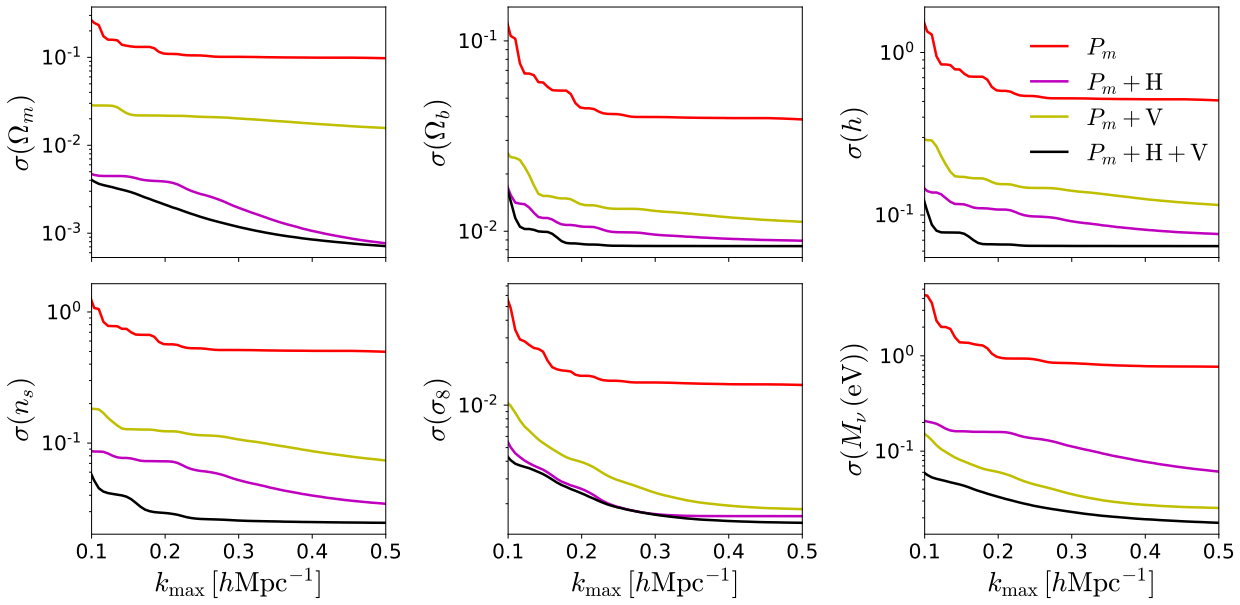

Figure 3.7: The 1D marginalized error for each of the cosmological parameters as a function of $k_{\max}$. We consider 4 scenarios: P_m alone (red), P_m + HMF (magenta), P_m + VSF (yellow), and P_m + HMF + VSF (black). While the constraints from P_m alone saturate at $k_{\max} \simeq 0.2\,h\mathrm{Mpc}^{-1}$, the combined constraints for M_ν (and Ω_m) continue to improve until $k_{\max} = 0.5\,h\mathrm{Mpc}^{-1}$, and likely beyond.

(3) considering the full covariance (full). We find that using only the diagonal components of the covariance matrix, effectively ignoring both the correlation between the probes and between different bins of the same probe, leads to a factor of 1.7 increase on the error on the neutrino mass. Using only block cross-correlations, i.e. ignoring the correlation between the probes, leads to a factor of 1.2 increase on the error on the neutrino mass. Therefore, to obtain the tightest constraints, it is crucial to model the full covariance matrix. It is interesting to note that when considering the matter power spectrum alone, correlations cause an increase in errors due to the *positive* correlation between different scales (see Fig. 3.4). However, it is the complex correlation structure, notably the anti-correlations, introduced by considering the HMF and VSF that leads to a reduction in error, both for the HMF and VSF individually, and in turn when combining all probes. The association of anti-correlation with the tightening of constraints was also pointed out by [101].

In Table 3.2 we quantify the improvement of the combined constraints compared to those achieved from P_m alone. We find the improvements to be a factor of 137, 5, 8, 20, 10, and 43, for Ω_m, Ω_b, h, n_s, σ_8, and M_ν, respectively. Thus we achieve 43 times tighter constraints on neutrino mass by combining all three probes. Specifically, the marginalized errors on M_ν are 0.77eV (P_m alone) and 0.018eV (P_m+HMF+VSF). We provide an additional plot in the

Marginalized Fisher Constraints						
Probe(s)	Ω_m	Ω_b	h	n_s	σ_8	M_ν(eV)
P_m	0.098	0.039	0.51	0.50	0.014	0.77
HMF	0.034	0.042	0.28	0.12	0.082	1.6
VSF	0.31	0.12	1.3	0.42	0.083	1.1
P_m + HMF	0.00077	0.0089	0.076	0.034	0.0016	0.061
P_m + VSF	0.016	0.011	0.12	0.074	0.0018	0.025
HMF + VSF	0.0063	0.037	0.23	0.10	0.0069	0.096
P_m + HMF + VSF (diag)	0.0015	0.0088	0.066	0.028	0.00061	0.031
P_m + HMF + VSF (auto)	0.0015	0.0086	0.071	0.033	0.0016	0.025
P_m + HMF + VSF (full)	0.00071	0.0084	0.064	0.025	0.0015	**0.018**
Multiplicative improvement	137	5	8	20	10	43

Table 3.2: Marginalized errors of cosmological parameters for $k_{\rm max} = 0.5h{\rm Mpc}^{-1}$ using different probe combinations. Note, we list the constraints obtained by combining all 3 probes while: 1) only using the diagonals of the covariance matrix (diag), 2) only considering auto-covariance (auto), and 3) considering the full covariance (full). We highlight in bold the full constraints on the sum of the neutrino masses. We also list the multiplicative improvement in the constraints from the full covariance compared to those from P_m alone.

appendix for this chapter (Section 3.8) to show the confidence ellipses when combining only two of the probes at a time.

3.6 Discussion and Conclusions

Upcoming galaxy surveys will map large volumes of the Universe at low redshifts, with the potential to drastically improve our understanding of the underlying cosmological model. With the unprecedentedly precision achievable by these surveys, it is expected that a very large amount of cosmological (and astrophysical) information will lie in the mildly to fully non-linear regime, where analytic methods are often intractable. It remains an open question which observable(s) will lead to the tightest bounds on the cosmological parameters.

In this paper, we use the *Quijote* simulations, based on the Fisher formalism, to quantify the information content embedded in the non-linear matter power spectrum, the halo mass function, and the void size function, both individually and when combined, at $z = 0$. We find that the HMF and VSF have different degeneracies to each other and to the matter power spectrum, particularly in the M_ν–σ_8 plane (Figs. 3.5 & 3.6). In terms of measuring neutrino mass, we find the void size function to be the more complementary probe to combine with the matter power spectrum. This is consistent with findings that void properties are particularly sensitive to matter components that are less clustered, such as neutrinos [258, 222].

By combining the non-linear matter power spectrum ($k_{\rm max} = 0.5\ h{\rm Mpc}^{-1}$), with the halo mass function ($M \gtrsim 2 \times 10^{13}\ h^{-1}M_\odot$), and the void size function ($R \geqslant 10.4 h^{-1}{\rm Mpc}$), we achieve significantly tighter constraints on the cosmological parameters compared to P_m alone (Fig. 3.7). In particular, we find that with a volume of just 1 $(h^{-1}{\rm Gpc})^3$, the error on the sum of neutrino masses from the combined probes is at the 0.018eV level, compared to 0.77eV from the matter power spectrum alone — a factor of 43 improvement. We emphasize that this value mainly demonstrates the information content in the late-time statistics, and they are not forecasts for any particular survey.

Also of particular interest is the factor 137 improvement in the error on Ω_m. This is driven by the information in the HMF, and gives a marginalized error of $\sigma(\Omega_m) = 7.1 \times 10^{-4}$, which is almost 100 times smaller than the error obtained from a joint large-scale structure analysis by DES Y1 [$\sigma(\Omega_m) \approx 0.04$, 362], and 8 times smaller than Planck 2018, [$\sigma(\Omega_m) \approx 5.6 \times 10^{-3}$ (TT,TE,EE+lowE+lensing+BAO), 292]. In addition, we found $\sigma(h) = 0.064$ by combining the three probes, which is 8 times tighter than the constraints from the matter power spectrum alone. This could provide a new angle to investigate the Hubble tension.

There are several caveats in this work. Firstly, we assumed perfect knowledge of the three-dimensional spatial distribution of the underlying matter field in real-space. However, in reality, one observes either tracers of the matter field in redshift-space, or the projected matter field through lensing. Therefore, additional links must be made to bridge the galaxy–matter connection and the 2D lensing–3D matter distribution gaps. This effect is also relevant for voids: in this work we considered voids in the 3D matter field, which is not something current surveys are able to observe directly. Detecting voids in the matter field from photometric (2D lensing) data has been considered in works such as [295, 132]. Alternatively, one can measure voids in the 3D halo field [see, e.g. 271, 118]. If we were to instead have considered voids in the 3D CDM field, the combined error on M_ν slightly degrades to 0.025eV. However, considering voids in the CDM field versus halo field can lead to non-trivial differences in void properties, which might increase or decrease constraints [222]. We will consider voids in the halo field in a future work.

A further note regarding voids is that there are various conventions when it comes to defining voids [see, e.g. 293, 350]. It would thus be interesting further work to consider how the choice of void finder impacts constraints. A different void finder may be able to extract additional information compared to the spherical void finder applied here.

Another limitation of this work is that our simulations consider only gravitational interactions and hence ignore baryonic effects which can impact the small-scale matter distribution. This is particularly relevant for both clustering and halos [see, e.g. 381, 126, 138, and references therein], while it is expected that baryons have a lower impact on voids [277]. Furthermore, halo clustering is influenced by various properties, such as spin, concentration, and velocity anisotropy, which have not been considered in this work [see, e.g. 386, 167, 152, 229, 228, 282, 336].

Additionally, we have neglected super-sample covariance [352, 240], which could modify the errors reported in this work.

We also note that the constraints obtained here may be overly conservative due to the

limited number and resolution of simulations available. Firstly, this means that the number of bins used are likely suboptimal. Second, applying more aggressive bounds on the observables, e.g. a higher $k_{\rm max}$, a larger halo mass range, or a larger void size range, would likely also reduce the combined constraints. Third, we only considered a single redshift, $z = 0$: in practice, surveys measure $z > 0$ where the universe is more linear and the constraints will thus be weaker, however, combining multiple redshifts could tighten the constraints as found in works such as [246]. Fourth, we considered a volume of only $1\,(h^{-1}{\rm Gpc})^3$, whereas surveys such as Euclid and DESI will cover volumes of around $10^2\,(h^{-1}{\rm Gpc})^3$, so, conservatively, the error on the parameters will shrink by a factor of $1/\sqrt{10^2} = 0.1$. Fifth, we have only considered three probes; using the same observations, one can derive other statistics such as the bispectrum, void profile, and BAO, which could be combined with the statistics considered here to further break degeneracies. Finally, considering redshift space distortions would also tighten constraints as neutrinos are distinguishable from CDM via their higher thermal velocity.

We have demonstrated that combining multiple probes of cosmological structure using their full covariance matrix provides remarkably tight constraints on the cosmological parameters, and helps extract much additional information from small scales. In particular, we have shown that there is, in principle, sufficient information to measure the sum of the neutrino masses at the minimum mass of 0.06 eV. Our results are in good agreement with [310] who found that combining halo and void abundances can yield $\mathcal{O}(0.01\,{\rm eV})$ constraints on the neutrino mass. This approach opens a promising pathway to measure neutrino mass, potentially without relying on CMB-based measurements which require accurate knowledge of the optical depth, τ. In addition, comparing constraints from different combinations of observables, e.g., CMB+P_m and P_m+HMF+VSF, will help identify systematic issues and provide robust evidence for any discovery. We thus hope our work will motivate galaxy survey collaborations to build the simulations and analytic tools necessary to implement this approach on upcoming observational data.

Acknowledgements

We thank Ravi Sheth for useful conversations on the early stages of this project, and Alice Pisani for fruitful discussion regarding voids. The *Quijote* simulations can be found at https://github.com/franciscovillaescusa/Quijote-simulations. The analysis of the simulations has made use of the *Pylians* libraries, publicly available at https://github.com/franciscovillaescusa/Pylians3. The work of D.N.S., B.D.W., and S.H. is supported by the Simons Foundation. L.V. acknowledges support from the European Union Horizon 2020 research and innovation program ERC (BePreSySe, grant agreement 725327) and MDM-2014-0369 of ICCUB (Unidad de Excelencia Maria de Maeztu). M.V. is supported by INFN PD51 Indark grant and by the agreement ASI-INAF n.2017-14-H.0.

3.7 Appendix A: Robustness of results to numerical systematics

In this section we verify the stability of our results to reduction in the number of simulations used to compute the covariance matrix and derivatives. In Fig. 3.8 we show the derivatives of the matter power spectrum (top), halo mass function (middle), and void size function (bottom) with respect to the cosmological parameters when using a different number of realizations. For the matter power spectrum, the derivatives are already converged when the mean values for each model are computed with 300 realizations. Results are slightly noisier for the halo mass function and the void size function, but still sufficiently converged by 500 realizations.

Next, we comment on the convergence of our simulated results with theory. The convergence of the matter power spectrum in *Quijote* has been thoroughly tested [383, 30, 178]. For the void size function there is no theoretical formula accurate enough to compute derivatives, but we have checked results are robust to the parameters used in the void finder. Therefore, we only compare our measured HMF to theoretical predictions. For the HMF, we plot the theoretical predictions of Sheth-Tormen (ST) [335, 334] and Tinker [361]. We use the prescription of [119] in the case of massive neutrino cosmologies by replacing $\Omega_m \to \Omega_{cb}$ and $P_m \to P_{cb}$ as neutrinos have negligible contribution to halo mass. There is good agreements between these predictions and *Quijote*. We have also checked that there is good agreement for different choice of step size (not shown). Note that these theoretical formulae provide a guideline rather than exact predictions, as they were fitted to simpler simulations or calibrated on spherical overdensity halos, as opposed to FoF here.

In Fig. 3.9 we show the convergence of the Fisher matrix elements with respect to the number of realizations used to compute the covariance, $N_{\rm cov}$, and derivatives $N_{\rm der}$. We consider the Fisher matrix components for P_m (red), the HMF (blue), the VSF (green), and the combined probes (black). The gray bands corresponds to the $\pm 5\%$ interval. While there is some noise in the σ_8 component of the Fisher matrix for P_m as function of $N_{\rm cov}$, good convergence is achieved by 15000. Likewise the Fisher matrix is well converged as a function of $N_{\rm der}$. Crucially, the Fisher matrix elements for the combined probes (black) all show good convergence. Note that when combining probes we scale the power spectrum by a factor of 10^{-10} to ensure that the condition number of the covariance matrix is sufficiently low for accurate inversion.

Finally, we comment on the choice of finite difference scheme used to compute the derivative of probes with respect to M_ν. Throughout the paper we used Eq. 3.6 with $\delta M_\nu = 0.2\,{\rm eV}$, thus making use of simulations with $M_\nu = 0, 0.2$, and 0.4 eV. Using this scheme we found the full combined constraint on M_ν is 0.018 eV, as shown in Table 3.2. To illustrate robustness to this choice of finite difference scheme, we also performed the analysis using Eq. 3.6 with $\delta M_\nu = 0.1\,{\rm eV}$ and found it to give an identical constraint of 0.018 eV. Additionally, we tried a forward difference scheme between $M_\nu = 0$ and 0.1 eV, which also gave identical constraints. Hence, the results are consistent with the choice of finite difference scheme.

We do also note that since the joint constraints on the parameters given in Table 3.2 are smaller than the step sizes used to compute derivatives, it would be interesting to investigate the effect of smaller step sizes on the joint constraints. This would reduce the error in the numerical derivatives, and thus may slightly modify the joint constraints.

Given these results, we believe that our conclusions are robust against potential numerical systematics. We note again that our bin configuration has been chosen with these results in mind, to ensure sufficiently converged derivatives and Fisher matrix components, but in principle one could consider more bins over a wider range to potentially obtain tighter constraints.

3.8 Appendix B: Combining two probes at a time

Fig. 3.10 shows the 2D Fisher contours to illustrate the effects of only combining two out of the three probes at a time. In most cases, the constraints obtained by combining the halo mass function with the void size function are the weakest, indicating that it is important to use the information from the non-linear matter power spectrum to break degeneracies.

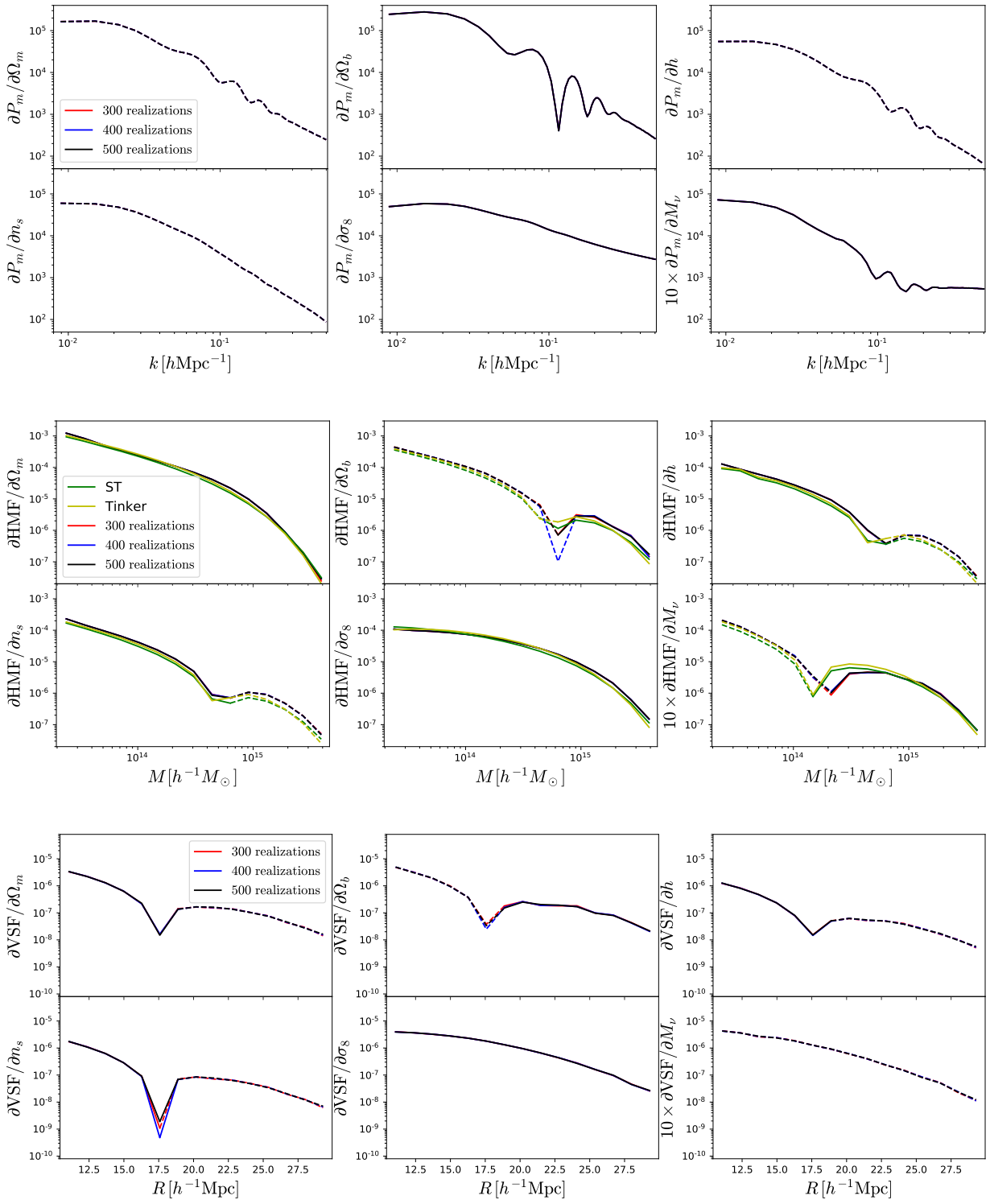

Figure 3.8: Derivatives of the matter power spectrum (top), halo mass function (middle), and void size function (bottom) with respect to the different cosmological parameters at $z = 0$. We show results when the mean values are estimated using 300 (red), 400 (blue), and 500 realizations (black). Solid/dashed lines indicate that the value of the derivative is positive/negative. While the derivatives for the matter power spectrum are well converged already with 300 realizations, more simulations are required for halos and voids.

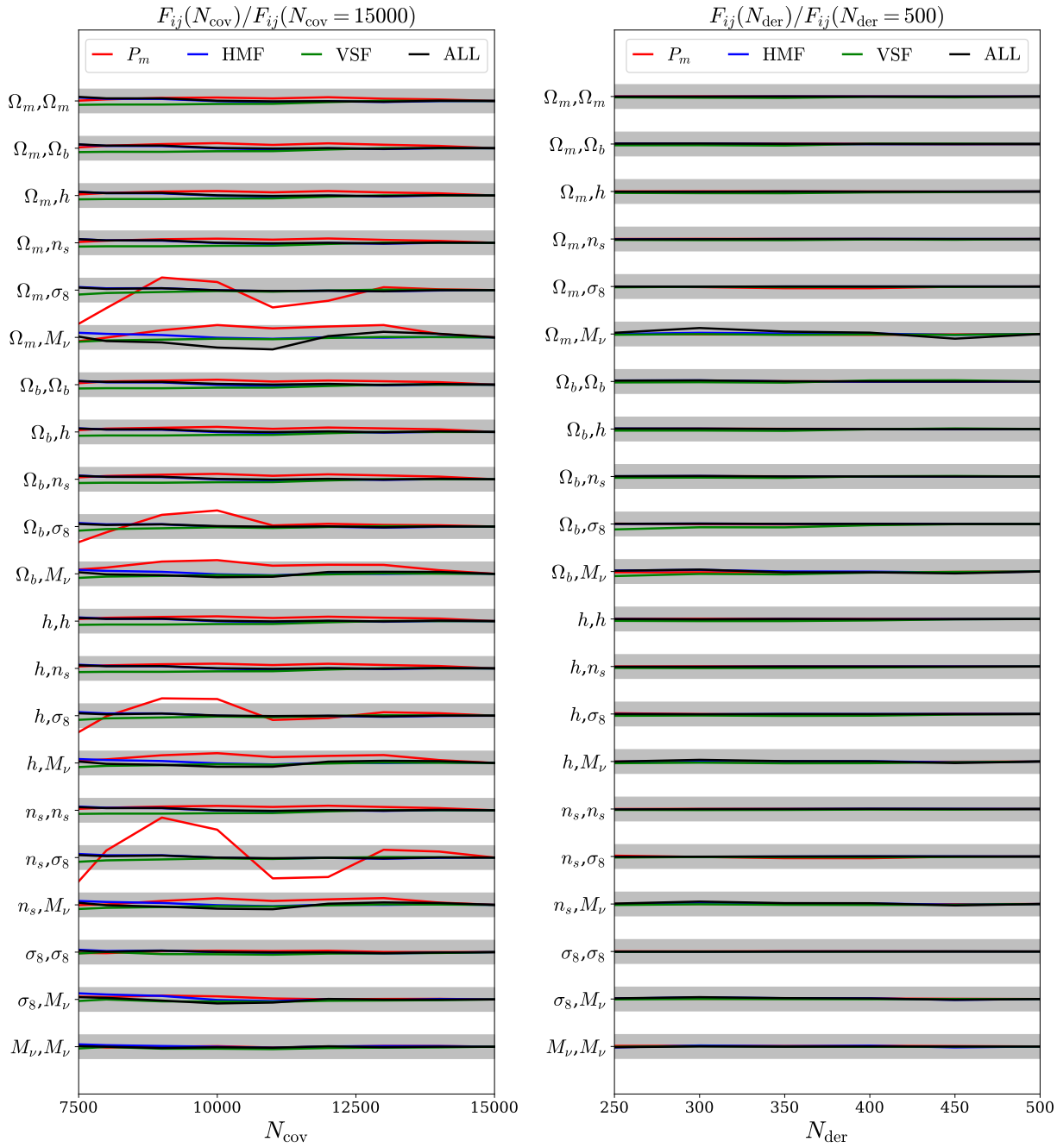

Figure 3.9: Left: Convergence of all Fisher matrix components as a function of number of simulations used to compute the covariance matrix, $N_{\rm cov}$. Each line shows the ratio between the Fisher matrix elements computed using $N_{\rm cov}$ simulations and 15,000 simulations (as used in the paper). Right: Convergence of all Fisher matrix components as a function of number of simulations used to compute derivatives, $N_{\rm der}$. Each line shows the ratio between the Fisher matrix elements computed using $N_{\rm der}$ simulations and 500 simulations for each cosmology (as used in the paper). In both cases, we plot the Fisher matrix components for P_m (red), the HMF (blue), the VSF (green), and the combined probes (black). The gray bands correspond to the $\pm 5\%$ interval. While there is some noise in the σ_8 component of the Fisher matrix for P_m as a function of $N_{\rm cov}$, good convergence is achieved by 15,000. Likewise the Fisher matrix is well converged as a function of $N_{\rm der}$. Crucially, the Fisher matrix elements for the combined probes (black) all show good convergence.

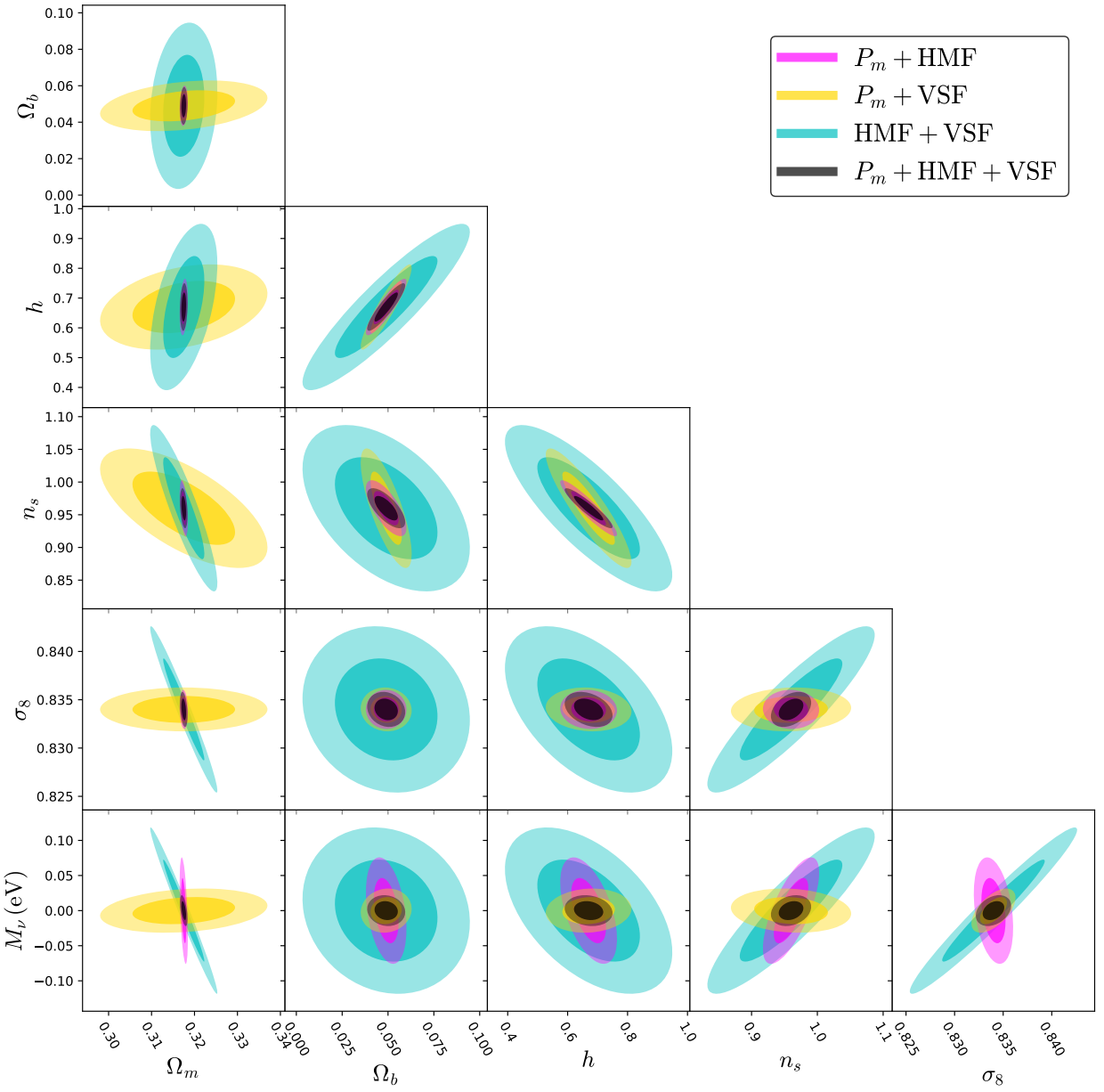

Figure 3.10: Same as Fig. 3.5 but for pair combinations of the probes: power spectrum + halo mass function (magenta), power spectrum + voids size function (yellow), halo mass function + void size function (cyan) and power spectrum + halo mass function + void size function (black).

Chapter 4

Beware of fake νs: The effect of massive neutrinos on the nonlinear evolution of cosmic structure

The contents of this chapter was originally published in [58],

> Beware of Fake $\boldsymbol{\nu}$'s: The Effect of Massive Neutrinos on the Non-Linear Evolution of Cosmic Structure
> *Bayer A.E., Banerjee A., Seljak U.* (arXiv:2108.04215) PRD 105, 123510 (2022)

Massive neutrinos suppress the growth of cosmic structure on small, non-linear, scales. It is thus often proposed that using statistics beyond the power spectrum can tighten constraints on the neutrino mass by extracting additional information from these non-linear scales. In this chapter, we study the information content regarding neutrino mass at the field level, quantifying how much of this information arises from the difference in non-linear evolution between a cosmology with 1 fluid (CDM) and 2 fluids (CDM + neutrinos). We do so by running two N-body simulations, one with and one without massive neutrinos; both with the same phases, and matching their linear power spectrum at a given, low, redshift. This effectively isolates the information encoded in the linear initial conditions from the non-linear cosmic evolution. We demonstrate that for $k \lesssim 1\,h/\mathrm{Mpc}$, and for a single redshift, there is negligible difference in the real-space CDM field between the two simulations. This suggests that all the information regarding neutrino mass is in the linear power spectrum set by the initial conditions. Thus any probe based on the CDM field alone will have negligible constraining power beyond that which exists at the linear level over the same range of scales. Consequently, any probe based on the halo field will contain little information beyond the linear power. We find similar results for the matter field responsible for weak lensing. We also demonstrate that there may be much information beyond the power spectrum in the 3d matter field, however, this is not observable in modern surveys via dark matter halos or weak lensing. Finally, we show that there is additional information to be found in redshift space.

4.1 Introduction

Upcoming cosmological missions such as DESI[1], Euclid[2], LSST[3], PFS[4], SKA[5], and WFIRST[6], will probe progressively smaller scales of cosmic structure. It is hoped that by probing these small, non-linear, scales one will be able to detect much information regarding the total neutrino mass. To fully realize the potential of these surveys, an urgent task is thus to quantify and optimally extract this information from the observed cosmological fields.

In a cosmology with massive neutrinos [236], we can define ρ_{cb} as the contribution to the energy density due to cold dark matter (CDM) and baryons, ρ_ν as the contribution due to neutrinos, and ρ_m as the total matter contribution. Given the lower bound on the sum of the neutrino masses coming from oscillation experiments is $M_\nu = 60\text{meV}$ [166, 9, 25, 10, 22], neutrinos are non-relativistic at low redshift. Defining $\bar{\rho}_X$ as the mean energy density in species X, where $X = \{cb, \nu, m\}$, we can further define the relative overdensity of species X at redshift 0 as $\delta_X = (\rho_X - \bar{\rho}_X)/\bar{\rho}_X$, and the fraction of the total matter density in species X as $f_X = \bar{\rho}_X/\bar{\rho}_m = \Omega_X/\Omega_m$. This gives

$$\Omega_m \delta_m = \Omega_{cb}\delta_{cb} + \Omega_\nu \delta_\nu, \tag{4.1}$$

and the matter overdensity as

$$\delta_m = (1 - f_\nu)\delta_{cb} + f_\nu \delta_\nu. \tag{4.2}$$

In practice, we cannot measure δ_ν directly, as we do not have direct access to fluctuations in the cosmic neutrino background. We also cannot measure $f_\nu = \Omega_\nu/\Omega_m$ directly at low redshifts from the redshift-distance relations, since neutrinos are non-relativistic and their density has the same redshift dependence as cold dark matter and baryons. This leaves density perturbations in the total matter, δ_m, and CDM+baryon, δ_{cb}, fields as ways to probe neutrino mass at low redshifts. So the success of upcoming surveys measuring neutrino mass hinges on their ability to measure the effects of neutrinos on the total matter and CDM+baryon perturbations, as well as on their ability to measure Ω_m from the redshift-distance relation (which can also be extracted from perturbations, such as from Baryonic Acoustic Oscillations).

On large scales neutrinos cluster analogously to CDM, whereas on small scales they do not cluster. The scale at which this transition occurs is known as the free streaming scale and is due to the neutrino thermal velocities erasing their own perturbations. We can thus divide perturbations into scales larger than the neutrino free streaming scale, where $\delta_\nu \sim \delta_{cb}$, and scales smaller than that, where $\delta_\nu \sim 0$. One can see that if one could measure δ_m and δ_{cb}

[1] https://www.desi.lbl.gov
[2] https://www.euclid-ec.org
[3] https://www.lsst.org
[4] https://pfs.ipmu.jp/index.html
[5] https://www.skatelescope.org
[6] https://wfirst.gsfc.nasa.gov/index.html

on small scales in the absence of noise, then any difference between the two would give strong constraints on neutrino mass via $\delta_m \sim (1 - f_\nu)\delta_{cb}$. However, this poses several observational difficulties.

A first difficulty is that the matter overdensity field δ_m is not directly observable. Weak lensing probes the convergence, given by

$$\kappa(\chi_*, \hat{\boldsymbol{n}}) = \frac{3H_0^2\Omega_m}{2c^2} \int_0^{\chi_*} d\chi \, \frac{\chi}{a(\chi)} \left(1 - \frac{\chi}{\chi_*}\right) \delta_m(\chi\hat{\boldsymbol{n}}), \tag{4.3}$$

where χ is the comoving distance, χ_* is the comoving distance to the source, $\hat{\boldsymbol{n}}$ is the direction on the sky, H_0 is the Hubble constant, c the speed of light, $a(\chi)$ is the expansion factor, and we assume zero curvature. Hence, κ can be viewed as a measurement of $\Omega_m\delta_m$ averaged over a radial window along the line of sight between the observer and the source. This dilutes the information contained in the total matter field.

A second issue is that we also cannot measure δ_{cb} directly. What we can typically measure from galaxy observations is a biased version, where at the linear level we have $\delta_g = b_1\delta_{cb}$, with galaxy overdensity δ_g being modulated by the linear bias b_1. The linear bias is constant on large scales, but has complicated scale dependence on small scales which cannot be predicted ab initio and thus has to be marginalized over to obtain constraints on cosmological parameters. One way to measure it is using redshift-space distortions (RSD), which at the linear order probes density-velocity correlations. Velocity can be related to the matter overdensity via $\delta_v = f\delta_m$, where f is the linear growth rate which depends on the matter density Ω_m. The growth rate is also affected by neutrinos, which slow down the growth of structure on small scales. However, on small scales, i.e. beyond linear order, this relation also becomes more complicated due to higher order velocity-density correlators [see e.g. 107], once again making it difficult to isolate the effects of neutrino mass.

Multi-tracer analyses, combining δ_g from spectroscopic or photometric surveys, with weak lensing κ from the cosmic microwave background (CMB) or large-scale structure (LSS), suggest that LSS surveys have the power to separate neutrino mass from other parameters, and that sampling variance cancellation is helpful on large scales [319, 395]. Nevertheless, this approach is limited to about 20meV precision on the sum of neutrino masses for surveys such as LSST, suggesting it may not be able to give a neutrino mass detection at more than 3 sigma for the minimum theoretical mass of 60meV.

This limited precision from multi-tracer probes has revived interest in measuring neutrinos from a single tracer using non-linear information. By studying the non-linear effects of massive neutrinos on structure formation [311, 82, 83, 337, 375, 18, 68, 69, 119, 380, 382, 94, 95, 47, 26, 89, 368, 6, 150, 203, 330, 396, 48, 247, 130, 103, 104, 57], several such statistics have been proposed, including the bispectrum, halo mass function, void size function, probability distribution function, and marked power spectrum [222, 246, 242, 120, 254, 11, 178, 177, 366, 257, 62, 223]. The reasoning is that a single tracer may have access to different types of information in different density regions. For example, while high density regions may be mostly sensitive to the CDM+baryons, which cluster and gravitationally collapse into virialized objects, low density regions such as voids may be more sensitive to neutrinos,

which cluster weakly in comparison. Crucially, this implies that a full description of the system requires a two-fluid model, that of CDM+baryons and of neutrinos, which cannot be mimicked by a single CDM+baryon component. The hope of this approach is that by effectively combining information from different density regimes one might be able to determine neutrino mass to a much higher precision than predicted by just the two point statistics, the power spectrum.

The goal of this paper is to investigate this single tracer proposal by comparing a single-fluid CDM simulation to a two-fluid simulation with CDM and neutrinos (for the purpose of this paper we assume baryons trace CDM). We examine whether the presence of massive neutrinos has a unique *non-linear* effect that differentiates the two at late times, or if the impact of the massive neutrino component can be faked by a solitary CDM component. To this end, we set up the two simulations with matched linear power spectrum of the field in question, and equal phases, at a redshift of interest, which we will take to be $z = 0$. We compare the two simulations at the field level for three different fields: (i) δ_{cb}, which uniquely defines anything observable with galaxies, (ii) $\Omega_m\delta_m$, which is the corresponding field controlling weak lensing observables, and (iii) δ_m, the 3d total matter field which is not currently observable.

If at the field level the two simulations differ in their phases at $z = 0$, this would suggest there is information that has been created by the non-linear evolution that is unique to the presence of massive neutrinos, and that cannot be mimicked by a single CDM fluid. If, on the other hand, the final phases are matched exactly, then there is no information associated with the difference in non-linear evolution beyond the overall amplitude of the field, i.e. the power spectrum. If the power spectra at $z = 0$ are also identical between the two simulations then there is no non-linear information arising specifically from the presence of the neutrino component, and any information regarding neutrino mass must simply arise from the differing linear physics. A similar analysis was performed in the context of modified gravity by [97], which studied only the non-linear power spectrum. Earlier work in the context of neutrino mass includes a study of the halo mass function [96], and the non-linear matter power spectrum for the Ly-α forest [284]. We will generalize such analyses by considering the information at the field level.

The structure of this paper is as follows. In Section 4.2 we outline how to study the information content of cosmological fields. In Section 4.3 we apply this to understand the amount of neutrino mass information in the various aforementioned cosmological fields. In Section 4.4 we then comment on the benefits of probes beyond the power spectrum (for example related to halos and voids). In Section 4.5 we consider a Fisher analysis to compare constraints obtained from the linear and non-linear power spectrum. Finally, in Section 4.6 we conclude and discuss how our findings relate to constraints on M_ν presented in recent works.

4.2 Cosmological Information

The simplest tool used to quantify the information content of a field $\delta(\boldsymbol{k})$ is the (auto) power spectrum $P_{\delta\delta}(k)$, defined via

$$\langle \delta^*(\boldsymbol{k})\delta(\boldsymbol{k}')\rangle = (2\pi)^3 P_{\delta\delta}(k)\delta^{(D)}(\boldsymbol{k}-\boldsymbol{k}'), \tag{4.4}$$

where $\delta^{(D)}$ is the Dirac delta function. $P_{\delta\delta}(k)$ is the Fourier transform of the two-point correlation function $\xi(r)$, i.e. it measures the overdensity correlation between two arbitrary points of space separated by r. For a statistically homogeneous, isotropic, and Gaussian field, the power spectrum contains the entire information of the field. The standard model of cosmology assumes homogeneity and isotropy, and that the primordial Universe was described by a Gaussian random field (although we note that there are some extensions beyond this theory, for example positing primordial non-Gaussianity [253, 125, 219, 326, 115, 108, 260]). The overdensity field in Fourier space is in general complex, i.e. it can be written as $\delta(\boldsymbol{k}) = |\delta(\boldsymbol{k})|e^{i\phi(\boldsymbol{k})}$, where $|\delta(\boldsymbol{k})|$ is the magnitude and $\phi(\boldsymbol{k})$ the phase. The phases of a Gaussian random field have a uniform random distribution in the range $[0, 2\pi)$.

The Universe then evolves, and during the late stages of evolution, structure formation introduces non-Gaussianities on small scales due to the non-linear nature of gravitational collapse. The exact nature of this non-linear evolution depends on the cosmological parameters, for example the energy density of dark energy Ω_Λ, the Hubble constant H_0, and the total neutrino mass M_ν. There is thus much interest in studying higher-order statistics, in the hope that they contain additional information beyond the power spectrum. This is particularly true in the case of neutrinos due their signature on small, non-linear, scales. It is thus important to understand how much information neutrinos imprint on different cosmological fields, and furthermore how much of this information arises from non-linear cosmic evolution.

To set up the problem, let us consider two different universes at some late redshift z_f. We denote some generic field as $\delta_X(\boldsymbol{k}, z_f)$ in the first universe with cosmological parameters λ, and $\tilde{\delta}_X(\boldsymbol{k}, z_f)$ in the second universe with cosmological parameters $\tilde{\lambda}$. A question of interest is, if our Universe corresponds to δ, how well can we distinguish it from a universe with field $\tilde{\delta}$? Or in other words, how much information can we learn about the cosmological parameters by studying $\delta(\boldsymbol{k}, z_f)$? While a typical analysis, e.g. a Fisher analysis, considers both linear and non-linear information as one, we seek to isolate the non-linear information. More concretely, while a cosmological field may be sensitive to a change in cosmological parameters, if this sensitivity is purely at the linear level, then there will be no additional information compared to the linear power spectrum; one could consider non-linear probes, such as the halo mass function, void size function, the bispectrum, etc., but they will just be expressing the information content of the linear power spectrum in a different form. So it is interesting to study how much non-linear information there is and thus how much benefit one can expect to extract from non-linear observables.

To quantify how much non-linear information an entire field contains with regard to a change in cosmological parameters $\lambda - \tilde{\lambda}$, we match the linear physics at z_f between the two

cosmologies. We then backscale the fields to some earlier redshift, z_i, using linear theory twice: one time using the cosmology associated with λ, and one time using $\tilde{\lambda}$. Finally we perform an N-body simulation to evolve the two fields to z_f and obtain non-linear results: here again we use the appropriate choice of cosmology in each case. A schematic of these two simulations is as follows:

$$\delta_X^{(1)}(z_f) \xleftarrow[backscale]{\lambda} \delta_X^{(1)}(z_i) \xRightarrow[N-body]{\lambda} \delta_X(z_f), \tag{4.5}$$

$$\delta_X^{(1)}(z_f) \xleftarrow[backscale]{\tilde{\lambda}} \tilde{\delta}_X^{(1)}(z_i) \xRightarrow[N-body]{\tilde{\lambda}} \tilde{\delta}_X(z_f), \tag{4.6}$$

where $\delta_X^{(1)}$ labels the linear power spectrum of component X. The key difference between this approach and a typical analysis is the use of identical initial conditions for both universes to ensure the linear physics is the same at z_f after running the simulation. This means that any difference between $\delta_X(z_f)$ and $\tilde{\delta}_X(z_f)$ after the N-body simulation will be purely due to non-linear effects caused by using $\tilde{\lambda}$ instead of λ.

Having set up the problem, we now review how to quantify the difference between two fields. Rather than considering specific observables, we seek to study effects at the field-level. In order to compare the the two fields at z_f, we consider the (complex) coherence of the two fields, defined as

$$\zeta(k) = \frac{P_{\delta\tilde{\delta}}(k)}{\sqrt{P_{\delta\delta}(k)P_{\tilde{\delta}\tilde{\delta}}(k)}}, \tag{4.7}$$

where $P_{\delta\tilde{\delta}}(k)$ is the cross-power spectrum between δ and $\tilde{\delta}$, given by

$$\langle \delta^*(\boldsymbol{k})\tilde{\delta}(\boldsymbol{k}')\rangle = (2\pi)^3 P_{\delta\tilde{\delta}}(k)\delta^{(D)}(\boldsymbol{k}-\boldsymbol{k}'). \tag{4.8}$$

Unlike the auto power spectrum, the cross power spectrum can in general be complex. Note that statistical isotropy and homogeneity enforces the coherence to only be a function of the magnitude k.

Two fields are said to be *coherent* at scale k if $|\zeta(k)| = 1$. In such a case the power spectra of the two fields are linearly related as follows

$$P_{\tilde{\delta}\tilde{\delta}}(k) = \left|\frac{P_{\delta\tilde{\delta}}(k)}{P_{\delta\delta}(k)}\right|^2 P_{\delta\delta}(k), \tag{4.9}$$

where the $|\cdot|^2$ term can be thought of as a linear transfer function between the auto spectra of the two fields.

If the real part of the coherence is equal to 1, the phases of δ and $\tilde{\delta}$ are statistically identical. If the phases of the two cosmologies evolved identically, then the entire difference between the two fields is captured by any difference in the amplitude of the individual power spectra. Furthermore, if two fields are coherent, and the transfer function is identical to unity, $|P_{\delta\tilde{\delta}}(k)/P_{\delta\delta}(k)| = 1$, this implies that the power spectra are identical and that there is thus no non-linear information in the power spectrum. In such a case the two cosmologies are

statistically indistinguishable in terms of non-linear effects, and there will be no information beyond the linear power spectrum. By this we mean that for a given set of scales, the information content of any non-linear statistic cannot exceed the information content of the linear power spectrum over those same scales. While the linear power spectrum is not something one can generally observe for a particular field, it is useful to know whether or not there is information that exists beyond linear theory.

4.3 Massive Neutrino Information

Using the notation of the previous section we use λ to denote a universe with massive neutrinos, $M_\nu = 0.15$eV, and $\tilde{\lambda}$ to denote a universe without massive neutrinos, $M_\nu = 0$. We start by using a Boltzmann solver to compute the linear power spectrum for a cosmology with $M_\nu = 0.15$eV at $z_f = 0$. We then backscale this power spectrum to $z_i = 99$ twice, one time using the linear physics associated with massive neutrinos (using the REPS package [399]), giving $P_X(k, z_i)$, and the other time using the linear physics associated with massless neutrinos, giving $\tilde{P}_X(k, z_i)$. We generate realizations of the two fields at $z = 99$ with matched phases. Note that the massless neutrino cosmology is thus initialized with a power spectrum whose shape encodes the linear suppression of growth due to the presence of massive neutrinos in the other cosmology. Then we evolve $P_X(k, z_i)$ through to z_f using the Gadget N-body simulation [346] with massive neutrinos, yielding $\delta_X(\boldsymbol{k}, z_f)$, and we similarly evolve $\tilde{P}_X(k, z_i)$ through to z_f using the N-body simulation without massive neutrinos, yielding $\tilde{\delta}(k, z_f)$. Since the linear predictions of the two cosmologies have been matched as closely as possible, we can determine how much non-linear evolution is special to the presence of massive neutrinos by comparing the fields at z_f: $\delta(\boldsymbol{k}, z_f)$ and $\tilde{\delta}(\boldsymbol{k}, z_f)$. Furthermore, by comparing the power spectra of the fields, $P_{\delta\delta}(k, z_f)$ and $P_{\tilde{\delta}\tilde{\delta}}(k, z_f)$, we can assess the information in the power spectrum. We refer to the $M_\nu > 0$ simulation as the "real" simulation, and the $M_\nu = 0$ simulation as the "fake" simulation, because the purpose of the $M_\nu = 0$ simulation is to fake the effects of massive neutrinos by using a single-fluid CDM simulation with initial conditions associated with a massive neutrino cosmology. Note that for each considered field, δ_X, a different fake N-body simulation is run with matched linear physics for that particular field. We consider a box of volume $1\,(\mathrm{Gpc}/h)^3$ and a grid of dimension 1024^3 for both CDM and neutrinos.

In the case of lensing, the field δ_m is not directly measured. Instead, lensing measures $\Omega_m\delta_m$ averaged over a window function integrated over the line of sight, as described in Eqn. (4.3). We are therefore free to define the effective lensing field by rescaling by a constant factor, which we choose to be $(1 - f_\nu)$ as follows:

$$\kappa \sim \Omega_m\delta_m = \Omega_m(1 - f_\nu)\frac{\delta_m}{(1 - f_\nu)} = \Omega_c\frac{\delta_m}{(1 - f_\nu)}. \tag{4.10}$$

Assuming no a priori information regarding Ω_m, we can evaluate the lensing information by

considering the information in the field defined by

$$\delta_{\Omega m} \equiv \frac{\delta_m}{(1-f_\nu)}. \tag{4.11}$$

While we are free to choose any normalization, the reason for this choice is that we seek the option which most closely matches the real and fake cosmologies. From Eqn. (4.2) it is clear that $\delta_{\Omega m} \sim \delta_{cb}$ on small scales, thus this choice is inspired such that neutrino effects should be negligible on small scales (note that in the case of $\delta_{\Omega m}$ linear matching cannot be achieved at large scales, but rather on small scales).

In Fig. 4.1 we plot the real part of the coherence for each of the fields between the real and fake simulations. Specifically, when we match P_X at $z = 0$, we plot the coherence for the X overdensity field. It can firstly be seen that in the case of $X = cb$, the coherence is unity up to $k = 1\,h/\mathrm{Mpc}$ to $\lesssim 0.01\%$. This implies that the final phases of the cb field are equivalent regardless of whether massive neutrinos are included in the simulation. This is to say that the non-linear evolution of the cb field is identical in both the one-fluid (CDM) and two-fluid (CDM+ν) description. Thus there is negligible non-linear information in the cb field that goes beyond the power spectrum within scales of experimental interest. On the other hand, the coherence for δ_m begins to differ from 1 at a lower value of k, implying there is non-linear information beyond the power spectrum for the 3d matter field. However, one cannot measure the matter field directly, and one instead measures lensing which is related to $\delta_{\Omega m}$, as in Eqn. (4.11). In this case the picture is identical to δ_{cb}, with a coherence of 1 up to $k = 1\,h/\mathrm{Mpc}$ to $\lesssim 0.01\%$, implying negligible non-linear information beyond the power spectrum in this field at these scales.

Having established that, for scales of interest, the information in the case of δ_{cb} and $\delta_{\Omega m}$ is all in the power spectrum, we now consider how much information the power spectrum contains. In Fig. 4.2 we plot the ratio between the power spectra for the various fields at redshift z_f. We see that for P_{cb} the ratio is always 1, implying that there is negligible non-linear information regarding neutrinos in the cb power spectrum. (We note that the $\lesssim 0.1\%$ upturn for scales smaller than $k \approx 0.5\,h/\mathrm{Mpc}$ is a numerical artifact caused by a slight discrepancy between the growth factor implemented in backscaling and that effectively implemented by the N-body simulation. The magnitude of this discrepancy depends on M_ν, leading to this small effect.) Given the coherence of the cb field is 1, this means there is negligible non-linear information about neutrino mass in the entire cb field. On the other hand, there is a deviation of order 1% in $P_{\Omega m}$ for $k \lesssim 0.1\,h/\mathrm{Mpc}$. This implies there is some information on neutrino mass in the lensed matter power spectrum. This is the typical shape information associated with neutrinos, however, it mostly appears on large, linear, scales and will thus be sample variance limited. Finally, we see that the ratio for P_m differs from 1 on small scales, implying the presence of non-linear information about neutrinos beyond the the linear power spectrum of the 3d total matter field.

To summarize, whenever we consider the single-fluid CDM field, we find that there is no difference between the real and fake simulations. On the other hand, whenever we consider fields that explicitly depend on both fluids (CDM+ν) in the real simulations, we generally

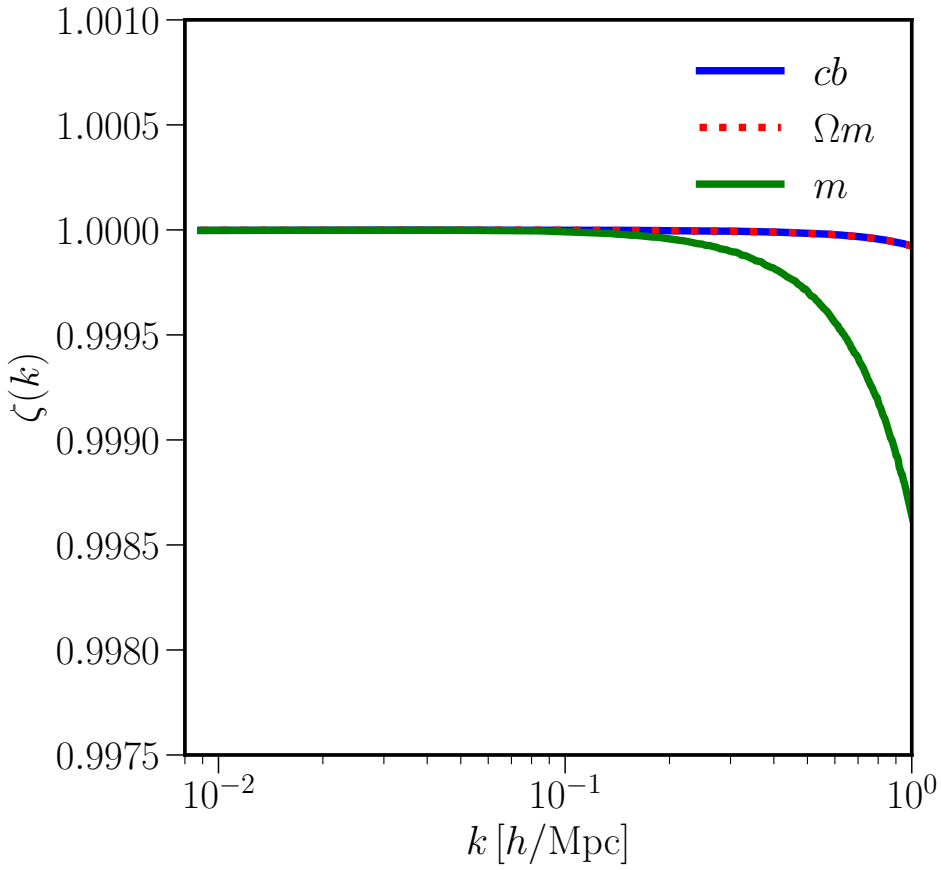

Figure 4.1: The real part of the coherence between fields from the real and fake simulations. If we match δ_X at $z = 0$, we plot the coherence for δ_X. It can be seen that the coherence in the case of δ_{cb} and $\delta_{\Omega m}$ is 1 up to $k = 1\,h/\mathrm{Mpc}$ to $\lesssim 0.01\%$. This implies that there is negligible non-linear information in the cb field or the lensed matter field at these scales that goes beyond the power spectrum. On the other hand, the coherence for δ_m begins to differ from 1 at a lower value of k, implying non-linear information beyond the power spectrum for the 3d matter field.

find that a single fake simulation cannot reproduce the statistics on all scales: either they remain matched on large scales, or they remain matched on small scales. The two choices we explore, m and Ωm, illustrate this clearly. For m, the small scales have a different non-linear behavior even though the linear statistics are exactly matched. For Ωm, the large scales are not matched even at the linear level, but crucially, the small scale matching is maintained both at the linear and non-linear level. Since for Ωm only the linear scales are not matched well, most of the information should be contained in the linear power spectrum.

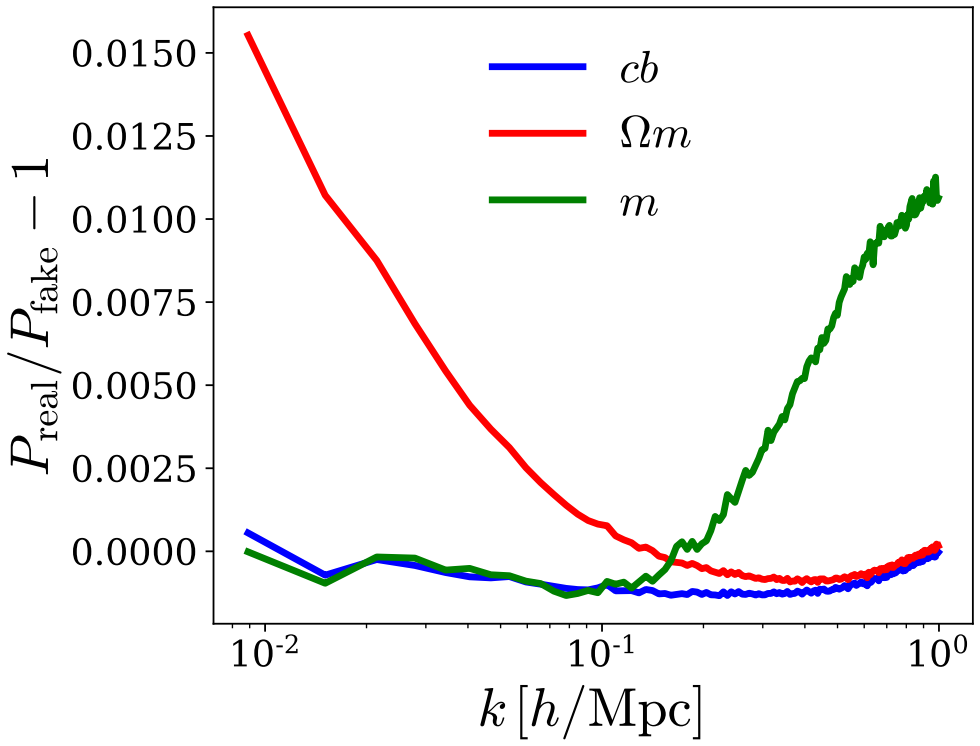

Figure 4.2: The ratio of the power spectra between the real and fake simulations. If we match δ_X at $z = 0$, we plot the corresponding power spectrum, P_X. It can be seen that the ratio is 1 for cb, while there is an approximately 1% deviation for $P_{\Omega m}$ on large scales. Only P_m differs form 1 on non-linear scales, implying information beyond the power spectrum in for the matter field. In the cases of cb and Ωm, the $\lesssim 0.1\%$ upturn on scales smaller than $k \approx 0.5\,h/{\rm Mpc}$ is a numerical artifact due to discrepancy between the backscaling and forward model; a similar effect can be seen in the case of m for which a downturn begins at $k \approx 0.5\,h/{\rm Mpc}$.

4.4 Higher-Order Statistics

We now illustrate the effect of the results of the previous section on various statistics beyond the power spectrum. While the results of the previous section are sufficient in determining the presence of information regarding M_ν in any non-linear statistic beyond the power spectrum, we now show this explicitly for various examples in the interest of clarity.

We start with the void size function (VSF), a commonly proposed source of information regarding neutrino mass [222]. We use the spherical void finder of [47] with a threshold of $\delta_{\rm th} = -0.7$, and look for voids in the three considered fields: cb, Ωm, and m. In Fig. 4.3 we find that there is no difference in the VSF between the real and fake simulations for both the cb and Ωm fields, but there is potentially some difference for the 3d matter field. Note that we find similar results regardless of the value of $\delta_{\rm th}$.

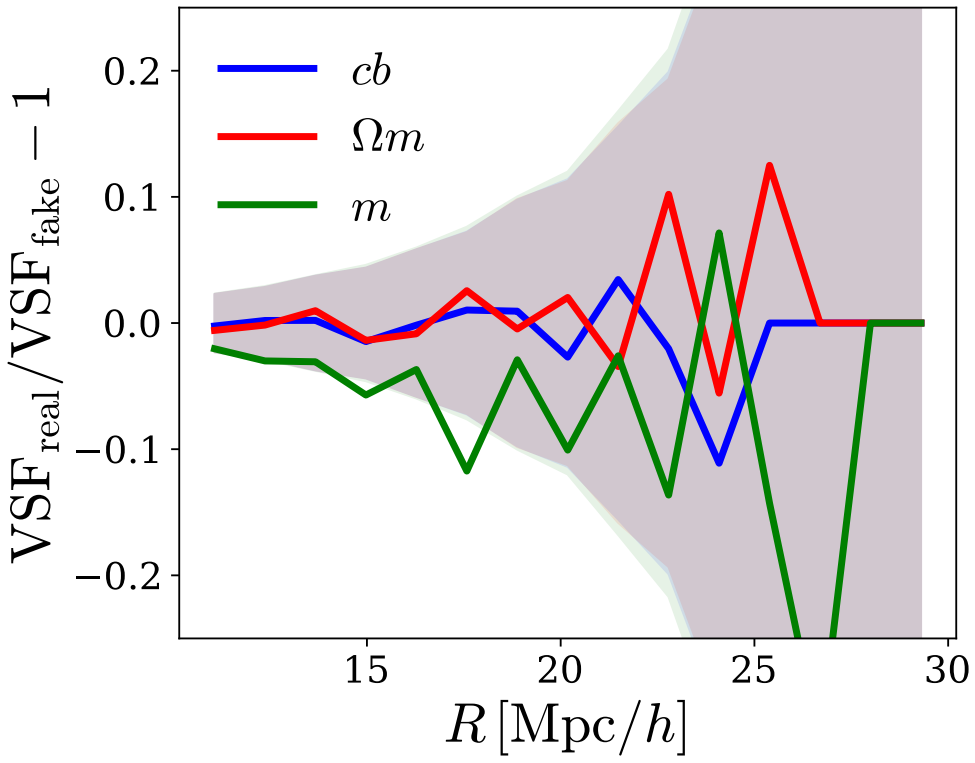

Figure 4.3: The ratio of the void size function between the real and fake simulations. If we match δ_X at $z = 0$, we plot the corresponding VSF in the X field. Bands represent Poisson errors. It can be seen that the ratio is 1 for cb and Ωm within the Poisson errors. Only the VSF in the 3d matter field shows a ratio that isn't unity, although it is still close to the Poisson error.

Next, in Fig. 4.4, we consider the marked power spectrum. We use the optimal choice of mark parameters found in [257], which uses the smoothed overdensity field with $10\,\mathrm{Mpc}/h$ smoothing window, thus small scale information is mixed into large scales. We again find little difference in the cb field. The Ωm field differs only on large, linear, scales. For the m field there is a difference on all scales. Again, this fits with our findings in the previous section.

A corollary of there being negligible information in the cb field is that there will also be negligible information in the halo field. The halo field is a function of the cb field and the bias parameters, hence, without knowledge of the bias, the information content of the halo field is just a re-expression of the information contained in the cb field. We illustrate this in Fig. 4.5, which shows the difference in the power spectrum, void size function, and marked power spectrum, for the halo field between the real and fake (cb-matched) simulations. We identify halos using the Friends-of-Friends (FoF) algorithm and apply a fixed number density cut. We find that there is no significant difference in any of the halo statistics between the real and fake simulations.

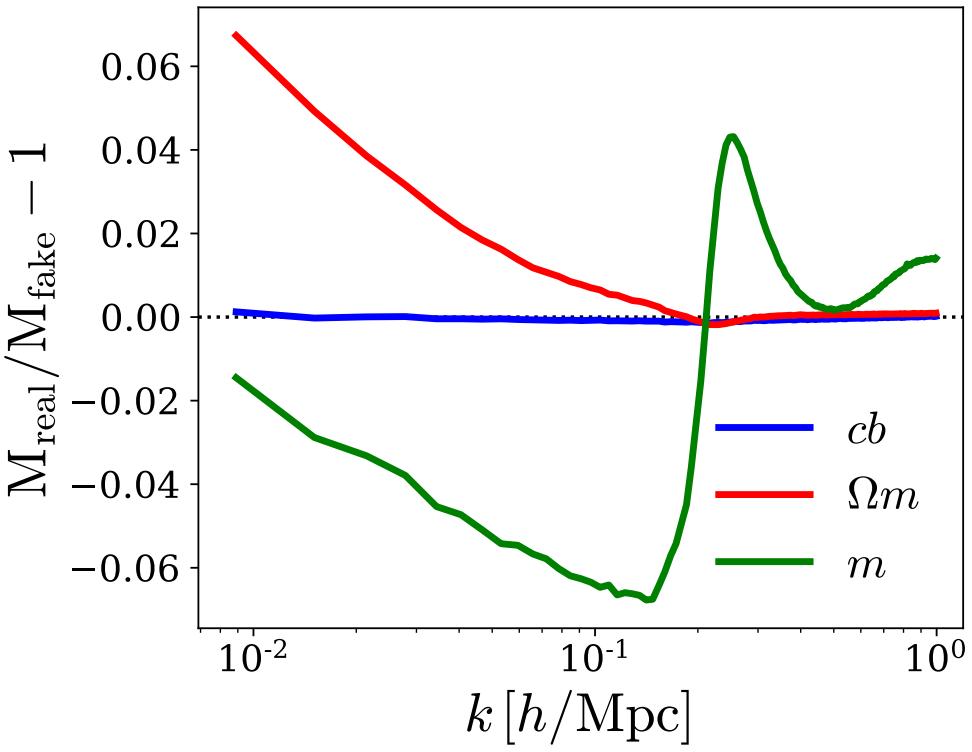

Figure 4.4: The ratio of the marked power spectrum (M) between the real and fake simulations. If we match δ_X at $z = 0$, we plot the corresponding M in the X field. It can be seen that the ratio is 1 for cb, while for the Ω_m it deviates form 1 on large scales, and for m is deviates from 1 on all scales.

Having shown there to be little information in the real-space halo field, we now consider redshift-space distortions (RSD), which includes the effects of the peculiar velocities of halos along the line of sight (LoS). The peculiar velocity field is sourced by the clustering of the total matter field, which includes neutrinos. The halo power spectrum in redshift space can, therefore, provide additional information on the total neutrino mass. In the left panel of Fig. 4.6 we show a bin-by-bin comparison of the redshift-space halo power spectrum from the real and fake (with matched cb field) simulations in the $(k_\parallel, k_\perp)$ plane, where parallel/perpendicular is in reference to the LoS. It can be seen that some bins along the LoS have a relatively large difference between the two simulations, but even small deviations from the LoS direction brings the size of the effect down to $\lesssim 1\%$, in line with the results obtained in real space. To better visualize the dependence on magnitude, k, and projection onto the LoS, $\mu = k_\parallel/k$, we bin the data into 3 bins of k and 4 bins of μ. The right panel shows the difference between the real and fake simulation increases with k and μ, signifying the information present at small scales due to RSD as one approaches the LoS. Therefore, we conclude that there is indeed additional non-linear information about neutrino mass that can

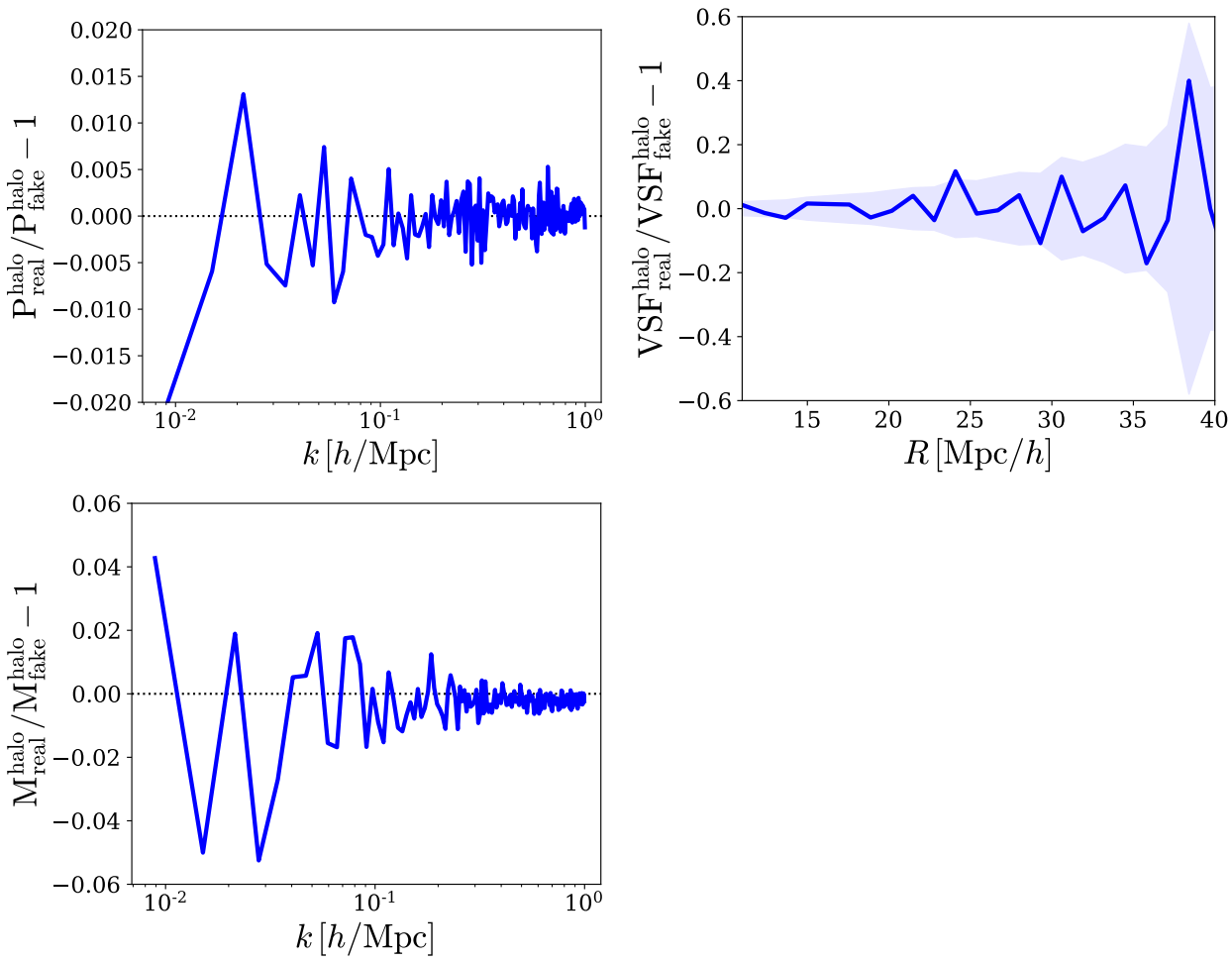

Figure 4.5: The ratio of the halo-traced power spectrum, void size function, and marked power spectrum (from left to right and then top to bottom) between the real and fake (*cb*-matched) simulations. It can be seen that the ratio is close to unity in all cases.

be obtained by studying clustering of biased tracers in redshift space. While this clustering can be difficult to model accurately, it may be a key source of information in upcoming surveys [160].

4.5 Fisher Analysis

As shown the previous two sections, without RSD, the *cb* field is statistically indistinguishable between the one-fluid (CDM) and two-fluid (CDM+ν) simulations. The same is also approximately true for the Ωm field, responsible for weak lensing, for which there is only a difference in the power spectrum on large scales. If there is negligible difference between the

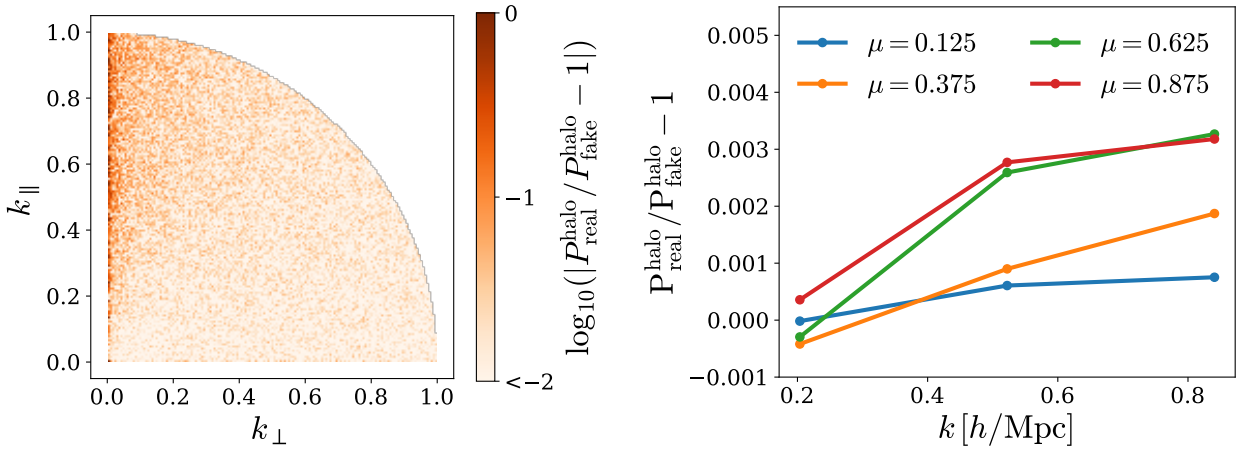

Figure 4.6: The ratio of the redshift-space halo power spectrum between the real and fake (*cb* matched) simulations. *Left*: bin-by-bin comparison in the $(k_\parallel, k_\perp)$ plane, where parallel/perpendicular is in reference to the LoS. *Right:* binning the data into 3 bins of magnitude, k, and 4 bins of LoS projection, $\mu = k_\parallel/k$. Both plots show a deviation of the ratio from unity when moving closer to the LoS and to smaller scales, but negligible deviation in the perpendicular direction, suggesting that the additional information on neutrino mass comes from the modified velocity field, or growth rate, which is sourced by the matter overdensity.

one-fluid and the two-fluid non-linear dynamics, the total information content is essentially maximized by that which arises from the linear physics. Nevertheless, as the linear power spectrum is not observable, it is instructive to compare the information content of the linear power spectrum to the non-linear power spectrum.

We perform a Fisher analysis in the $\{\Omega_m, \Omega_b, h, n_s, \sigma_8, M_\nu\}$ plane. We use a fiducial cosmology with $\Omega_m = 0.3175$, $\Omega_b = 0.049$, $h = 0.6711$, $n_s = 0.9624$, $\sigma_8 = 0.834$, and $M_\nu = 0.05$eV. To compute derivatives we use a central difference scheme at $\pm\delta\theta$ for each cosmological parameter. Specifically we use $\delta\Omega_m = 0.01$, $\delta\Omega_b = 0.002$, $\delta h = 0.02$, $\delta n_s = 0.02$, $\delta\sigma_8 = 0.015$, and $\delta M_\nu = 0.025$eV. For the linear covariance between probes x and y, we use $C_X = 2P_X^2/N_k$, where $N_k = 4\pi k^2 k_F/k_F^3$, and $k_F = 2\pi/L$ is the fundamental wavenumber which we take for a box of volume $1\,({\rm Gpc}/h)^3$. For the non-linear results we use the Quijote simulations [383], and for the linear results we use CAMB [239], using the same derivative computation method and binning as Quijote.

Fig. 4.7 shows the marginal error on M_ν as a function of $k_{\rm max}$ for the various linear and non-linear power spectra. As expected there is good agreement between the linear and non-linear results on large scales, where cosmic evolution is approximately linear. Moving to smaller scales, we see that the non-linear power spectra for *cb* and Ωm have a factor 2 times lower constraining power compared to their linear counterparts. Note that the non-linear power has worse constraints because its covariance has positive off-diagonal elements due to mode coupling, which in turn degrades the information content after marginalizing [62].

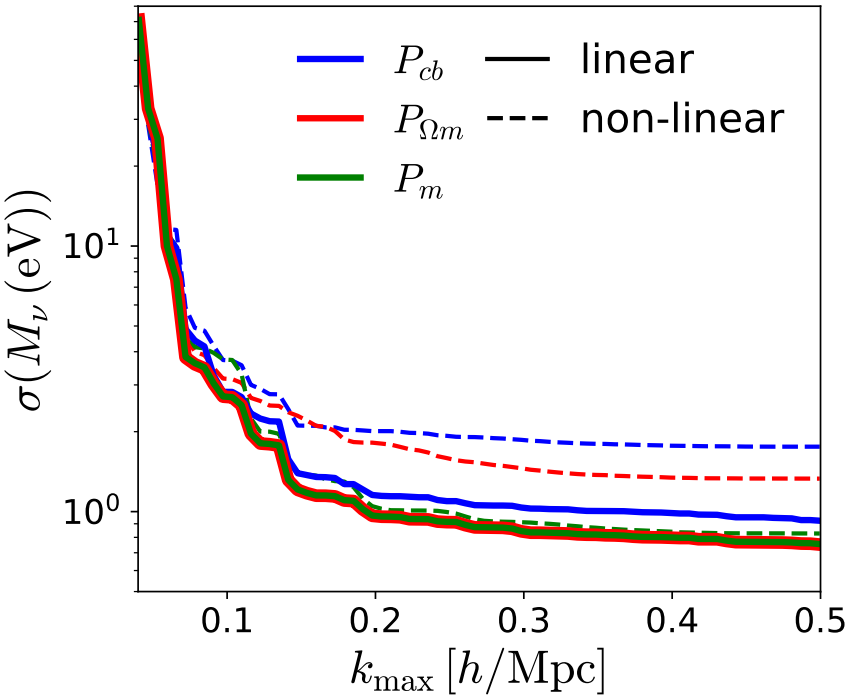

Figure 4.7: Marginal error on M_ν for P_{cb} (blue), $P_{\Omega m}$ (red), and P_m (green), in both the linear (solid) and non-linear (dashed) regime, for a volume of $1\,(\mathrm{Gpc}/h)^3$.

This means that there is still potential room for improvement upon the constraints from the non-linear power spectrum, and one may benefit from around a factor 2 by using statistics beyond the power spectrum. Thus, on their own, the cb and Ωm fields give a marginal error on the neutrino mass of just under 1eV in a $1\,(\mathrm{Gpc}/h)^3$ volume.

On the other hand, the linear and non-linear marginal error on P_m match well all the way to $k_{\mathrm{max}} = 0.5\,h/\mathrm{Mpc}$. But, regardless of this, it was shown in the previous section that there is additional information in the phases of the 3d matter field that is not fully captured by the power spectrum, and there is thus additional information to be found in higher-order statistics.

4.6 Discussion and Conclusions

In this paper we have investigated how much non-linear information regarding neutrino mass one can expect to find in various cosmological fields by comparing one-fluid (CDM) to two-fluid (CDM+ν) simulations with matched initial conditions. In real space, we found that the cb field and Ωm (lensing) field do not contain additional information regarding neutrino mass that is unique to the two-fluid dynamics up to $k \lesssim 1\,h/\mathrm{Mpc}$. Essentially, the evolutionary effect of including a massive neutrino fluid can be faked by a solitary CDM fluid.

This implies that the *cb* field, and derived quantities (e.g. the halo field), and weak-lensing convergence, contain little information regarding neutrino mass beyond that which exists in the linear power spectrum over the same scales. We have also shown that there is much non-linear information regarding neutrino mass in the 3d matter overdensity field, however, this is not currently experimentally detectable. The fundamental quantities we considered are the coherence and power spectrum ratio between the two simulations, summarized in Table 4.1, which alone quantify the amount of non-linear information at the field level. We then explicitly verified these findings for various higher-order statistics, including the void size function and marked power spectrum.

Consequently, one can expect constraints on neutrino mass a little lower than 1eV in a volume of $1\,(\mathrm{Gpc}/h)^3$ when using the *cb* or lensing fields alone at a single redshift. Hence, using only this information, a very large volume of $10^4\,(\mathrm{Gpc}/h)^3$ would be needed to reach an error of 0.01eV (corresponding to a ~ 5 sigma detection), which exceeds the available volume of currently realistic surveys.

We note that even in the face of these findings, there is still motivation to consider statistics beyond the power spectrum to detect neutrino mass. A first consideration is the choice of redshift(s). In our analysis we have matched the linear physics at a single redshift, $z = 0$. Similarly, [284] used hydrodynamical simulations to find one can fake the effects of massive neutrinos in the non-linear power spectrum up to $k \lesssim 10/\mathrm{Mpc}$ for the Ly-α forest ($z = 3$). While the effects of massive neutrinos can be faked at a single redshift, the real and fake universes have, in principle, different evolution. Therefore, combining the fields at multiple redshifts should help discriminate between the two and improve constraints on M_ν. It will thus be imperative to combine multiple redshifts (from CMB redshift of 1100 to today) and tracers (CMB, galaxies, and weak lensing) to obtain tight constrains on neutrino mass in upcoming surveys. For example, combining weak lensing and galaxy clustering, can reach 0.02eV with Rubin (LSST) and Stage IV CMB [319].

Second, the late-time linear power spectrum is not an observable quantity, as cosmic evolution is indeed non-linear. Hence, even though the linear power spectrum may provide a bound for the error in the case of *cb* and lensing, the non-linear power spectrum does not quite reach this bound. We have shown this effect corresponds to around a factor of 2, thus a different non-linear statistic may be able to obtain slightly better constraints than

Table 4.1: Summary of key results. The coherence and power spectrum ratio between the real and fake simulations for the *cb*, Ωm, and m fields, for $k \leq 1\,h/\mathrm{Mpc}$. Note that while the power spectrum ratio for the Ωm field differs from unity at the 1% level, this is only at low k which is sample variance limited.

Field	$\zeta - 1$	$P_{\mathrm{real}}/P_{\mathrm{fake}} - 1$
cb	$\lesssim 0.01\%$	$\lesssim 0.1\%$
Ωm	$\lesssim 0.01\%$	$\lesssim 1\%$ (low k)
m	$\lesssim 0.1\%$	$\lesssim 1\%$

the non-linear power spectrum. This factor of 2 can also be recovered by reconstructing the linear field from the non-linear field [329].

Third, we consider lensing measurements to be directly sensitive to the product $\Omega_m\delta_m$. This is exact if the sources are at low redshift. However, the comoving distance in Eqn. (4.3) implicitly depends on Ω_m as well, so for sources at higher redshift the relation is more complicated. Thus all possible combinations of Ω_m and δ_m that keep $\Omega_m\delta_m$ fixed may not be compatible with the observed lensing signal because they will modify the comoving distances. If one could obtain strong constrains on Ω_m from the redshift-distance relation, then combining it with lensing measurements may be able to probe δ_m directly, rather than the product $\Omega_m\delta_m$. We also note that neutrinos are non-relativistic at low redshift and thus will not induce a significant geometric effect on lensing observables which is known to arise in the context of dark energy [339, 404].

Fourth, we have motivated that redshift-space distortions (RSD) may provide non-linear information regarding neutrino mass, thus considering higher-order statistics in redshift space is a worthwhile pursuit. RSD adds new information because velocities are determined by the growth factor f, which is sensitive to matter density Ω_m and neutrino density Ω_ν. While RSD can be difficult to model, it could be a key source of information in upcoming surveys [160]. For example, [224] illustrates how halo velocities can aid in constraining neutrino mass. A further improvement on f may be possible from redshift dependence, which we did not consider in this paper.

Fifth, it might be argued that even for the cb or Ωm fields one could find information regarding M_ν beyond the linear power spectrum, as there may be a non-linear statistic with more favourable parameter degeneracies. For example, a particular non-linear statistic might constrain some other cosmological parameter much better than the linear power spectrum, and thus after marginalizing over this parameter, the constraints on M_ν will outperform the linear power spectrum. However, the other parameters of key relevance in the case of neutrino mass are Ω_m and A_s, for which non-linear cosmic evolution does not induce additional information beyond that which exists in the linear initial conditions. We illustrate this in the appendix for this chapter (Section 4.7). Thus it is not expected that degeneracies will cause a big improvement in the constraints on M_ν. One could also consider non-standard cosmological parameters, for example related to primordial non-Gaussianty or exotic neutrino interactions [221]. For the latter to have an effect there would likely need to be a mechanism which couples the non-linear evolutions of the cb and neutrino perturbations much more strongly than what happens through the Poisson equation. In principle this is possible given a sufficiently strong neutrino-baryon or neutrino-neutrino interaction, and this could help break-degeneracies with neutrino mass if one had a means to measure this non-standard effect.

Numerous recent works have proposed that one can obtain information regarding neutrino mass beyond the power spectrum [246, 242, 120, 254, 11, 178, 177, 366, 257, 62, 223, 109, 371]. Some forecast $\mathcal{O}(0.1\text{eV})$ constraints by employing tomography, which is in good agreement with our results. On the other hand, some works find constraints that are over an order of magnitude smaller than linear theory. Given our findings we are able to explain exactly

where this information comes from. In the case of [178, 177] the information arises from working in redshift space, while for [366, 257, 62] it comes from working with the 3d matter field. Regarding [223], which considers the real-space halo field, the information comes from assuming knowledge of the bias model as a function of cosmology. The bias model can be thought to transfer information on small scales in the *cb* field to larger scales in the halo field, thus information at $k > 1\,h/\mathrm{Mpc}$ in the *cb* field could move to scales of $k < 1\,h/\mathrm{Mpc}$ in the halo field. Hence, if one knew the bias model one could obtain tight constraints on the neutrino mass with modern measurements of the halo field. However, the bias model parameters can have strong degeneracies with the cosmological parameters, for example, the linear bias b_1 is essentially degenerate with σ_8. It is thus important to marginalize over bias, and apply halo mass or number density cuts, to obtain realistic constraints.

Many of the works that compute constraints on M_ν are based on Fisher forecasts, for which one must take great care to avoid inaccurate results [71, 393, 67]. Additionally, a Fisher analysis employs asymptotic limits using the Taylor expansion of the log likelihood, which may not be justified in a realistic data analysis where the posteriors are often non-Gaussian. Thus, while some practitioners do go to great lengths to show that their Fisher matrices have converged, it is unclear how credible such forecasts are for higher-order statistics. There is a growing trend in modern Statistical Inference and Machine Learning to use cross-validation as a golden standard for validation of results. The same standard should be adopted in cosmology as well. This means setting aside some fraction of simulations that are not used for training (i.e. not used to evaluate the covariance or derivatives of summary statistics), and performing an end-to-end analysis on these validation simulations all the way to the cosmological parameters of interest, where the result can be compared to the truth in terms of bias and variance. Such an analysis is expensive, even more so if the validation simulations are chosen to be produced by an independent simulation code, but this could be a worthwhile standard validation procedure. Another worthwhile verification strategy is to use null tests, in which one explicitly performs the analysis on setups where the signal is known to be null. An example is running non-Gaussian statistical analysis on Gaussian data to demonstrate that the Fisher analysis does not give more information than what is available in the Gaussian field. Thus a useful piece of future work would be to train a neural network to learn the effects of massive neutrinos on the various cosmological fields and perform all of these tests. In the absence of such work, we intend for our results to give a useful rule of thumb when proposing new statistics to measure the non-linear effects of massive neutrinos.

Acknowledgements

We thank Francisco Villaescusa-Navarro for many insightful discussions on the manuscript. AB (Arka Banerjee) was supported by the Fermi Research Alliance, LLC under Contract No. DE-AC02-07CH11359 with the U.S. Department of Energy, and the U.S. Department of Energy (DOE) Office of Science Distinguished Scientist Fellow Program. This material is based upon work supported by the National Science Foundation under Grant Numbers

1814370 and NSF 1839217, and by NASA under Grant Number 80NSSC18K1274. We acknowledge the use of CAMB for computations involving the linear power spectrum [239]. The *Quijote* simulations [383] can be found at https://github.com/franciscovillaescusa/Quijote-simulations. The analysis of the simulations has made use of the *Pylians* libraries, publicly available at https://github.com/franciscovillaescusa/Pylians3. Some of the computing for this project was performed on the Sherlock cluster. The authors would like to thank Stanford University and the Stanford Research Computing Center for providing computational resources and support that contributed to these research results.

4.7 Appendix A: Other Parameters (Ω_m and A_s)

In this paper we have studied the effect of neutrino mass, M_ν, on non-linear cosmic evolution. We now briefly discuss the effects of two other cosmological parameters relevant for disentangling the effects of neutrino mass from large-scale structure: Ω_m and A_s.

We first perform the real-versus-fake analysis on the cosmological parameter Ω_m. We seek to test if non-linear evolution leaves an imprint at the field level. To do so, we match the linear $P(k)$ of two $M_\nu = 0$ simulations, but differ the value of Ω_m by 10% between the two, during both backscaling and the forward N-body simulation. Fig. 4.8 shows the coherence between these two simulations, which is shown to be $\lesssim 0.01\%$ for $k \lesssim 1\,h/\mathrm{Mpc}$. This suggests there is negligible additional information regarding Ω_m coming from the non-linear evolution that would be present in higher-order statistics, since the agreement is exact at the field level. Interestingly, this is about the same value as the coherence found for the real-versus-fake M_ν coherence found in Fig. 4.1. Note that this analysis does not take into account any change in the shape of $P(k)$ due to a change in Ω_m, which is information contained in the initial conditions.

The other parameter of relevance when it comes to neutrino mass is the amplitude of linear fluctuations, A_s. As this is the amplitude of the initial linear power spectrum, it is a property of the initial conditions. Thus late-time non-linear evolution cannot produce additional information on A_s.

We thus conclude that there is little information regarding M_ν, Ω_m, or A_s coming from non-linear cosmic evolution. Hence, for $k \lesssim 1\,h/\mathrm{Mpc}$ in the cb or Ωm fields, there is no non-linear statistic that will constrain these parameters significantly better than the linear power spectrum, even after marginalizing.

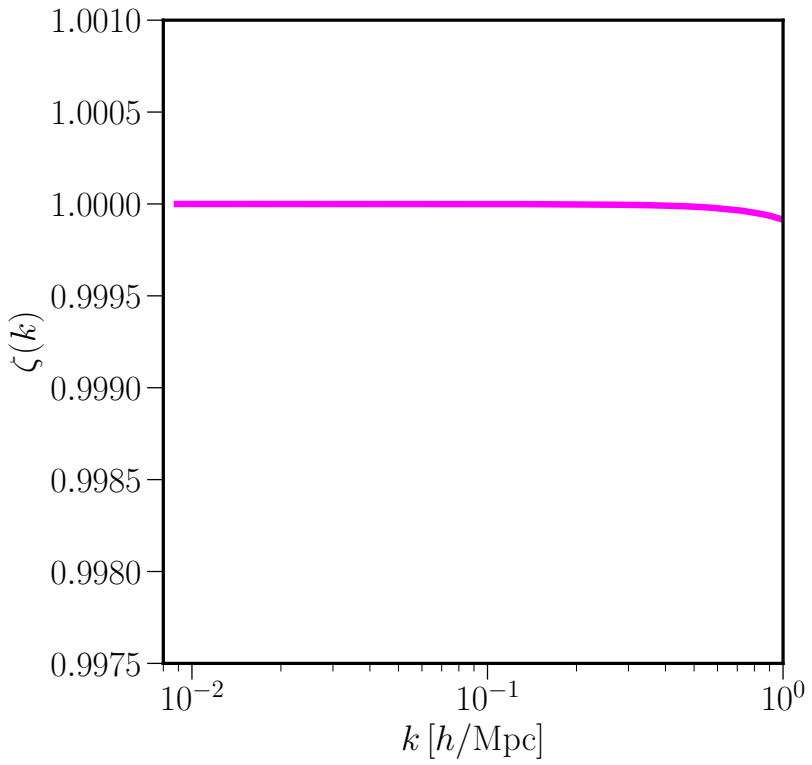

Figure 4.8: The real part of the coherence between the *cb* fields from two $M_\nu = 0$ simulations with matched linear $P(k)$ at $z = 0$, but with Ω_m differing by 10%. It can be seen that the coherence is 1 up to $k = 1\,h/\mathrm{Mpc}$ to $\lesssim 0.01\%$. This implies that there is negligible non-linear information regarding Ω_m in the *cb* field. The vertical range is identical to Fig. 4.1 to enable comparison.

Chapter 5

Joint velocity and density reconstruction of the Universe with nonlinear differentiable forward modeling

The contents of this chapter was originally published in [59],

> Joint velocity and density reconstruction of the Universe with nonlinear differentiable forward modeling
> *Bayer A.E., Modi C., Ferraro S.* (arXiv:2210.15649) (submitted to JCAP)

Reconstructing the initial conditions of the Universe from late-time observations has the potential to optimally extract cosmological information. Due to the high dimensionality of the parameter space, a differentiable forward model is needed for convergence, and recent advances have made it possible to perform reconstruction with nonlinear models based on galaxy (or halo) positions. In addition to positions, future surveys will provide measurements of galaxies' peculiar velocities through the kinematic Sunyaev-Zel'dovich effect (kSZ), type Ia supernovae, the fundamental plane relation, and the Tully-Fisher relation. In this chapter, we develop the formalism for including halo velocities, in addition to halo positions, to enhance the reconstruction of the initial conditions. We show that using velocity information can significantly improve the reconstruction accuracy compared to using only the halo density field. We study this improvement as a function of shot noise, velocity measurement noise, and angle to the line of sight. We also show how halo velocity data can be used to improve the reconstruction of the final nonlinear matter overdensity and velocity fields. We have built our pipeline into the differentiable Particle-Mesh `FlowPM` package, paving the way to perform field-level cosmological inference with joint velocity and density reconstruction. This is especially useful given the increased ability to measure peculiar velocities in the near future.

5.1 Introduction

Reconstructing the initial conditions of the Universe from cosmological data is a pressing task dating back many decades [183, 357, 76, 325]. Recently, there has been much work to achieve this in the context of galaxy surveys, weak lensing, and the Lyman alpha forest [329, 266, 267, 196, 197, 75, 207, 206, 296]. Furthermore, this procedure can be extended to infer cosmological parameters using a field-level approach, which improves constraints compared to a traditional 2-point analysis [329]. This is of particular interest as surveys probe progressively smaller scales in which the effects of nonlinear gravitational evolution moves cosmological information beyond the 2-point statistics. Many summary statistics beyond the power spectrum have been proposed to extract some of this information (see e.g. [353, 323, 64, 213, 315, 245, 248, 211, 332, 255, 246, 242, 222, 120, 310, 254, 11, 178, 177, 129, 366, 20, 175, 188, 46, 257, 110, 109, 62, 223, 58, 371, 370, 147]), however, a field-level approach could enable optimal extraction of this information [329].

While much of the reconstruction literature focuses on using only the galaxy overdensity field as the data, one could consider including additional information in the reconstruction process, such as galaxy peculiar velocities [297]. This is of particular interest as modern surveys begin to provide accurate measurements of peculiar velocities. Current galaxy peculiar velocity catalogs include the 6dF galaxy survey (using fundamental plane) [87] and Cosmicflows-4 (using the Tully-Fisher relation) [220]. In addition, the DESI Bright Galaxy Survey [143] will provide many more measurements of galaxy peculiar velocities using fundamental plane measurements.

Peculiar velocity information can also be obtained from type Ia supernovae measurements [304, 214], for example from the DSS survey [347]. This information can be combined with galaxy surveys [215] or gravitational waves [280] to understand the nature gravity. Furthermore, upcoming CMB experiments, such as Simons Observatory [360] and CMB-S4 [3] will provide accurate measurements of the kinematic Sunyaev-Zel'dovich (kSZ) effect from which peculiar velocities can be obtained [144, 341].

Adding information from peculiar velocities v_r to the reconstruction framework is expected to greatly reduce reconstruction error on large and intermediate scales. This is because (in linear theory) the reconstruction error from peculiar velocity scales as k^2 [341, 270], while density reconstruction has an approximately k-independent noise. Therefore on sufficiently large scales, depending on the shot noise of the galaxy field, the reconstruction from velocity will have lower noise. This is of great importance for measuring parameters like primordial non-Gaussianity of the local type which are sensitive to the very large scales and can be measured with reduced sample variance [326]. This has been studied analytically in linear theory in [270], and it is also a sensitive probe of more general models of multi-field inflation as discussed in [157, 24]. Reconstruction from velocities may also help improve the reconstruction of the Baryon Acoustic Oscillations (BAO), especially close to the boundary of the survey [403].

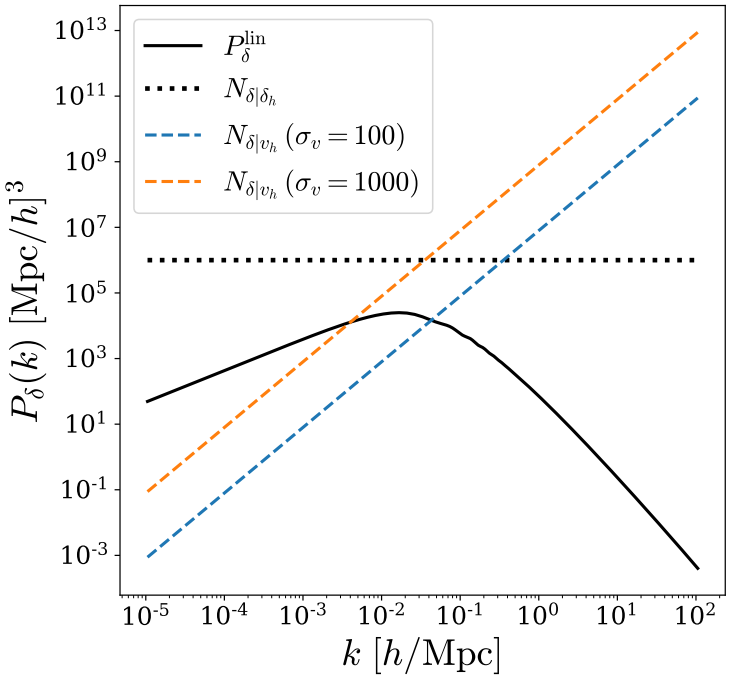

Figure 5.1: The linear matter power (solid black), shot noise (dotted black), and the error on radial modes of density ($\mu = 1$) given measurements of the line-of-sight velocity with varying measurement error σ_{v_r} [in km/s] (dashed colored). The shot noise corresponds to a halo comoving number density of $\bar{n} = 10^{-6}(h/\text{Mpc})^3$.

The shot noise[1], which corresponds to the error in reconstructing the initial field δ using halo overdensity δ_h information, is approximately given by

$$N_{\delta|\delta_h}(k) \approx \frac{1}{\bar{n}}, \tag{5.1}$$

where $\bar{n}$ is the comoving number density of halos[2]. Note that this is an approximation, and the shot noise can differ from the value above due to the effects of exclusion and non-linear clustering [35]. The effect of this term is to decorrelate the true halo density field from the model halo density field, in turn limiting the accuracy of reconstruction from the density field.

[1]Shot noise arises due to the discrete sampling of objects such as halos and galaxies, and is apparent in finite resolution simulations such as our own. In principle, if one had perfect knowledge of halo and galaxy formation, and could run a high enough resolution simulation to model it, one could elude shot noise. While such an endeavour is beyond the scope of this paper, we note that there have been advances in forward modeling progressively smaller scales with progressively higher resolutions to better map the positions of halos from the initial conditions [313].

[2]We will work with halos, as opposed to galaxies, in the remainder of this paper.

Similarly, it is often the case that peculiar velocities are measured for each halo with independent errors of size σ_{v_r} [341]. In this case, the noise on the velocity power spectrum is simply given by

$$N_{v_r|v_{h,r}} \approx \frac{\sigma_{v_r}^2}{\bar{n}}. \tag{5.2}$$

We note that σ_{v_r} may depend on the redshift of the tracer, with the error growing with distance.

The velocity and density power spectra, in linear theory, are related by the continuity equation:

$$P_{v_r}(\mathbf{k}) = \mu^2 \left(\frac{faH}{k}\right)^2 P_\delta(\mathbf{k}), \tag{5.3}$$

where f is the growth rate, a is the scale factor, H is Hubble's constant, μ is the cosine of the angle between $\mathbf{k}$ and the line of sight (so that $\mu = 1$ corresponds to radial modes), and P_δ is the linear density power spectrum. Thus, the error in the density power spectrum given measurements of the velocity is given by

$$N_{\delta|v_{h,r}}(\mathbf{k}) = \frac{1}{\mu^2}\left(\frac{k}{faH}\right)^2 N_{v_r|v_{h,r}} \approx \frac{1}{\mu^2}\left(\frac{k}{faH}\right)^2 \frac{\sigma_{v_r}^2}{\bar{n}}. \tag{5.4}$$

From the previous equations, we can see that while $N_{\delta|\delta_{h,r}}$ is independent of k (Eqn. 5.1), $N_{\delta|v_{h,r}}$ scales as k^2 (Eqn. 5.4) and thus we expect the latter to be smaller on sufficiently large scales.

We illustrate this in Fig. 5.1, where we show the linear matter power spectrum, together with the error expected from measurement of the halo density field (i.e. shot noise), as well as the error on density from measurements of the halo velocity for two different velocity measurement errors. It is apparent that for the configuration considered here, we expect reconstruction from velocity to dominate on large scales.

Note that here we focus on reconstruction based on measurements of the *radial* velocities v_r given the more immediate observational prospects. Velocities can also be measured via direct astrometry, e.g. with GAIA [275]. Moreover, future experiments might also provide measurements of the *transverse* velocity through the "moving lens" effect [70, 198]. The same formalism described in this paper applies to reconstruction using transverse velocities or a combination of radial and transverse, potentially with different noise properties.

In recent applications of field-level inference, the 3-dimensional initial density field is typically comprised of many millions of modes (or voxels); this is determined by computational memory and speed, but in principle an arbitrarily fine resolution could be used. Furthermore, the dynamics of the forward model linking the initial field to the final halo field are highly nonlinear. Reconstruction thus involves performing inference over a highly nonlinear parameter space with millions of dimensions. Knowledge of the derivative of the forward model with respect the initial modes is therefore crucial to efficiently perform the inference. We use `FlowPM`[3] [264], a differentiable Particle Mesh (PM) simulation code writ-

[3] https://github.com/DifferentiableUniverseInitiative/flowpm

ten in TensorFlow. This is entirely differentiable end-to-end, providing derivatives of the forward model, and in turn the data likelihood, with respect to the initial modes and the cosmological parameters. It is also GPU-accelerated which aids the computational efficiency on modern computing nodes[4]. Furthermore, this technique can readily be combined with machine learning techniques to go beyond the traditional N-body approach, and for example connect galaxies and halos [266].

In this work we apply a fully nonlinear forward model using `FlowPM` together with a bias model to map matter to halos. We assess how reconstruction of the initial field from velocity data compares to reconstruction from density data. We also investigate joint reconstruction of the initial field from a combination of both datasets. We consider the effects of the amount of shot noise, velocity measurement noise, and angle to the line of sight.

A further application one might be interested in is predicting another field given the reconstructed initial field. This can simply be achieved by running high fidelity forward models using the reconstructed initial density field. For example, to reconstruct the final matter velocity field, one can run a forward model for the final matter velocity field on the reconstructed initial field. In the case of reconstructing the final matter velocity field, attempts have been made using various approximations, such as assuming the velocity follows a scalar potential and linear approximations [65, 139, 391, 390, 195, 344, 397, 165, 122, 364, 365, 194, 234, 232, 235]. Applications to kSZ measurements are also considered in [274]. More recently [297] used the BORG algorithm [207, 206] to perform reconstruction with peculiar velocity data using a Lagrangian perturbation theory (LPT) forward model on dark matter, finding it to outperform linear reconstruction. To improve upon this, in this work we extend to using a fully nonlinear forward model and consider performing reconstruction with joint velocity and density data.

The structure of the paper is as follows. In Section 5.2 we review the Bayesian procedure of initial mode reconstruction, and discuss how it can be used to perform field-level inference. We also discuss the datasets and forward model employed in this paper. In Section 5.3 we show the quality of reconstruction for both, the initial (linear) and final (nonlinear) matter density fields, as well as the matter velocity field. We conclude in Section 5.4.

5.2 Method

In this section we review the Bayesian methodology of field-level inference. We start in the context of reconstruction from density data, and then describe how velocity can be included. We then discuss the data and forward model employed in this work. For the sake of clarity we will present the theory both for reconstruction and cosmological parameter inference, even though we will not perform cosmological parameter inference in this particular work.

[4]For example, on Cori KNL at NERSC, which has 64 CPUs and 4 GPUs per node, one N-body step using GPU-accelerated code is substantially faster. We invite the reader to consult [264] for a thorough analysis of the GPU performance of `FlowPM`, and [154] for the performance of `FastPM` (which performs the same computation but is written in C and intended to be run on CPU).

Field-level inference

Bayes theorem states that posterior distribution of parameters θ given data d is given by

$$p(\theta|d) = \frac{p(d|\theta)p(\theta)}{p(d)}, \tag{5.5}$$

where $p(d|\theta)$ is the likelihood, $p(\theta)$ is the prior, and $p(d)$ is the evidence. In the cosmological context θ might refer to cosmological parameters which we wish to estimate, and d can refer to the galaxy overdensity field or peculiar velocity data. In principle one would like to estimate the full posterior distribution of θ, but a more tractable, approximate, alternative is to estimate the best-fit value, i.e. the maximum a posteriori (MAP) value, of θ, denoted $\hat{\theta}$, and the width of the posterior around the maximum to quantify the uncertainty of the estimate.

In addition to the parameters we wish to infer, θ, there will typically also be nuisance parameters which we wish to marginalize out of the problem, denoted z. In cosmology, this corresponds primarily to the initial conditions i.e. the initial density distribution of the Universe. Then, the marginalized likelihood required to evaluate Eqn. 5.5 is computed by integrating out the nuisance parameters from the joint likelihood $p(d|\theta, z)$ as follows,

$$p(d|\theta) = \int dz\ p(z, d|\theta) = \int dz\ p(d|\theta, z)p(z|\theta), \tag{5.6}$$

where $p(z|\theta)$ is the prior of z conditioned on θ.

Thus, to compute the posterior, $p(\theta|d)$, Eqns 5.5 and 5.6 show that there are 3 ingredients required: the joint likelihood, $p(d|\theta, z)$, the prior of z conditioned on θ, $p(z|\theta)$, and the prior on θ, $p(\theta)$.[5] We will now discuss each ingredient separately.

Likelihood

In this work we consider data corresponding to the halo overdensity field δ and peculiar velocity along the line of sight v, thus $d = \{\delta, v\}$[6]. The halo overdensity field data corresponds to a 3D mesh containing the value of the overdensity field in each voxel. For clarity of notation, the overdensity field δ can be thought of as a vector consisting of each pixel in the map. On the other hand the peculiar velocity data corresponds to the peculiar velocity of each halo. We thus note that the overdensity is considered at the field level, while the velocity data is considered at the object level.

To reconstruct any latent field, one must make use of a forward model of the data $f(\theta, z, ...)$ which in general depends on the parameters θ and z. This typically corresponds

[5]Note that the evidence, $p(d)$, is a constant and can be dropped when one is only interested in finding the maximum or width of the posterior.

[6]From now on, we use the shorthand v to denote the peculiar velocity. While we only consider peculiar velocities in this work, we note that our formalism is general and can be applied to any component of the velocity field.

to a perturbative model or N-body simulation. There will inevitably be some error in the forward model, as well as some noise in the data. In this analysis we assume these to be Gaussian and uncorrelated, with variance

$$\sigma^2 = \sigma^2_{\rm model} + \sigma^2_{\rm data}. \tag{5.7}$$

Under this assumption, the negative log likelihood is given by the chi-squared difference between the data and a forward model. For halo overdensity data alone the likelihood is thus given by

$$-2\log p(\delta|\theta, z) = \sum_{\boldsymbol{k}} \frac{|\tilde{\delta}(\boldsymbol{k}) - f_{\tilde{\delta}}(\boldsymbol{k};\theta,z)|^2}{\sigma^2_{\tilde{\delta}}(\boldsymbol{k})}, \tag{5.8}$$

where the sum is performed over all modes $\boldsymbol{k}$, and the δ subscript denotes this is the forward model and error for the overdensity data. Note that the density field is evaluated in Fourier space (denoted by the $\sim$), as the model error is typically k dependent. Eqn. 5.8 corresponds to approximating the likelihood to be Gaussian; this assumption works well on large and intermediate scales, but a more sophisticated likelihood would be required to accurately describe small scales, and also to account for observational systematics such as masking, light cones, luminosity dependence, and depth modulation.

Analogously to Eqn. 5.8, for velocity data alone we have

$$-2\log p(v|\theta, z) = \sum_i \frac{[v_i - f_v(\boldsymbol{x}_i;\theta,z)]^2}{\sigma_v^2}, \tag{5.9}$$

where the sum is over all velocity tracers, labeled by i. The forward model for the velocity of the $i^{\rm th}$ halo, $f_v(\boldsymbol{x}_i;\theta,z)$, depends on the position of the the halo $\boldsymbol{x}_i$, as will be described in Sec. 5.2. We note that a more sophisticated likelihood may be required when analyzing data, for example depending on the technique used to extract the signal (e.g. matched filter), and including a tracer-dependent error (here σ_v is the same for all tracers).

While we will consider both density and velocity data individually in this paper, we will also consider the effects of combining density and velocity data. In such a case, and under the assumption of independence, the log likelihood is simply given by the sum of Eqns. 5.8 and 5.9,

$$-2\log p(\delta, v|\theta, z) = -2\log p(\delta|\theta, z) - 2\log p(v|\theta, z) \tag{5.10}$$

$$= \sum_{\boldsymbol{k}} \frac{|\tilde{\delta}(\boldsymbol{k}) - f_{\tilde{\delta}}(\boldsymbol{k};\theta,z)|^2}{\sigma^2_{\tilde{\delta}}(\boldsymbol{k})} + \sum_i \frac{[v_i - f_v(\boldsymbol{x}_i;\theta,z)]^2}{\sigma_v^2}. \tag{5.11}$$

Priors

There are two priors to consider, the prior of z conditioned on θ, $p(z|\theta)$, and the prior on θ, $p(\theta)$. In the context of density reconstruction, the nuisance parameters z refer to the initial

overdensity field modes, and θ corresponds to the cosmological parameters. The prior on θ will typically be motivated by previous measurements of the cosmological parameters from experiments such as Planck [292]. In our analysis we will not do inference on θ and fix θ at it's true value, thus this prior term is not relevant. Based on inflationary theory, which has been verified by Planck, the prior on the initial modes z is taken to be Gaussian with mean 0 and variance proportional to the power spectrum. Hence the prior of z conditioned on θ is given by

$$-2\log p(z|\theta) = \sum_{\boldsymbol{k}} \frac{|\tilde{z}(\boldsymbol{k})|^2}{P(k;\theta)}, \tag{5.12}$$

where $P(k;\theta)$ is the power spectrum of the initial modes, and depends on the cosmological parameters θ. Note that the initial modes are written in Fourier space, and can thus be complex.

Posterior

Adding Eqns 5.11 and 5.12 gives the posterior of z given θ for joint density and velocity inference,

$$-2\log p(z, d = \{\delta, v\}|\theta) = \sum_{\boldsymbol{k}} \left\{ \frac{|\tilde{\delta}(\boldsymbol{k}) - f_{\tilde{\delta}}(\boldsymbol{k};\theta,z)|^2}{\sigma^2_{\tilde{\delta}}(\boldsymbol{k})} + \frac{|\tilde{z}(\boldsymbol{k})|^2}{P(k;\theta)} \right\} + \sum_i \frac{[v_i - f_v(\boldsymbol{x}_i;\theta,z)]^2}{\sigma_v^2}. \tag{5.13}$$

Note that it is z, not $\tilde{z}$, that appears on the left hand side as we perform the inference on the initial field in configuration space, enforcing the physical constraint that z is a real field. To compute the posterior on θ from Eqn. 5.5, one must compute the integral over initial modes from Eqn. 5.6. This can be done using traditional sampling methods, such as HMC as in the BORG method [207], or via optimization with the Laplace approximation or other approximation schemes [329, 262]. In this work we are motivated by the optimization approach, whereby the first step to performing the marginalization integral is to find the value of z which maximizes $p(z, d|\theta)$. In this approach one will first iterate to find the MAP θ, and then find the MAP z to perform the inference. We refer the reader to [329] for further details, but we mention this to note that finding the MAP z is not only interesting from the perspective of reconstructing the initial modes, but also in terms of parameter inference. In this work we will focus on computing $\hat{z} \equiv \max_z p(z, d|\theta)$, in which the cosmological parameters θ are fixed at their true values. We will explore performing inference on θ in future work. We will additionally be interested in the quality of the reconstructed final matter density and velocity fields, which we compute by running the forward model using $\hat{z}$ as the initial conditions. We note that these results could be biased as the forward modeled MAP initial field is not necessarily the MAP final field – the MAP does not generally commute with the nonlinear forward model. A more thorough, unbiased, approach would be to go beyond the MAP and obtain the full posterior, but that is beyond the scope of this work.

We note that one could also use more accurate forward models to obtain the final fields from the inferred initial field to improve the small scale agreement (e.g. [29]).

Data

We consider simulated halo field as data observables. To generate this data, we use `FastPM` [154, 57] with $N_{\rm cdm} = 1024^3$ CDM particles, and a $N_{\rm grid} = 2048^3$ resolution force grid. We use the following cosmological parameters: $\Omega_m = 0.3175$, $\Omega_b = 0.049$, $h = 0.6711$, $n_s = 0.9624$, $\sigma_8 = 0.834$, and $M_\nu = 0$. We begin the simulation at redshift of 9 and use 20 steps to evolve to redshift 0. The halo catalog is computed using the Friends-of Friends (FoF) algorithm with a linking length of 0.2. The halo positions $\boldsymbol{x}_h^{\rm data}$ and velocities $v_h^{\rm data}$ are computed at the halo center-of-mass using `nbodykit` [185]. The halo overdensity field is computed using the cloud-in-cell (CIC) method. We do not consider the effect of redshift space distortions in the main text, as they have little effect on the reconstruction accuracy for halo models on the scales considered in this work. We explicitly show the effect of redshift space distortions on reconstruction in the appendix for this chapter (Section 5.5).

Our fiducial data considers a box of side length $L = 4\,{\rm Gpc}/h$. We select the 67,000 most massive halos, corresponding to a number density of $\bar{n} \simeq 10^{-6}\,(h/{\rm Mpc})^3$, a minimum halo mass of $M_{\rm min} \simeq 6.8 \times 10^{14} M_\odot/h$, and a bias of $b_1 \simeq 4.0$. This data has a high shot noise to illustrate the benefits of velocity reconstruction.

We inject white noise into the velocity data with standard deviation $\sigma_{v,{\rm data}} = 300\,{\rm km/s}$. We inject no noise into the density field $\sigma_{\delta,{\rm data}} = 0$ (although there is still the natural Poisson shot noise due to considering discrete tracers) to understand how helpful velocity data is in this limiting case. Throughout the results we will consider the effect of perturbing individual components of this fiducial setup.

We will additionally consider a lower shot-noise example consisting of a $L = 400\,{\rm Mpc}/h$ box with the 10,000 most massive halos, corresponding to a number density of $\bar{n} \simeq 1.6 \times 10^{-4}\,(h/{\rm Mpc})^3$, a minimum halo mass of $M_{\rm min} \simeq 3.0 \times 10^{13} M_\odot/h$, and a bias of $b_1 \simeq 0.67$. This example allows studying effects on smaller, nonlinear, scales.

Forward models and errors

In this subsection we describe our forward models for both the halo overdensity, $f_{\tilde{\delta}}(\boldsymbol{k}; \theta, z)$, and the peculiar velocity, $f_v(\boldsymbol{x}_i; \theta, z)$. We will also discuss the model errors.

Halo overdensity field forward model

We first compute the matter overdensity field using `FlowPM` [264] (a TensorFlow version of the `FastPM` simulation which was used to generate the data). For the matter-halo connection we use a linear LPT effective field theory model, namely

$$f_{\tilde{\delta}}(\boldsymbol{k}) = \tilde{\delta}_h^{\rm model}(\boldsymbol{k}) = \int d^3\boldsymbol{q}\,[1 + b_1\delta_1(\boldsymbol{q})]\,e^{-i\boldsymbol{k}\cdot(\boldsymbol{q}+\boldsymbol{\psi}(\boldsymbol{q}))}, \tag{5.14}$$

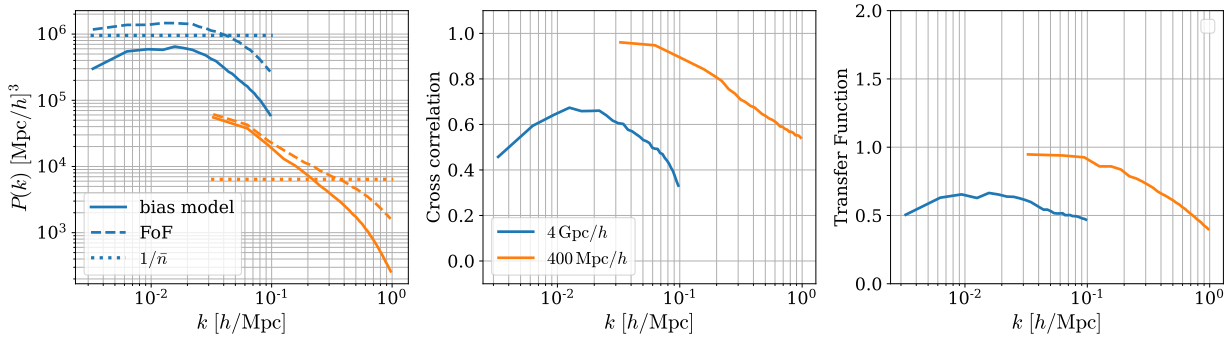

Figure 5.2: Comparison of the best fit bias model to the true FoF halo data. We plot for the fiducial 4 Gph/h box (blue) when the signal-to-noise is low, and a 400 Mph/h box with higher signal-to-noise (orange). Left: bias model (solid), true FoF (dashed) power spectra, and Poisson shot noise (dotted). Middle: cross correlation between model and truth. Right: transfer function between model and truth.

where k is the wavenumber, q are the grid coordinates, b_1 is the linear bias parameter, δ_1 is the linear matter overdensity field (which corresponds to z in the Sec. 5.2 discussion), and $\psi(q)$ is the Lagrangian displacement field. The cosmological parameter θ dependence enters via δ_1 and ψ, but we drop this from the notation to avoid clutter.

We compute ψ with `FlowPM` using 5 steps between redshift 9 and 0. We use $N_{\rm cdm} = 128^3$ matter particles, and force grid with resolution $N_{\rm grid} = 128^3$. Note the differences in the forward model and the data generation described in Sec. 5.2.

We define the model error as the 'stochastic' or 'shot noise' term,

$$\epsilon(\boldsymbol{k}) = \tilde{\delta}_h^{\rm model}(\boldsymbol{k}) - \tilde{\delta}_h^{\rm data}(\boldsymbol{k}). \tag{5.15}$$

The variance on the the overdensity model is thus given by the error power spectrum, defined as

$$\sigma^2_{\tilde{\delta},{\rm model}}(k) = P_{\rm err}(k) = \frac{1}{N_{\rm modes}(k)} \sum_{\boldsymbol{k}:|\boldsymbol{k}|=k} |\epsilon(\boldsymbol{k})|^2, \tag{5.16}$$

where $N_{\rm modes}(k)$ in the number of modes in the k bin.

The integrand in the square brackets of Eqn. 5.14 is equal to the Lagrangian field at q, namely $\delta_h^L(q) = 1 + b_1\delta_1(q)$. One could extend this relation to include higher order bias terms (see e.g. [319, 318, 317]), however for the number density of tracers considered here, we did not find improvement from using higher order terms and thus a linear bias model is sufficient for the purpose of this work. We note that the choice of model can bias the reconstruction in various ways (see [273] for a detailed review).

We fit the bias parameter b_1 before performing reconstruction by minimizing $|\epsilon(\boldsymbol{k})|^2$, from Eqn. 5.15, as in [319]. Note we only fit using scales with $k \leq N_{\rm cdm}/L$; this choice is

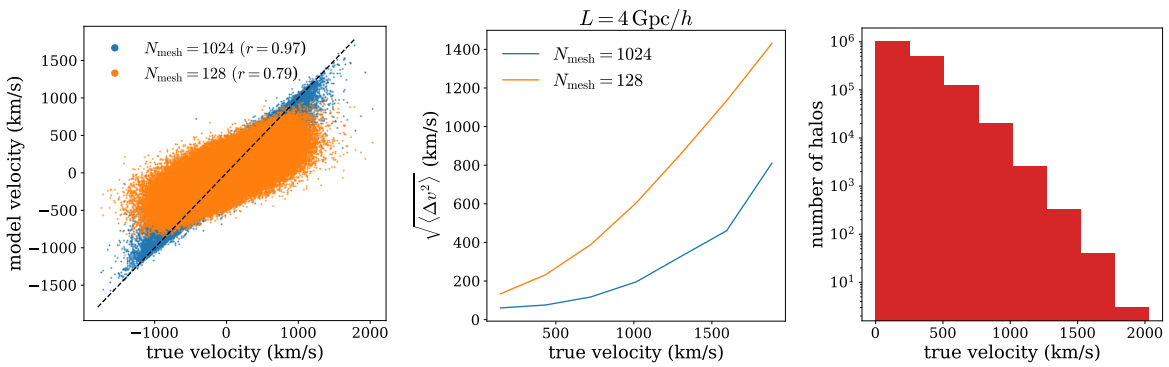

Figure 5.3: Velocity information for fiducial ($4\,{\rm Gpc}/h$) setup. Left: scatter plot of the model velocity compared to the true (data) velocity for an $N_{\rm mesh} = 1024$ (blue) and $N_{\rm mesh} = 128$ (orange) forward model. Middle: the corresponding root squared difference between the truth and model velocities. Right: halo abundances as a function of velocity.

somewhat arbitrary, but is chosen because the bias model breaks down at high k (see e.g. [319, 273]). We then fix the bias parameter to this best fit value throughout the analysis – in a full analysis one would infer the bias parameters, along with the cosmological parameters, while performing the reconstruction. We show the best fit model for our fiducial $4\,{\rm Gpc}/h$ and $400\,{\rm Mpc}/h$ setups (described in Section 5.2) in Fig. 5.2. It can be seen that the cross correlation and transfer function between the model and truth is considerably less than unity in regions of high shot noise. This occurs on all scales for the bigger box, while only on small scales for the small box. This is expected since we are trying to fit a discrete tracer field with a continuous bias model field, which limits the ability of density-based reconstruction. Note that one could also consider a neural network forward model for the halo overdensity field [266] to model individual discrete objects, but we do not pursue that here. The velocity-based reconstruction however does not suffer from this noise.

Peculiar velocity forward model

For the velocity forward model we assume an unbiased forward model, i.e. we ignore halo velocity bias terms which should vanish on the large scales considered here due to the equivalence principle [33]. We thus assume that the halo velocity directly traces the underlying matter velocity field. This is not a complete description because, for example, halos may be in multistreaming regions (see e.g. [112, 116, 232]), but such effects are beyond the scope of this work. The matter velocity field can be computed using the final matter particle position and velocities from `FlowPM` as follows. First we define the momentum field at position $\boldsymbol{x}$ as,

$$\mathcal{V}(\boldsymbol{x}) = [1 + \delta(\boldsymbol{x})]v(\boldsymbol{x}). \tag{5.17}$$

This is evaluated using CIC interpolation to paint the overdensity field δ weighted appropriately by the velocities v. As noted above, in this work we assume that the velocity is only measurable along the line of sight, thus, we simply use v to denote the velocity projected along the line of sight. Given data corresponding to halo positions $\boldsymbol{x}_h^{\rm data}$ and peculiar velocities $v_h^{\rm data}$, the model velocity is computed as

$$f_v(\boldsymbol{x}_h^{\rm data}) \equiv v_h^{\rm model} = \frac{\mathcal{V}(\boldsymbol{x}_h^{\rm data})}{1+\delta(\boldsymbol{x}_h^{\rm data})}. \tag{5.18}$$

The error in this forward model has a strong dependence on the mesh resolution used in the CIC assignment scheme. Fig. 5.3 shows the scatter between the true halo velocity and model velocity (left), and the corresponding error variance on the model velocity as a function of the true velocity (middle). The model velocity corresponds to Eqn. 5.18 using $N_{\rm mesh} = 1024$ (blue) and $N_{\rm mesh} = 128$ (orange) for CIC. It can be seen that the model error is larger when using a coarser mesh. For $N_{\rm mesh} = 128$, the error approximately scales as $0.7 v_h^{\rm true}$, i.e. it is an 70% effect. On the other hand, for $N_{\rm mesh} = 1024$ the error is almost negligible for low velocity halos, and only a $\sim 20\%$ effect for the fastest halos. In this work, for tractable reconstruction, we use a $N_{\rm mesh} = 128$ for the forward model. Since the majority of halos have low velocities (Fig. 5.3 right panel) this is able to provide a sufficient quality of reconstruction for this work, however one can expect the results to improve as we use higher resolution forward models. We thus use an interpolated form of the orange line for velocity model error, $\sigma_{v,\rm model}$.

Optimization

Given the data and forward model, we maximize the posterior to get the MAP estimate of the initial field. Since the parameter space consists of $128^3 \approx 2$ million dimensions, we need to use optimization algorithms which make use of the gradient information that is readily provided by our differentiable PM code. In this work, we use the LBFGS-B algorithm [86] which uses gradients at each step, and additionally keeps track of them over the trajectory to approximate the Hessian with a low memory cost. As there is much noise, and many more modes to be fitted on small scales than large scales, we anneal the posterior as we optimize to iteratively fit the modes up to a give scales $k < k_{\rm iter}$. Mathematically, we multiply the density term in the loss from Eqn. 5.11 with a step function $A(k - k_{\rm iter})$ as follows,

$$\begin{aligned} -2\log p_{\rm iter}(z, d=\{\delta, v\}|\theta) = &\sum_{\boldsymbol{k}} \left\{ A(k-k_{\rm iter}) \frac{|\tilde{\delta}(\boldsymbol{k}) - f_{\tilde{\delta}}(\boldsymbol{k};\theta,z)|^2}{\sigma^2_{\tilde{\delta}}(\boldsymbol{k})} + \frac{|\tilde{z}(\boldsymbol{k})|^2}{P(k;\theta)} \right\} \\ &+ \sum_i \frac{[v_i - f_v(\boldsymbol{x}_i;\theta,z)]^2}{\sigma_v^2}, \end{aligned} \tag{5.19}$$

where $A(k) = 1$ if $k \leq 0$ and 0 if $k > 0$. We iteratively increase $k_{\rm iter}$ in steps of the fundamental frequency, $k_F = 2\pi/L$, up to some maximum value $k_{\rm max}$ beyond which convergence

breaks down. In this work we use $k_{\max} = 16k_F$. We note that the cutoff for scales smaller than $k_{\max}$ could affect the quality of reconstruction on scales larger than the cutoff [273], but such effects are sufficiently small to not affect the conclusions of this work. A similar annealing approach has been applied and studied in [266]

To ensure the optimizer has sufficiently converged we average over 5 datasets, each starting at a different initial guess for the initial modes. We refer the reader to [153] for a thorough analysis of the posterior surface and how the starting position of the optimizer can affect the reconstruction, but such effects are sufficiently small to not affect the conclusions of this work.

5.3 Results

We quantify the reconstruction accuracy in terms of the cross-correlation coefficient and transfer function between the reconstructed field and the data field. Because we consider only line-of-sight velocity data, we plot the reconstruction quality in three equally sized μ bins (with centers $\mu = 0.17, 0.5, 0.83$), where $\mu \equiv \cos\theta$ and θ is the angle to the line of sight for an infinitely far away observer (corresponding to the flat-sky approximation). Given the power spectrum of the data field $P_{\rm data}$, the power spectrum of the reconstructed field $P_{\rm recon}$, and the cross power spectrum $P_{\rm data,recon}$, the cross-correlation coefficient is given by

$$r(k,\mu) = \frac{P_{\rm data,recon}(k,\mu)}{\sqrt{P_{\rm data}(k,\mu)P_{\rm recon}(k,\mu)}}, \tag{5.20}$$

and the transfer function is given by

$$T(k,\mu) = \sqrt{\frac{P_{\rm recon}(k,\mu)}{P_{\rm data}(k,\mu)}}. \tag{5.21}$$

We note that the power spectrum in a (k,μ) bin is calculated as

$$P(k,\mu) = \frac{1}{N_{k,\mu}V} \sum_{\boldsymbol{k}\ :\ |\boldsymbol{k}|=k,\ \boldsymbol{k}\cdot\boldsymbol{k}_z=\mu} \delta(\boldsymbol{k})\delta(-\boldsymbol{k}), \tag{5.22}$$

where the sum runs over all wavevectors with magnitude k (plus/minus the bin width) and angle to line of site μ (plus/minus the bin width), $N_{k,\mu}$ is the number of modes in the (k,μ) bin, and V is the volume of the box.

Fig. 5.4 shows the reconstruction for the fiducial setup described in Section 5.2, namely: $4\,{\rm Gpc}/h$, $\bar{n} = 10^{-6}\,(h/{\rm Mpc})^3$, and $\sigma_v = 300\,{\rm km/s}$. We first focus on the top row which shows the reconstruction of the initial linear field. The middle panel shows that the reconstruction using only density data (dashed) is decorrelated on large and intermediate scales with a correlation value of approximately 0.6. Furthermore, the right panel shows the transfer

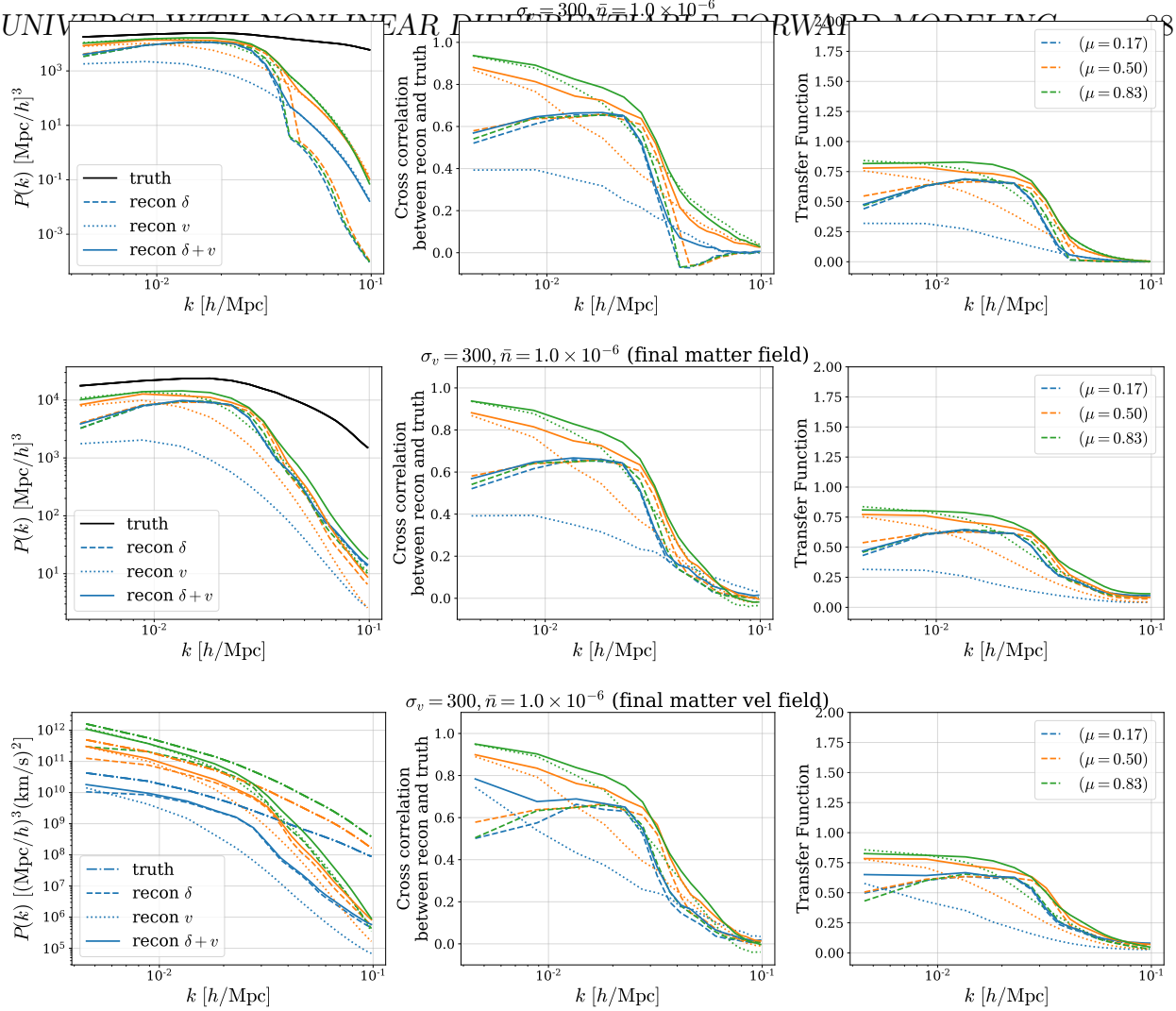

Figure 5.4: Reconstruction of the initial linear matter field, final matter field, final matter velocity field (top to bottom) for fiducial setup: $4\,\mathrm{Gpc}/h$, $\bar{n} = 10^{-6}\,(h/\mathrm{Mpc})^3$, and $\sigma_v = 300\,\mathrm{km/s}$. All future plots will perturb one feature of this. The left panel shows the true power spectrum (black), and the reconstructed power spectra using density-only (dashed), velocity-only (dotted), and joint density+velocity (solid). Three μ bins are considered, centered at $\mu = 0.17$ (blue), $\mu = 0.5$ (orange), and $\mu = 0.83$ (green). The middle panel shows the cross-correlation between the reconstructed and true fields, while the right panel shows the transfer function between the two.

function to be approximately 0.5 on the largest scales. The poor reconstruction on small scales is due to the annealing described in Sec. 5.2 – as we only anneal to $k_{\mathrm{max}} = 16k_F$ smaller scales are never fitted. On the other hand, doing velocity-only reconstruction (dotted) greatly improves the large-scale correlation in the two highest μ bins, while the $\mu = 0.17$ performs worse than density-only reconstruction. The story is similar for the transfer function which

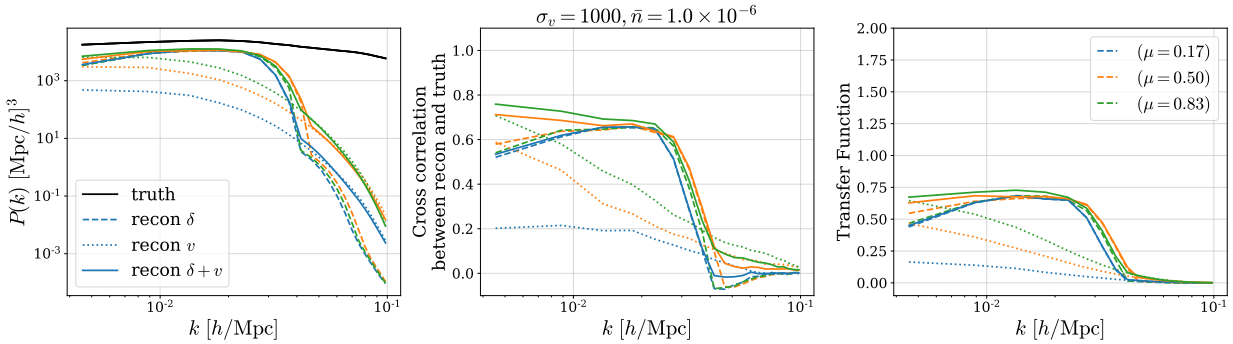

Figure 5.5: Like top row of Fig. 5.4 but with a higher velocity noise of $\sigma_{v,\mathrm{data}} = 1000\,\mathrm{km/s}$.

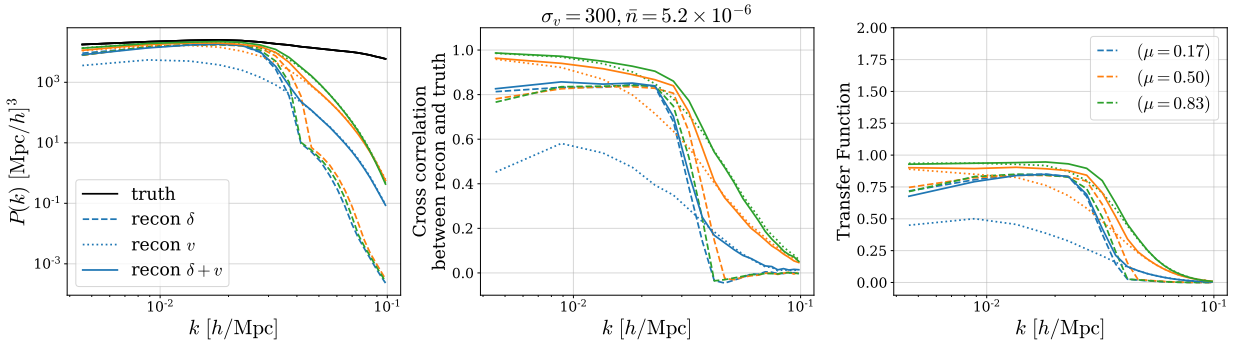

Figure 5.6: Like top row of Fig. 5.4 but with a ≈5 times higher number density of $\bar{n} = 5.2 \times 10^{-6}\,(h/\mathrm{Mpc})^3$.

is increased to 0.8 in the highest μ bin on large scales. Finally we consider reconstruction using the joint density and velocity data, using the full posterior of Eqn. 5.13 (solid). It can be seen that this further improves the correlation and transfer function on intermediate scales. The second and third rows of Fig. 5.4 show the reconstruction of the final matter density and velocity fields respectively. In both cases the reconstruction quality is of similar quality to the initial field.

Fig. 5.5 considers increasing the error in the velocity data from 300 to 1,000 km/s. In this case the density reconstruction is unchanged compared to the fiducial. The velocity-only reconstruction now performs worse than in the fiducial setup and is only slightly better than density-only for both the correlation coefficient and transfer function along the line of sight. Combining velocity with density slightly improves upon density-only reconstruction along the line of sight.

Fig. 5.6 considers increasing the number of halos in the fiducial setup by a factor of

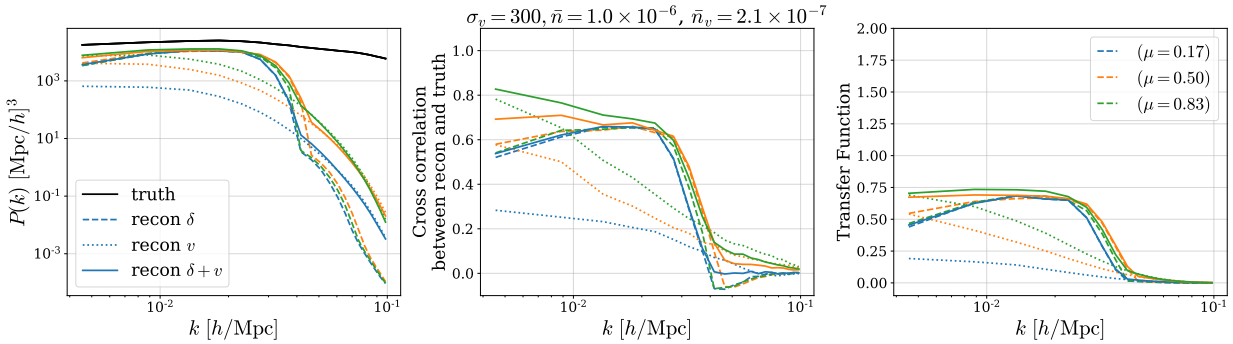

Figure 5.7: Like top row of Fig. 5.4 but with ≈5 times fewer halos with peculiar velocity data.

approximately 5 by lowering the minimum mass cut; the number density is now $\bar{n} = 5.2 \times 10^{-5}\,(h/\mathrm{Mpc})^3$ and the minimum halo mass $M_{\mathrm{min}} \simeq 1.2 \times 10^{15} M_\odot/h$. Given the higher number density, there is now lower shot noise and density-only reconstruction performs better compared to the fiducial case. Velocity-only reconstruction also benefits from the inclusion of additional halos, as there is now 5 times more halo data to use for reconstruction. The correlation coefficient and transfer function are now unity on the largest scales in the two highest μ bins.

So far we have assumed all halos used for density reconstruction also have velocity data. This is the case for kSZ, however, galaxy surveys are typically only able to measure the velocities of the most massive halos. Fig. 5.7 considers the modification of the fiducial setup such that only the most massive 20% of halos have velocity data. It can be seen that the velocity, and thus joint, reconstruction performs worse, with the correlation coefficient and transfer function dropping to around 0.8 on the largest scales for largest μ bin.

Until now we have considered velocity reconstruction on large scales in the high shot noise regime. To study the effect on smaller scales we divide the box size of the fiducial setup by a factor of 10, giving a $400\,\mathrm{Mpc}/h$ box. We also consider a lower velocity noise of $\sigma_{v,\mathrm{data}} = 100\,\mathrm{km/s}$. The top row of Fig. 5.8 shows the reconstruction of the initial linear field. It can be seen that there is a small gain from joint reconstruction compared to density-only in terms of the correlation coefficient and transfer function. We also plot the reconstruction of the final (nonlinear) matter density and peculiar velocity fields in the second and third rows of Fig. 5.8. The reconstruction of the final density can be seen to be of comparable quality to the initial field reconstruction. One the other hand, the reconstruction of the velocity field is improved on small scales by performing joint reconstruction compared to using density or velocity alone.

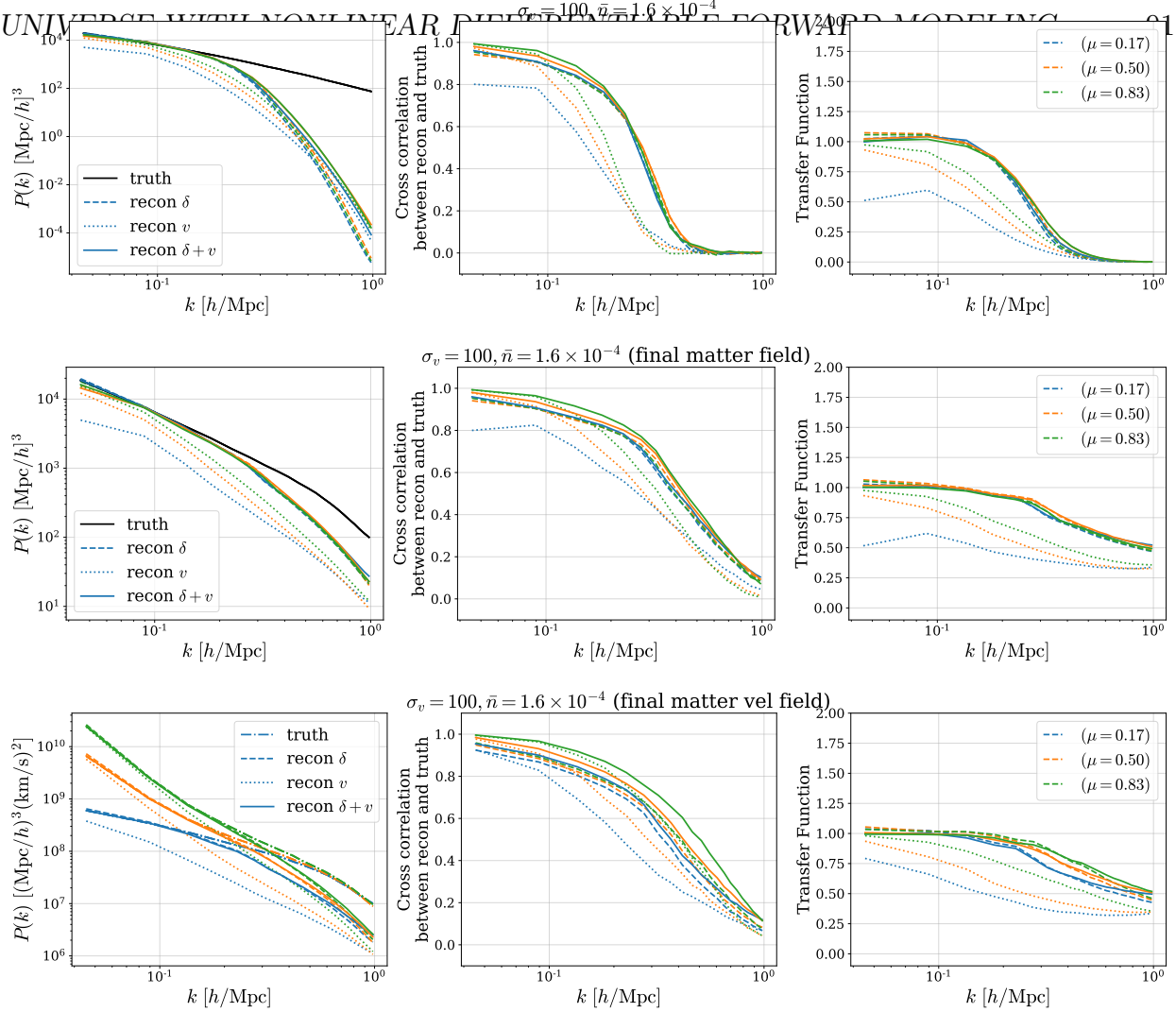

Figure 5.8: Like Fig. 5.4 but with 10 times smaller box size (number of halos different now due to different resolution of simulation), and $\sigma_v = 100$km/s.

5.4 Discussion and Conclusions

In this paper we have developed the formalism for including peculiar velocity information in field-level reconstruction of the initial conditions of the Universe. We have implemented it in the differentiable forward modelling code `FlowPM` to reconstruct the initial conditions using halo overdensity data, halo peculiar velocity data, and a combination of the two. We also considered the reconstruction of the final matter density and velocity fields. We showed that in shot noise dominated cases, reconstruction from density data alone is decorrelated from the truth, but this is greatly improved by including velocity data.

We studied this as a function of shot noise, error on velocity, and number of velocity

tracers. We found that the benefit of including velocity data is very much dependent on these quantities. We also showed that even in cases of low shot noise, combining velocity and density data, together with the non-linear model implemented in `FlowPM`, benefits the reconstruction of the final matter velocity field on nonlinear scales ($k > 0.1\,h/\mathrm{Mpc}$).

We expect this work to have wide applicability to future surveys: for example, upcoming observations by DESI [143], together with CMB maps from the Simons Observatory [360], will produce kSZ measurements with signal-to-noise $\gtrsim 100$: indeed the reconstruction from velocities is expected to dominate in statistical power for scales $k \lesssim 0.02h/\mathrm{Mpc}$ [341] and be even more powerful with CMB-S4 [3]. In addition, Rubin Observatory's LSST will discover hundreds of thousands of type Ia supernovae, for which a redshift can be obtained by DESI and individual peculiar velocities can be obtained with a few percent scatter [215]. In each case, forecasts based on linear theory suggest large improvements for measurement of cosmological parameters such as growth rate f [215] and local primordial non-Gaussianity f_{NL} [270]. This is in part due to the lower noise overall in the reconstruction of the initial conditions, and in part to the fact that galaxy positions and velocities trace the same underlying matter density, and therefore quantities like f and f_{NL} can be measured with reduced sample variance when combining the two measurements [326]. The joint reconstruction formalism developed in this work is a natural way to optimally combine the data available, and inference of cosmological parameters from this is an important next step, which is left to future work. To enable such future work, our code will be made publicly available upon publication of the paper.

Acknowledgements

We thank Boryana Hadzhiyska, Alex Kim, Uroš Seljak, and Kendrick Smith for very useful discussion. SF is funded by the Physics Division of Lawrence Berkeley National Laboratory and by the U.S. Department of Energy (DOE), Office of Science, under contract DE-AC02-05CH11231. This research used resources of the National Energy Research Scientific Computing Center, which is also supported by the Office of Science of the U.S. DOE under Contract No. DE-AC02-05CH11231.

5.5 Appendix A: Redshift Space Distortions

Here we repeat the analysis performed in the main paper for the fiducial example, but include redshift space distortions in the data and the modeling. Redshift space distortions include some velocity information and so it is appropriate to see what effect this has relative to the peculiar velocity information.

To transform the data to redshift space, assuming the flat sky approximation, we shift the configurations space coordinate of each halo by $v_z/(aH)$, where v_z is the velocity in the z direction. To model redshift space distortions, we make two alterations to the model in

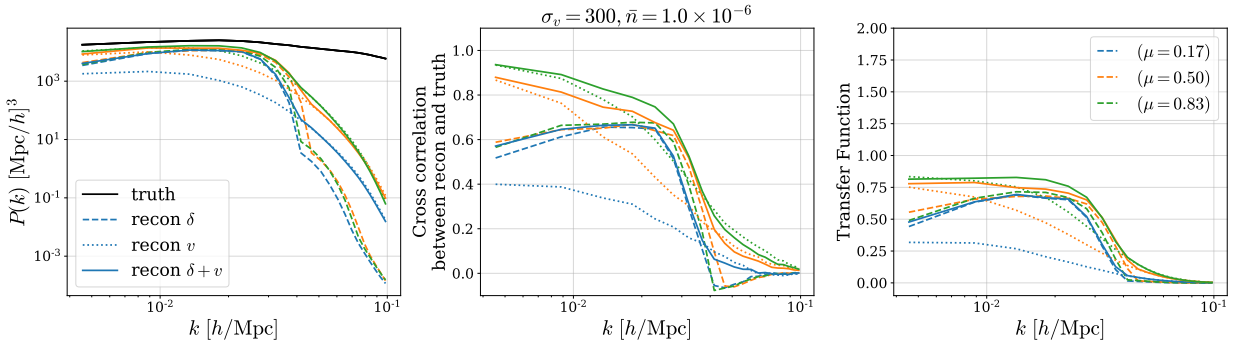

Figure 5.9: Like top row of Fig. 5.4 but with redshift space distortions.

the main paper. Firstly, while we use the same bias model as in the main paper, we now fit the bias parameters to the *redshift space* halo field (see [320] for application of the bias model to redshift space, note in particularly that the form of the model is the same to linear order, which is the order used in this work). Secondly, we displace the matter particles by their redshift space displacement, $v_z/(aH)$, before applying the bias model.

The results including redshift space distortions are shown in Fig. 5.9. By comparison to its configuration space analog, found in the top row of Fig. 5.4, it can be seen that the quality of reconstruction is comparable. This is to be expected for halos on the scales considered in this work, however redshift space distortions will become more important when considering lower mass objects.

Chapter 6

Super-sample covariance of the power spectrum, bispectrum, halos, voids, and their cross-covariances

The contents of this chapter was originally published in [63],

> Super-sample covariance of the power spectrum, bispectrum, halos, voids, and their cross-covariances
> *Bayer A.E. et al.* (arXiv:2210.15647) (submitted to PRD)

In this chapter, we study the effect of super-sample covariance (SSC) on the power spectrum and higher-order statistics: bispectrum, halo mass function, and void size function. We also investigate the effect of SSC on the cross-covariance between the statistics. We consider both the matter and halo fields. Higher-order statistics of the large-scale structure contain additional cosmological information beyond the power spectrum and are a powerful tool to constrain cosmology. They are a promising probe for ongoing and upcoming high precision cosmological surveys such as DESI, PFS, Rubin Observatory LSST, Euclid, SPHEREx, SKA, and Roman Space Telescope. Cosmological simulations used in modeling and validating these statistics often have sizes that are much smaller than the observed Universe. Density fluctuations on scales larger than the simulation box, known as super-sample modes, are not captured by the simulations and in turn can lead to inaccuracies in the covariance matrix. We compare the covariance measured using simulation boxes containing super-sample modes to those without. We also compare with the Separate Universe approach. We find that while the power spectrum, bispectrum and halo mass function show significant scale- or mass-dependent SSC, the void size function shows relatively small SSC. We also find significant SSC contributions to the cross-covariances between the different statistics, implying that future joint-analyses will need to carefully take into consideration the effect of SSC.

6.1 Introduction

Ongoing and upcoming cosmological missions such as DESI[1], PFS[2], Rubin Observatory LSST[3], Euclid[4], SPHEREx[5], SKA[6], and Roman Space Telescope[7] will probe ever larger volumes of cosmic structure in the small-scale, nonlinear regime. These data contain rich information that can be used to constrain fundamental physics, such as dark energy, dark matter, and neutrino mass. To fully realize the potential of these surveys, many higher-order (or non-Gaussian) statistics have been proposed to extract additional information beyond the power spectrum (2-point function). These include, for example, the bispectrum (3-point function), halo mass function, void size function, probability distribution function, marked power spectrum, and wavelet scattering transform [353, 323, 64, 213, 315, 245, 248, 211, 332, 255, 246, 242, 222, 120, 310, 254, 11, 178, 177, 129, 366, 20, 175, 188, 46, 257, 110, 109, 62, 223, 58, 371, 370, 147]. They have been studied intensively in recent years and are becoming standard tools for cosmological inferences. Moreover, there is increased interest in joint analysis, in which second and higher-order statistics are combined to maximize the information gain (see e.g. [62, 182, 146, 281, 142]).

Models for higher-order statistics usually rely on simulations for validation of analytic theories, calibration of semi-analytic models, or as the base of simulation-based inferences. To compute the covariance matrix of higher-order statistics, one typically requires a large set of simulations with different random initial conditions [383]. Such simulations assume periodic boundary conditions and are normally much smaller than the typical observed volumes of the Universe. Importantly, the mean density of these simulations is the cosmic one, and so by construction, they do not take into account the effects of perturbations with wavelengths longer than the size of the simulation. These so-called "super-sample modes" can however contribute sizeably to the covariance matrix; this effect is called the super-sample covariance (SSC) effect and must be carefully included to achieve accurate results.

To make contact with past literature, SSC has been studied as the "beat-coupling" (BC) effect in the mildly nonlinear regime using perturbation theory, as "halo sample variance" (HSV) in the highly nonlinear regime using the halo model, and was sometimes called the DC mode effect as an analogy between the constant background fluctuation and constant electric Direct Current [340, 169]. It was first studied in the context of the power spectrum [184, 352], and its effects have since been quantified using direct simulations [137], perturbation theory [355, 34, 54, 53], and separate universe simulations [240, 384, 359]. It has also been studied in relation to cluster counts [200, 315, 287], the matter bispectrum [323, 99, 50, 121], the matter one-point probability density function [367], the redshift space galaxy power

[1] https://www.desi.lbl.gov
[2] https://pfs.ipmu.jp/index.html
[3] https://www.lsst.org
[4] https://www.euclid-ec.org
[5] https://www.jpl.nasa.gov/missions/spherex
[6] https://www.skatelescope.org
[7] https://wfirst.gsfc.nasa.gov/index.html

spectrum [15, 14, 241], and the lensing power spectrum [52]. Furthermore, the effects of baryons on the SSC has been studied in [268, 56, 180]. The SSC effect associated with gravitational potential perturbations in cosmological with local primordial non-Gaussianity (i.e. $f_{\rm NL} \neq 0$) has also recently been studied [93]. Fast, approximate, methods exist to account for the SSC in forecasts for upcoming lensing and photometric surveys [226, 173, 227].

In this paper, we study the effects of SSC for the power spectrum and several higher-order statistics: the bispectrum, halo mass function, and void size function. To do so, we compare the statistics measured using small periodic boxes, which ignore SSC, to those using equally-sized boxes that are embedded in a much larger simulation, which include SSC. We study the effect of SSC in both the total matter field and the halo field. We also validate our results against the separate universe (SU) approach, in which the SSC contribution is calculated semi-analytically using the response of the statistics to certain changes in the cosmological parameters.

Our work is the first to investigate the effects of SSC for voids. Cosmic voids have been studied intensively in recent years [294, 80, 349, 208, 288, 277, 310, 118, 374, 222, 62, 223] and have achieved cosmological constraints with observational data [182, 181, 117, 78]. The bias parameters of voids have also been recently studied using SU simulations [205, 100]. Our work is also the first to study the effects of SSC on the cross-covariance between the combinations of all of these statistics.

The paper is organized as follows. Section 6.2 outlines the methods employed to compute the SSC of the power spectrum, bispectrum, halo mass function, and void size function. Section 6.3 presents the results for SSC of these statistics and their cross-covariances. We conclude in Section 6.4.

6.2 Method

In this section, we describe the methods used to run the N-body simulations, to compute the statistics, and to compute the SSC. We also briefly describe the SU approach.

Covariance

The covariance matrix between an observable $\mathcal{O}_\alpha$ and another observable $\mathcal{O}_\beta$ is given by

$$C_{\alpha\beta} = \langle (\mathcal{O}_\alpha - \langle \mathcal{O}_\alpha \rangle) (\mathcal{O}_\beta - \langle \mathcal{O}_\beta \rangle) \rangle \,, \tag{6.1}$$

where $\langle\rangle$ denotes the mean over realizations. The α and β subscripts can refer to different bins of particular statistic, or two completely different statistics. The covariance can be estimated by evaluating Eqn. 6.1 using an ensemble of simulations with different random realizations of the initial conditions.

We quantify the SSC effect by comparing the following two sets of simulations:

- **sub**-boxes that are embedded in a much larger simulation, where the effect of SSC is properly captured;
- **small** boxes that are of the same size and resolution as the sub-boxes, but are independently simulated with periodic boundary conditions and have no super-sample modes.

Because the super-sample modes are only present in the former and not in the latter, the SSC is given by

$$C_{\rm SSC} = C_{\rm sub} - C_{\rm small}, \tag{6.2}$$

where $C_{\rm sub}$ is the covariance computed using sub-boxes and $C_{\rm small}$ is that using small boxes.

N-body simulations

We use `FastPM` [154, 57], a particle-mesh (PM) N-body simulation, to simulate a big box of side-length $5\,{\rm Gpc}/h$ with 2048^3 matter particles. We then split this big box into $8^3 = 512$ sub-boxes. We compare these to 512 independent, periodic, small boxes of size $5000/8{=}625\,{\rm Mpc}/h$, each with 256^3 particles. The resolution of the small boxes is chosen to match the big box. We consider a maximum scale cut of $k_{\rm max} = 0.8\,h/{\rm Mpc}$ as scales with lower k than this are well modeled by our simulations. In all cases, we begin the simulations at $z = 9$ and take 60 steps to $z = 0$. The resolution of the force mesh is 2 times the number of particles. For simplicity, we consider and discuss only the results at $z = 0$, which is when SSC is expected to be the strongest.

Our cosmological parameters are $h = 0.6774$, $\Omega_m = 0.3089$, $\Omega_b = 0.0486$, $\sigma_8 = 0.8159$, $n_s = 0.9667$, $M_\nu = 0$.

To identify halos, we use the Friends-of-Friends (FoF) algorithm with a linking length of 0.2. We generate the particle and halo overdensity fields using the Cloud-in-Cell (CIC) method with $N_{\rm mesh} = 256$ using `nbodykit` [185]. Further, we consider the halo field in real and not in redshift space. We compensate the field for window effects before calculating the statistics [210]. For matter the overdensity field is computed as $\delta = (\rho - \bar{\rho})/\bar{\rho}$, where ρ is the matter density, while for halos it is computed as $(n - \bar{n})/\bar{n}$, where n is the number density. In order to compute the overdensity field in the sub-boxes, there are two choices for the mean density $\bar{\rho}$ (or $\bar{n}$): using the "**g**lobal" mean of the big box, or the "**l**ocal" mean of the sub-box. Realistically, for weak lensing surveys it is appropriate to use the global mean as the mean density can be directly calculated from the cosmological model, moreover, the measured weak-lensing shear field is sensitive to the global-mean density. However, for galaxy surveys, since one does not know how to predict from first principles the total number of galaxies, the local mean is what is most appropriate as we measure the galaxy statistics w.r.t. the observed galaxy number density in the survey. We consider both cases in our analysis.

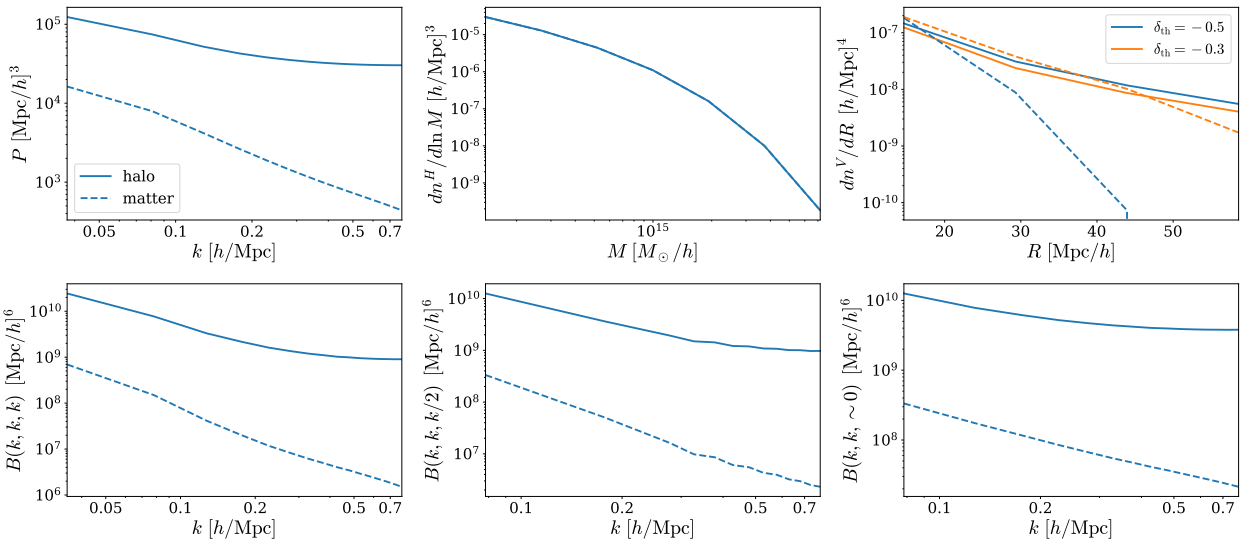

Figure 6.1: Power spectrum (top left), halo mass function (top middle), void size function (top right), and three bispectrum configurations (bottom) averaged over all small boxes. We show results for the halo (solid) and matter (dashed) field. For the void size function, we show results for two density thresholds used for void searching, $\delta_{\rm th} = -0.5$ (blue) and $\delta_{\rm th} = -0.3$ (orange).

Statistics

We now define the various statistics considered in this work, and the method used to compute them. We plot the statistics averaged over all the sub-boxes in Fig. 6.1.

- The **power spectrum** $P(k)$ is defined as the Fourier transform of the 2-point correlation function $\xi(x_1 - x_2) \equiv \langle \delta(x_1)\delta(x_2)\rangle$, where δ is the overdensity field. Defining the fundamental frequency of the box as $k_F \equiv 2\pi/625\,h$/Mpc, we use 15 linear bins between bin edges $[0,\, 80k_F]$, with $\Delta k = 5k_F$.

- The **bispectrum** $B(k_1, k_2, k_3)$ is defined as the Fourier transform of the 3-point correlation function $\langle \delta(x_1)\delta(x_2)\delta(x_3)\rangle$. We consider three particular configurations of the bispectrum: equilateral ($k_1 = k_2 = k_3$), isosceles ($k_1 = k_2 \neq k_3$), and squeezed ($k_1 = k_2, k_3 \sim 0$). More concretely, we consider the squeezed mode as $k_3 = 3.6 \times 10^{-2}\,h$/Mpc. We compute the bispectrum using the `bskit` package [162], which employs the FFT-based bispectrum estimators of [321, 322]. The k binning is the same as for the power spectrum.

- The **halo mass function** (HMF), denoted $dn^H/d\ln M$, is defined as the comoving number density of halos n^H per unit of log halo mass $\ln M$. We consider 7 logarithmic bins bounded by $M_{\rm min} = 10^{14} M_\odot/h$ and $M_{\rm max} = 10^{16} M_\odot/h$.

- The **void size function** (VSF), denoted dn^V/dR, corresponds to the comoving number density of voids n^V per unit of void radius R. We consider spherical voids in smoothed density fields. The $N_{\rm mesh} = 256^3$ field is smoothed with top-hat filters of size $R_{\rm filter}$ in 7 linear bins between $R_{\rm min} = 12.2\,{\rm Mpc}/h$ and $R_{\rm max} = 100.1\,{\rm Mpc}/h$, with $\Delta R = 6 d_{\rm grid}$ where $d_{\rm grid} \approx 2.44\,{\rm Mpc}/h$ is the grid size. We search hierarchically, first finding the largest voids and then the, more abundant, smaller voids. Voids of size $R = R_{\rm filter}$ are defined as local minima in the $R_{\rm filter}$-filtered field, with values lower than a predefined threshold $\delta_{\rm th}$, unless they overlap with existing larger voids. In this work we investigate thresholds of $\delta_{\rm th} = -0.3$ and -0.5. The void finding algorithm was developed by [47] and we use the implementation in Pylians3 [378].

Separate universe simulations

We now briefly summarize an alternative method to compute the SSC using Separate Universe (SU) simulations; we refer the reader to [240, 384] for more details. In this approach, the effect of a super-survey density mode that is constant inside the box and has amplitude δ_b is mimicked by adjusting the cosmological parameters such that $\bar{\rho}_m \to \bar{\rho}_m\,(1+\delta_b)$, where $\bar{\rho}_m$ is the mean physical matter density; if the fiducial cosmology is a spatially flat universe, this implies that the separate universe has non-zero curvature ($\Omega_k \neq 0$). The response of any summary statistic $\mathcal{O}$ to δ_b is computed by considering the difference between simulations run with different δ_b. The SSC is then approximated by

$$C^{ij}_{\rm SSC-SU} \simeq \sigma_b^2 \frac{d\mathcal{O}_i}{d\delta_b}\frac{d\mathcal{O}_j}{d\delta_b}, \tag{6.3}$$

where σ_b^2 is the variance of the linear matter density fluctuations on the size of the survey described by a window function W,

$$\sigma_b^2 \equiv \frac{1}{V_W^2}\int \frac{d^3\boldsymbol{k}}{(2\pi)^3}|W(\boldsymbol{k})|^2 P_{\rm linear}(k), \tag{6.4}$$

where $V_W = \int d^3\boldsymbol{x}\; W(\boldsymbol{x})$ is the survey volume and $P_{\rm linear}(k)$ is the linear matter power spectrum. The window function used in this work corresponds to a 3d cube of side-length $625\,{\rm Mpc}/h$, giving $\sigma_b^2 = 6.8\times10^{-5}$. Concretely, we evaluate the responses using finite difference methods on simulations with $\delta_b = \pm 0.03$, and averaging over 20 realizations of the initial conditions.

The δ_b mode modifies the background expansion history, which implies some care when choosing the box size of the simulations with $\delta_b \neq 0$. In our simulations here, we choose the comoving box size to match at all times in Mpc units. This corresponds to the "growth-dilation" methods in the notation of [240], or equivalently, with our SU simulations we measure the so-called "growth-only" responses in the language of [384]. Importantly, when identifying halos in the simulations, the FoF linking length in the separate universe needs to be rescaled by the ratio of the scale factors in the two simulations to guarantee matching halo definitions (see e.g. [55] for a discussion.)

Since the simulations with different δ_b values can be initialized with the same random phases of the initial conditions, the SU approach has the significant advantage of converging with much fewer simulations than the sub-box approach (discussed in Section 6.2). We note, however, that our SU simulations account only for the impact of isotropic density perturbations as super-survey modes, i.e., they do not account in particular for the effect of super-survey tidal fields [316, 256, 13]. Here, we consider angular averaged spectra in real space, for which the impact of super-survey tidal fields averages out, but we note that for analyses in redshift space [15, 241, 14] and weak-lensing applications [52] this is not the case and the super-survey tidal fields can have a non-negligble effect. Further, the super-survey tidal fields contribute also to the SSC effect of halo and void counts, although in a weak manner since this happens only at second order. On the other hand, the SSC calculated using the sub-box approach automatically includes both the effects of density and tidal fields.

6.3 Results

Here we show the effect of SSC for individual statistics as well as their cross-covariances. In all plots, error bars are computed using bootstrapping and correspond to the 95% confidence interval.

Matter field statistics

Fig. 6.2 shows the results of the SSC contribution for the power spectrum, void size function, halo mass function, and three bispectrum configurations. Each statistic contains two panels, the upper of which shows the diagonal term of the covariance computed with and without SSC, using sub- and small boxes, respectively. For sub-boxes, we show results using both the local mean and the global mean density. The lower panel shows the ratio between the sub and small box, which is equal to $C_{\rm sub}/C_{\rm small} = 1 + C_{\rm SSC}/C_{\rm small}$ (using Eqn. 6.2). Shown also is the result from the SU approach (marked by the red crosses and pluses for the local and global mean cases respectively), which agrees reasonably well overall with the SSC estimated from the sub-box approach.

For the case of our power spectrum and bispectrum results, we note also that while the SSC does not depend to first order on the size of the wavenumber bins, other contributions to the covariance typically do, which can have an impact on the exact values of $C_{\rm sub}/C_{\rm small}$ (note this does not mean there is a dependence of the signal-to-noise on the bin size). This does not have an impact on the main takeaway points of our results, but it is useful to keep in mind especially when comparing quantitatively to results obtained previously in the literature.

Power spectrum

The SSC of the matter power spectrum can be seen to be a just under a 100% effect at $k = 0.7\,h/\mathrm{Mpc}$ in the local mean case. However, the SSC is much larger when using the

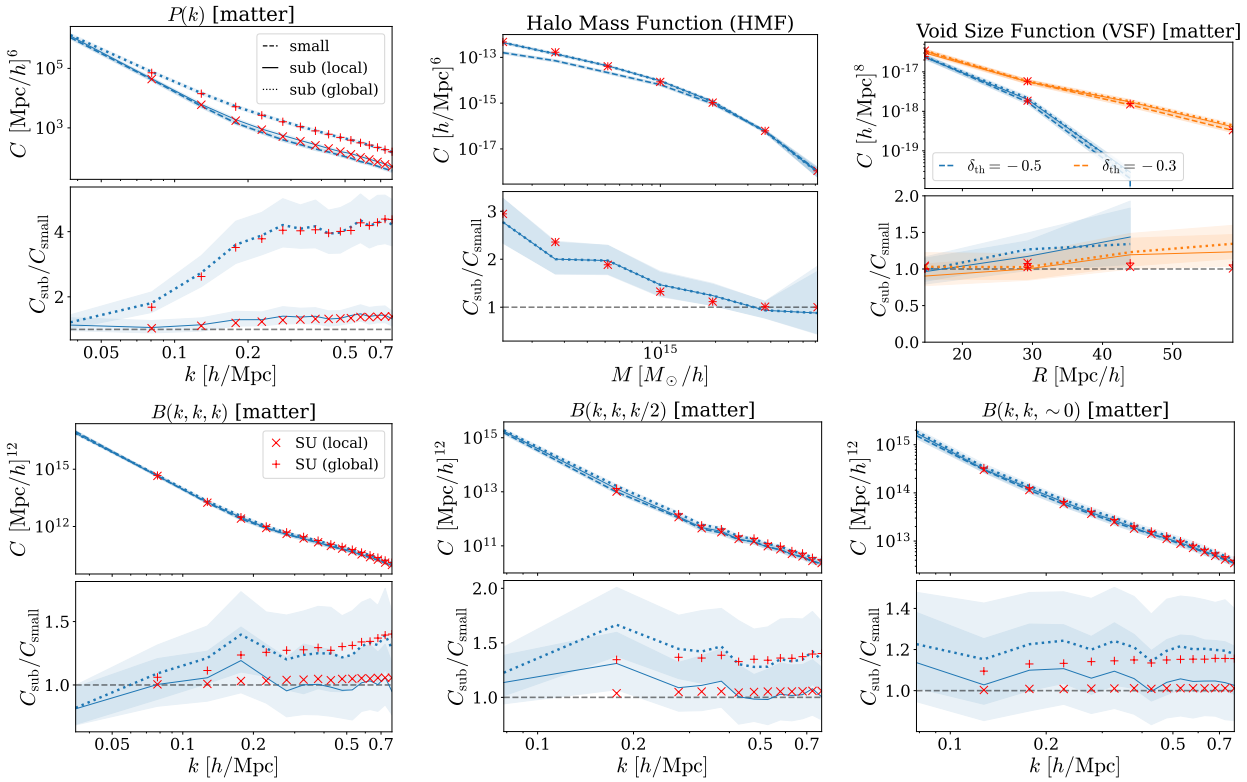

Figure 6.2: Covariances for the power spectrum (top left), halo mass function (top middle), void size function (top right), and three bispectrum configurations (bottom) in the matter field. Each statistic contains two panels, the top of which shows the diagonal term of the covariance computed in the small box (dashed), and in the sub-box using local (solid) and global (dotted) mean. The lower panel shows the ratio between the sub- and small boxes, where the dashed horizontal grey line indicates no SSC. For the void size function, we show results for two density thresholds used for void searching, $\delta_{\rm th} = -0.5$ (blue) and $\delta_{\rm th} = -0.3$ (orange). We also show separate universe results for SSC computed using local and global mean densities, marked in "x" and "+", respectively. Shaded regions correspond to 95% confidence intervals.

global mean, with the ratio $C_{\rm sub}/C_{\rm small}$ increasing with k to a factor of ~ 4 at $k = 0.7\,h/{\rm Mpc}$. This can be explained as follows: the local mean density is modified with respect to the global mean by the background density δ_b, as $\bar{\rho}_{\rm local} = \bar{\rho}_{\rm global}(1+\delta_b)$. Thus the power spectrum with respect to the local mean is given by $P_{\rm local}(k) = (1+\delta_b)^{-2} P_{\rm global}(k)$, where $P_{\rm global}(k)$ is the power spectrum with respect to the global mean. The local and global responses are then related as,

$$\frac{d\ln P_{\rm local}(k)}{d\delta_b} \approx \frac{d\ln P_{\rm global}(k)}{d\delta_b} - 2, \tag{6.5}$$

where we use the fact that $\delta_b = 0.03 \ll 1$. It can be shown with perturbation theory that the global response is close to 2 for the scales considered in this paper [137, 352, 240], hence

the local response is much suppressed in comparison to the global. Recall, the SSC of the power spectrum referenced to the global density is what is relevant to weak lensing analysis, and this strong response is ultimately the reason why SSC is the most important piece of the off-diagonal covariance in cosmic shear 2-point function studies [52, 51].

Halo Mass Function

We find the SSC has very little contribution to the counts of massive halos $\gtrsim 10^{15} M_\odot/h$, while it increases towards less massive halos, with the ratio $C_{\rm sub}/C_{\rm small}$ becoming roughly a factor of 3 for masses of $\sim 10^{14} M_\odot/h$. The fluctuation in the number density of halos $\delta n = \delta n(M)$ in a mass bin M is a biased tracer of the underlying matter field δ_m

$$\frac{\delta n}{\bar{n}} = b\delta_m, \tag{6.6}$$

where $b = b(M)$ is the halo bias in the mass bin. The diagonal term of the sub-box HMF covariance divided by the shot noise $C_{\rm small} = \bar{n}/V$ is thus [200],

$$C_{\rm sub}/C_{\rm small} = 1 + \sigma_b^2 b^2 \bar{N}, \tag{6.7}$$

where $\bar{N} = \bar{n}V$ is the number (or abundance) of halos. The second term on the right is the SSC contribution. While massive halos tend to be more biased (by a factor of few compared to low mass halos), their abundance is exponentially suppressed. Thus, the most massive halos are in the shot-noise dominated regime with little contribution from SSC.

Void size function

Recall, we consider spherical voids with density thresholds $\delta_{\rm th} = -0.5$ and -0.3, whose results are shown on the top right panel of Fig. 6.2. The SSC contribution for voids is generally small for all void radii shown. Following from the discussion above for the HMF, this is as expected since voids are approximately 100 times less abundant than the halos, and their bias values remain of order $0 - 10$ [205, 100]. Thus, the covariance is strongly dominated by shot noise and the SSC effect is negligible. Note further that for certain void radii, and contrary to the case of halos, the void bias can be zero, in which case even an infinite abundance would have no SSC.

Bispectrum

All three configurations of the bispectrum that we consider have a much smaller SSC than the power spectrum, typically a factor 1–2 effect for $k \lesssim 0.7h/\text{Mpc}$. Similarly to the power spectrum, the global case has a larger effect than local. This agrees within error bars with [99], which considered the equilateral and isosceles cases using an analytical response approach and simulations. Also [121] found the SSC effect to be small for the matter bispectrum. Our results are also in good agreement with [50] which analytically derived that the SSC effect is small for the squeezed matter bispectrum.

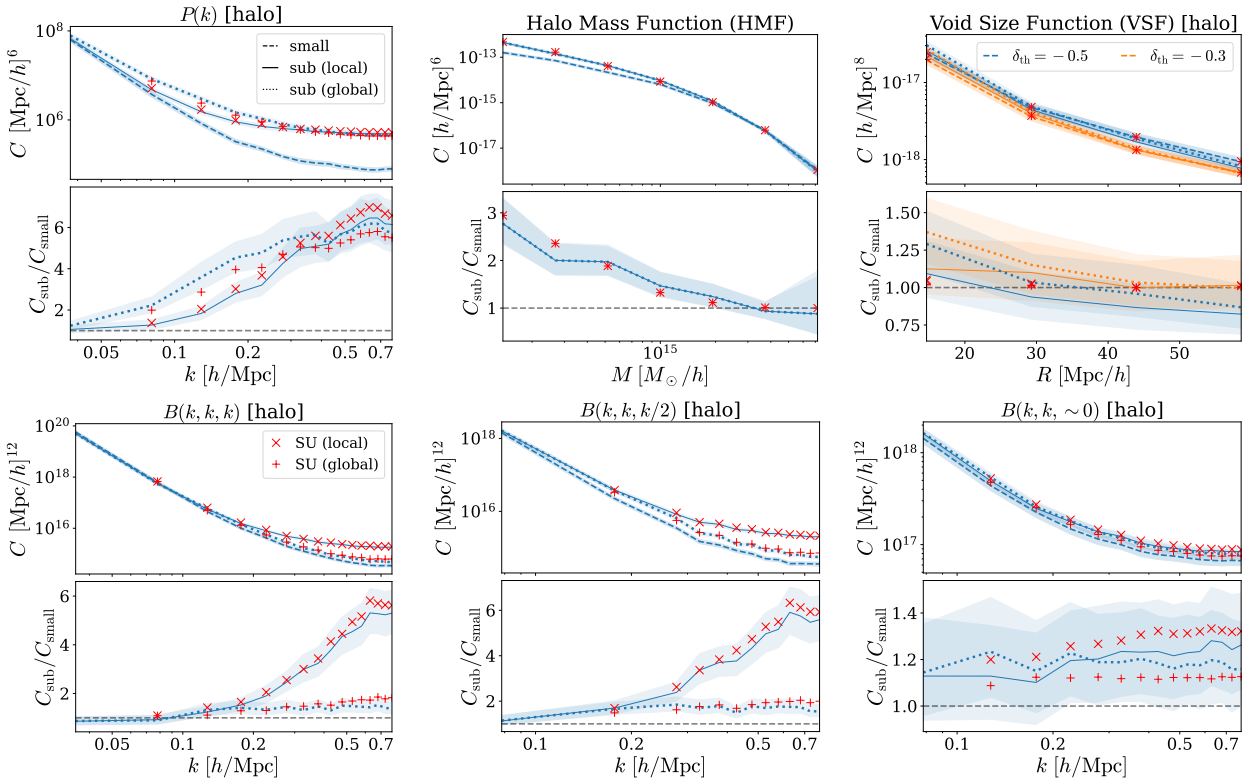

Figure 6.3: Same as Fig. 6.2 but for the halo field, with a minimum mass cut $M_{\rm min} = 10^{14} M_\odot/h$. The halo mass function plot (top middle) is duplicated here for completeness.

Halo field statistics

Fig. 6.3 shows the results for statistics computed from the halo field, which is a biased tracer of the matter field. Here, we apply a minimum mass cut of $M_{\rm min} = 10^{14} M_\odot/h$. The halo field SSC shows qualitative similarities to that of the matter field, namely larger SSC on smaller scales for the power spectrum and the bispectrum, and a small SSC for voids.

One noteworthy difference concerns the relative size of the local and global cases for the power spectrum, which is now comparable. This is because the local and global responses for halos are related as

$$\frac{d \ln P^h_{\rm local}(k)}{d\delta_b} \simeq \frac{d \ln P^h_{\rm global}(k)}{d\delta_b} - 2b. \tag{6.8}$$

Unlike in the case of the local matter response (Eqn. 6.5), where the global response nearly cancels with the -2 term, the global halo response is corrected by $-2b_1$. In our work the bias is $b_1 \approx 2.5$, which leads to a negative local halo response (see Fig. 12 of [36] and derivation therein), so much so that the local effect is now comparable to the global effect after taking the square of the response in Eqn. 6.3. The exact value of the response is mass and redshift

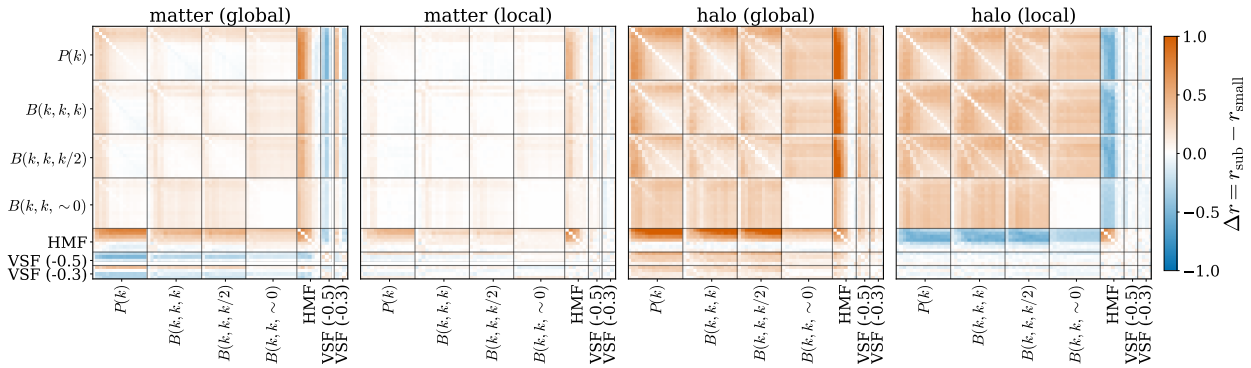

Figure 6.4: Difference in correlation coefficient between the sub and small box $\Delta r \equiv r_{\rm sub} - r_{\rm small}$. From left to right: matter global, matter local, halo global, halo local.

dependent, as halo bias increases when considering more massive halos and/or halos at higher redshift. Analogous considerations apply to the case of the halo bispectrum.

Cross-covariance

In this section, we study the SSC contribution to the cross-covariance of the statistics. Higher-order statistics typically have large off-diagonal terms in the covariance compared to the power spectrum and are usually studied jointly with other statistics. Therefore, it is important to study not only the variances of individual statistics, but also their cross-covariance. To focus on the off-diagonal terms, we normalize the covariance using the diagonal term to obtain the correlation matrix

$$r_{ij} \equiv C_{ij}/\sqrt{C_{ii}C_{jj}}, \tag{6.9}$$

where C is the covariance matrix with indices run through the bins of all the statistics studied here. Fig. 6.4 shows the difference in correlation coefficient between the sub- and small box, $\Delta r \equiv r_{\rm sub} - r_{\rm small}$, for the both the matter and halo fields, using local and global mean densities.

For the power spectrum and the bispectrum, the SSC contribution to the cross-covariance is positive in all cases. The amplitude is the smallest in the local matter case, due to the local mean response cancellation discussed in Sec. 6.3. The local and global halo cases see comparable contributions, also similarly to what was observed for the diagonal term in Sec. 6.3.

For the HMF, we see a large contribution of SSC to the cross-covariance with other statistics. The effect is positive in all cases, except for the halo local-mean case.

For the VSF, we observe a relatively small contribution of SSC to the cross-covariance with other statistics, except for the global-mean matter field. This is consistent with the observation of almost negligible SSC in VSF variances in Sec. 6.3, where we discussed that the SSC is low due to the low abundance of voids.

6.4 Conclusions

We study the effect of SSC on the power spectrum, bispectrum, halo mass function, and void size function, as well as on the cross-covariance between them. We consider both the total matter and halo fields. We compare the covariance that includes the SSC (computed using 625 Mpc/h *sub*-boxes of a 5 Gpc/h simulation), to the covariance without (computed using 625 Mpc/h *small* periodic boxes). We now summarize our main results together with additional discussion:

- We see an increasing impact of SSC on smaller scales for the matter and halo power spectrum, reaching a factor of a few compared to the covariance that ignores SSC from $k \approx 0.2h$/Mpc and beyond. The exception is the case of the power spectrum referenced to the local-mean density, for which the effect is less than 100%. This is in agreement with previous studies [184, 137, 352, 240, 355].

- For the halo mass function, the SSC has little effect on massive halos above $10^{15} M_\odot/h$, as they are dominated by shot noise. However, the effect of SSC increases to a factor of 2–3 for lower mass halos, consistently with previous studies [200, 315, 287].

- For the void size function, we found a relatively small SSC effect, due to the low number density of voids compared to that of halos for a given survey volume, and also the low bias. This is an attractive feature of voids, making their covariance simpler to approximate without considering the SSC. In this work we considered spherical voids, however one might consider different void finding algorithms, such as `VIDE` [350]. However, as long as the void abundance and bias are comparable to those we considered, our general conclusions should hold for other void definitions as well.

- While the *m*atter bispectrum receives less SSC contribution, an approximately 50% effect, in good agreement with previous studies [99, 50], the *h*alo bispectra shows dependence on the bispectrum configuration and the choice of local or global mean density. Concretely, the level of SSC remains low (≈10% level) for all three halo bispectrum configurations (equilateral, isosceles, and squeezed) when using the global-mean density and for squeezed bispectrum using the local mean. However, the SSC contribution becomes a factor of a few for the equilateral and isosceles halo bispectra referenced to the local mean.

- For the cross-covariances, we see non-negligible contribution of SSC, in particular for the halo field statistics. We also observe a negative effect of SSC (or reducing the off-diagonal terms) for HMF × other statistics in the halo-local mean case, and VSF × other statistics in some radius bins. This indicates the importance of including the effect of SSC in future joint-statistic analysis.

In summary, our work shows that future cosmological analyses with the power spectrum and higher-order statistics, as well as their joint analysis, should need to carefully consider

the effect of SSC. Analyses where many higher-order statistics are combined can have large data vector sizes, which puts pressure on simulation-based methods for the covariance because of the need to have a sufficiently converged covariance matrix that is stable under inversion (which is what is needed in parameter inference analyses). This therefore strongly motivates more simulation-based works like ours here towards a robust understanding of the super-sample covariance and cross-covariance of higher-order statistics. The level of impact of SSC depends in particular on the box/survey size, halo sample and redshift, which would be interesting to investigate with more detail. In the future, it would also be interesting to assess the impact of the SSC on these statistics at the level of final parameter posteriors in simulated likelihood inference analyses for ongoing and future surveys. It would also be fruitful to use our simulations to quantify the SSC of other higher-order statistics. To enable such future works, our simulations will be made publicly available upon publication of the paper.

Acknowledgements

We thank Masahiro Takada, Eiichiro Komatsu, Yue Nan, Uroš Seljak, James Sullivan, and Francisco Villaescusa-Navaro for insightful discussion. AEB thanks Kavli IPMU for hosting him during the duration of this project. AB acknowledges support from the Excellence Cluster ORIGINS which is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy - EXC-2094-390783311. This research used resources of the National Energy Research Scientific Computing Center (NERSC), a U.S. Department of Energy Office of Science User Facility located at Lawrence Berkeley National Laboratory, operated under Contract No. DE-AC02-05CH11231 using NERSC award HEP-ERCAP0023125. This work was partly performed at the Aspen Center for Physics, which is supported by National Science Foundation grant PHY-1607611. We use `FastPM` [154, 57] to simualte large-scale structure. We use `nbodykit` [185] to compute overdensity fields and power spectra. The bispectrum is computed using the `bskit` package [162], which employs the FFT-based bispectrum estimators of [321, 322]. We use Pylians3[8] [378] to find voids and compute the void size function.

[8] https://pylians3.readthedocs.io/

Chapter 7

The look-elsewhere effect from a unified Bayesian and frequentist perspective

The contents of this chapter was originally published in [60],

> The look-elsewhere effect from a unified Bayesian and frequentist perspective
> *Bayer A.E., Seljak U.* (arXiv:2007.13821) JCAP 10 (2020) 009

When searching over a large parameter space for anomalies such as events, peaks, objects, or particles, there is a large probability that spurious signals with seemingly high significance will be found. This is known as the look-elsewhere effect and is prevalent throughout cosmology, (astro)particle physics, and beyond. To avoid making false claims of detection, one must account for this effect when assigning the statistical significance of an anomaly. This is typically accomplished by considering the trials factor, which is generally computed numerically via potentially expensive simulations. In this chapter, we develop a continuous generalization of the Bonferroni and Šidák corrections by applying the Laplace approximation to evaluate the Bayes factor, and in turn relating the trials factor to the prior-to-posterior volume ratio. We use this to define a test statistic whose frequentist properties have a simple interpretation in terms of the global p-value, or statistical significance. We apply this method to various physics-based examples and show it to work well for the full range of p-values, i.e. in both the asymptotic and non-asymptotic regimes. We also show that this method naturally accounts for other model complexities such as additional degrees of freedom, generalizing Wilks' theorem. This provides a fast way to quantify statistical significance in light of the look-elsewhere effect, without resorting to expensive simulations.

7.1 Introduction

A common problem in statistical analysis is to find evidence for a physical signal in a large, continuous parameter space, where the true position of the signal is not known a priori. By searching over a wide parameter space one increases the probability of finding large signals caused by random statistical fluctuations, as opposed to a physical source. This is known as the look-elsewhere effect – or sometimes the "problem of multiple comparisons" in discrete cases – and must be accounted for when performing a hypothesis test [263, 331]. Ignoring this effect would lead to an overestimation of the statistical significance, sometimes by a considerable amount, and thus incorrectly concluding the detection of a physical signal.

The look-elsewhere effect is prominent throughout (astro)particle physics and cosmology. One of the most commonly known occurrences is in collider searches for new particles, for example it was a key consideration in the Higgs boson discovery [1, 102]. In this example, one searches a large range of masses for a resonance, without a priori knowledge of the true mass of the particle. Similarly, in astrophysical searches for particles one seeks resonances in the energy flux of various astrophysical spectra, where the true energy signature of the particle is unknown. Examples include: constraining the dark matter self-annihilation cross-section via gamma ray emission from galaxy clusters [23], searching for WIMPs via charged cosmic rays [302], searching for non-baryonic dark matter via X-ray emission from the Milky Way [324], and explaining the source of high energy astrophysical neutrinos [2, 151]. In terms of cosmology, the look-elsewhere effect occurs in searches for gravitational wave signals from black hole or neutron star mergers [88, 4, 261]. Here one searches large time series for a signal, where the time and shape of the event are unknown. A further cosmological example is searching for signatures of inflation in the primordial power spectrum [155, 156, 201].

The look-elsewhere effect is also prevalent in other areas of physics and beyond, for example: in astronomy it occurs when detecting exoplanets via stellar photometry, where the period and phase of the planets' transits are unknown (e.g. [307]); in biology it occurs when considering large DNA sequences to study genetic association [32, 348]; in medicine it occurs when testing the effectiveness of drugs in clinical trials [298]; and in theology it occurs when attempting to find hidden prophecies in religious texts [259]. Therefore, given the apparent ubiquity of the look-elsewhere effect, there is much motivation for a fast method to account for it.

Many simple general methods exist to mitigate for the look-elsewhere effect in the case of discrete problems, for example if one is testing multiple drugs for their effectiveness at treating a disease [298]. The number of drugs tested, more generally known as the trials factor, quantifies the extent of the look-elsewhere effect. The larger the trials factor, i.e. the more drugs tested, the larger the chance of a false positive arising due to a statistical fluctuation. Methods such as the Bonferroni correction [77] and Šidák correction [338] use the trials factor to correct the conclusions of a hypothesis test in light of this effect. There is however no unique definition of the trials factor when searching a continuous parameter space for a signal, making it unclear how to implement these corrections in such cases. Therefore, a common, brute-force approach to account for the look-elsewhere effect in continuous problems

is to perform many simulations of an experiment assuming there is no signal. One can then estimate the p-value of a chosen test statistic, usually related to the maximum likelihood, and in turn define a relation between the significance of a signal and the test statistic. This means that to conclude a detection at the 5-sigma level, corresponding to a p-value of order 10^{-7}, one would need to simulate more than $\sim 10^7$ realizations of the experimental data, which is computationally expensive. A faster method, developed in the context of high energy physics, is to approximate the asymptotic form of the p-value by counting upcrossings, requiring fewer simulations [174]. In both of these cases new simulations are required each time a new model is considered, and the simulations may not be an accurate representation of the data. In this paper we seek an approach that can be directly applied to experimental data, without the need for simulations.

Our approach applies Bayesian logic to tackle the look-elsewhere effect. The Bayesian evidence is equal to the prior-weighted average of the likelihood over the parameter space, which can be considerably lower than the maximum likelihood if the prior is broad. This integration over the prior accounts for the look-elsewhere effect by penalizing large prior volumes. When considering large prior volumes, the likelihood is typically multimodal, with most of the peaks corresponding to noise fluctuations rather than physical sources. In order to estimate the location of a physical signal, and its associated statistical significance, one typically considers a point estimator, such as the maximum a posteriori (MAP) estimator which maximizes the posterior density. By applying the Laplace approximation, we introduce a Bayesian generalization of the MAP estimator, referred to as the maximum posterior mass (MPM) estimator, which corrects the MAP estimator by the prior-to-posterior volume ratio. Then, by drawing an analogy between Bayesian and frequentist methodology, we present a hybrid of the MAP and MPM estimators, called the maximum posterior significance (MPS) estimator, which determines the most significant peak in light of the look-elsewhere effect. The frequentist properties of the MPS estimator are shown to be independent of the look-elsewhere effect, providing a universal way to quantify the p-value, or statistical significance, without the need for expensive simulations.

The outline of this paper is as follows. In section 7.2 we review Bayesian posterior inference and hypothesis testing for a multimodal posterior, by discussing MAP estimation and then introducing MPM estimation. We then draw an analogy between Bayesian and frequentist philosophy in section 7.3 to motivate MPS estimation as the appropriate technique to tackle the look-elsewhere effect. The following three sections then apply this method to various examples: section 7.4 considers a resonance search, which can be thought of as a toy example of a collider or astrophysical particle search; section 7.5 considers a white noise time series, which can be thought of as a toy example of a gravitational wave search; and section 7.6 considers a search for non-Gaussian models of cosmological inflation using Planck data [290]. Note that section 7.4 is the main example, as it illustrates the key advantages of MPS, with the other examples complementary. Finally, we summarize and conclude in section 7.7.

7.2 Bayesian posterior inference and hypothesis testing

Two of the main tasks of Bayesian statistical analysis are posterior inference and hypothesis testing. Consider a model with parameters $\boldsymbol{z} = \{z_j\}_{j=1}^{M}$, and data $\boldsymbol{x} = \{x_i\}_{i=1}^{N_d}$ that depends on $\boldsymbol{z}$. The inference of $\boldsymbol{z}$ is given by its posterior

$$p(\boldsymbol{z}|\boldsymbol{x}) = \frac{p(\boldsymbol{x},\boldsymbol{z})}{p(\boldsymbol{x})} = \frac{p(\boldsymbol{x}|\boldsymbol{z})p(\boldsymbol{z})}{p(\boldsymbol{x})}, \tag{7.1}$$

where $p(\boldsymbol{x}|\boldsymbol{z})$ is the likelihood of the data, $p(\boldsymbol{z})$ is the prior of $\boldsymbol{z}$, and $p(\boldsymbol{x}) = \int d\boldsymbol{z}\; p(\boldsymbol{x}|\boldsymbol{z})p(\boldsymbol{z})$ is the Bayesian evidence, also known as the normalization, marginal likelihood, or partition function.[1] Typically, one can evaluate the joint probability $p(\boldsymbol{x},\boldsymbol{z})$, but not the evidence, which makes the posterior inference analytically intractable. This is usually handled using simple approximations or Monte Carlo Markov Chain methods [105].

A related problem is that of a hypothesis testing. In this case there are two different hypotheses, H and H_0, each with their own, potentially different, set of model parameters. In Bayesian methodology, hypothesis testing is performed using the Bayesian evidence ratio of the two hypotheses, which gives the Bayes factor

$$B \equiv \frac{p(\boldsymbol{x}|H)}{p(\boldsymbol{x}|H_0)}, \tag{7.2}$$

where the Bayesian evidence, or marginal likelihood, for hypothesis H is given by

$$p(\boldsymbol{x}|H) = \int d\boldsymbol{z}\; p(\boldsymbol{x}|\boldsymbol{z},H)p(\boldsymbol{z}|H). \tag{7.3}$$

If the prior on each hypothesis is equal, i.e. $p(H) = p(H_0) = 0.5$, then the Bayes factor is equal to the posterior odds ratio, $B = p(H|\boldsymbol{x})/p(H_0|\boldsymbol{x})$.

The Bayesian evidence and Bayes factor are also analytically intractable and harder to evaluate than posteriors, especially for high dimensional $\boldsymbol{z}$, although recent numerical methods such as Gaussianized Bridge Sampling [209] have made the problem easier. For the sake of exposition we will not consider such methods in this work, but instead use analytical approximations that give the Bayes factor an intuitive meaning. It is worth keeping in mind however that the full Bayes factor calculation can always be performed numerically, without any approximations.

Maximum a Posteriori (MAP) estimation

Given the analytical intractability of posterior inference and hypothesis testing, one often chooses an estimator to extract useful information from the posterior. A common estimator is

[1] Note that throughout this paper we use the letter p to refer to likelihood, prior, posterior, and evidence, even though they each refer to different functions. This notation has the advantage of making the probabilistic nature of each function clear, while the identity of the function should be clear from its argument.

the maximum a posteriori (MAP) point estimator, which corresponds to the global maximum of the posterior. If the prior is flat, as it will always be in this paper, this equals the maximum likelihood estimator (MLE), which maximizes the likelihood. Mathematically, MAP is defined via

$$\text{MAP} : \arg\max_{\boldsymbol{z}} p(\boldsymbol{z}|\boldsymbol{x}), \tag{7.4}$$

i.e. the MAP is located at the argument, z, that maximizes the posterior. For the purpose of comparing data to a null hypothesis, a useful quantity to define is

$$q_L(\boldsymbol{z}) \equiv 2 \ln \frac{p(\boldsymbol{x}|\boldsymbol{z})}{p(\boldsymbol{x}|\boldsymbol{z}_n)}, \tag{7.5}$$

where $\boldsymbol{z}_n$ represents the values of the parameters under the null hypothesis, and a subscript of L is used because the argument of the logarithm is the Likelihood ratio. To assess the significance of a result one considers the maximum value of q_L, which in the case of a flat prior is equal to q_L evaluated at the MAP: $\hat{q}_L = q_L(\boldsymbol{z}_{\text{MAP}})$. For a Gaussian likelihood, this is equal to the chi-squared (χ^2), and in the absence of the look-elsewhere effect $\sqrt{\hat{q}_L}$ typically gives the statistical significance. However, we will see that this test statistic greatly suffers from the look-elsewhere effect. This is because q_L effectively assumes the prior on $\boldsymbol{z}$ to be a delta function located at the best fit value. Such a prior is unreasonable because it has been determined a posteriori (after seeing the data), and ignores the fact that when looking elsewhere in parameter space the prior will be broad and in turn penalize the significance. We will now explore this in more depth, and later discuss what a reasonable prior means.

Maximum Posterior Mass (MPM) estimation

MAP is often a good point estimator in low dimensions if there is a single mode in the posterior. However, if the posterior has several modes, a more reasonable point estimator associates with the highest posterior mass. We refer to this as the maximum posterior mass (MPM) estimator.

For the purposes of this work, we will consider the example of a multimodal posterior consisting of a sum of multivariate Gaussian distributions; this has been shown to be a good approximation in many practical cases [327]. We thus consider a posterior of the following Gaussian mixture form,

$$p(\boldsymbol{z}|\boldsymbol{x}) = \sum_l w^l N(\boldsymbol{z}; \boldsymbol{\mu}^l, \boldsymbol{\Sigma}^l), \tag{7.6}$$

where $N(\boldsymbol{z}; \boldsymbol{\mu}, \boldsymbol{\Sigma})$ is a multivariate normal distribution with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$.Note that working with a posterior of this form is equivalent to applying the Laplace approximation to a general multimodal posterior in the upcoming derivations. In this model, the mass of mode l is proportional to the weight w^l, which is normalized such that $\sum_l w^l = 1$.

Assuming that only one component contributes at each peak, the weight of mode l is given by evaluating the posterior at the location of the mode, $\boldsymbol{z} = \boldsymbol{\mu}^l$,

$$\ln w^l = \ln p(\boldsymbol{\mu}^l|\boldsymbol{x}) - \ln N(\boldsymbol{\mu}^l; \boldsymbol{\mu}^l, \boldsymbol{\Sigma}^l) = \ln p(\boldsymbol{\mu}^l|\boldsymbol{x}) + \frac{1}{2}\left[\ln\det \boldsymbol{\Sigma}^l + M \ln(2\pi)\right]. \tag{7.7}$$

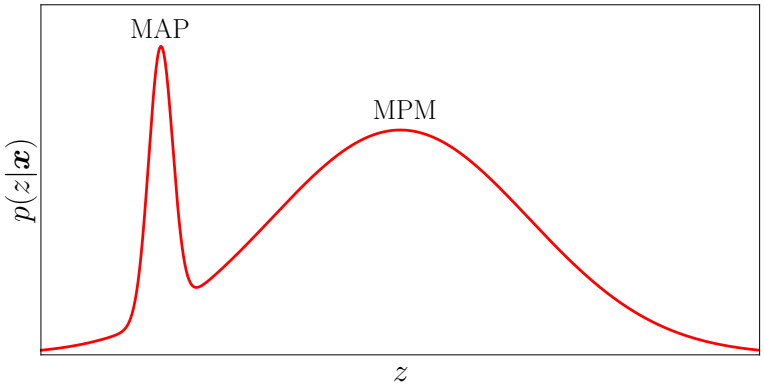

Figure 7.1: Plot of a bimodal Gaussian posterior for a 1d example in which 90% of the posterior mass is assigned to the right peak and 10% to the left. MPM yields the mode that maximizes the posterior mass and is close to the true mean, whereas MAP maximizes the posterior density and can be distant from the mean and represent only a small fraction of posterior mass.

Given the aforementioned analytical intractability of the posterior, we multiply each weight by the normalization $p(\boldsymbol{x})$ to give a quantity that can be readily computed. We call this quantity the mass m^l, and it is defined by

$$\ln m^l \equiv \ln w^l + \ln p(\boldsymbol{x}) = \ln p(\boldsymbol{x}|\boldsymbol{\mu}^l) + \ln p(\boldsymbol{\mu}^l) + \frac{1}{2}\left[\ln\det \boldsymbol{\Sigma}^l + M\ln(2\pi)\right]. \tag{7.8}$$

Thus the mass of each mode is equal to the likelihood multiplied by the product of the prior density and the posterior volume at the peak, where the posterior volume is defined as

$$V_{\text{posterior}} \equiv (2\pi)^{M/2}\sqrt{\det \boldsymbol{\Sigma}}. \tag{7.9}$$

The MPM estimator corresponds to the mode with the highest mass, thus

$$\text{MPM} : \arg\max_{\boldsymbol{z}} \left[p(\boldsymbol{x}|\boldsymbol{z})p(\boldsymbol{z})\sqrt{\det \boldsymbol{\Sigma}(\boldsymbol{z})}\right]. \tag{7.10}$$

To determine the MPM mode one would compute the $\ln m^l$ by first finding the positions of all local posterior maxima $\boldsymbol{\mu}^l$, and then computing $\boldsymbol{\Sigma}^l$ using the inverse of the Hessian at each peak. Qualitatively, MPM corresponds to maximizing the posterior density multiplied by the posterior volume $\sim \sqrt{\det \boldsymbol{\Sigma}}$, whereas MAP only maximizes the former. It is apparent that if there are multiple modes in the posterior, the one that has the largest posterior mass does not necessarily have the largest posterior density, as shown in figure 7.1. In some situations the MPM mode will dominate the posterior mass such that the MPM mode alone gives a useful way to summarize the posterior.

Hypothesis testing with MPM

Consider a model with parameters $z_1, z_2, ..., z_M$, with z_1 corresponding to the amplitude of a feature, and $z_2, ..., z_M$ corresponding to the properties of the feature. For example, z_1 might correspond to the amplitude of a signal detected in a time series at time z_2. We will use the notation $\boldsymbol{z} \equiv (z_1, ..., z_M)^T$ as the vector of all parameters, and $\boldsymbol{z}_{>1} \equiv (z_2, ..., z_M)^T$ as the subvector of non-amplitude parameters, i.e. excluding z_1. A typical analysis would scan over $\boldsymbol{z}_{>1}$, finding the best fit value for the amplitude z_1 at each point, giving rise to a multimodal posterior.

In this work we wish to determine whether or not a dataset contains a true anomaly. In the language of hypothesis testing, we wish to compare the hypothesis that there is an anomaly H, corresponding to $z_1 > 0$, to the null hypothesis that there is no anomaly H_0, corresponding to $z_1 = 0$. We assume the common case that the parameters of H_0 are a subset of the parameters of H, with H reducing to H_0 when $z_1 = 0$. There may also be parameters other than $\boldsymbol{z}$ that are common to both models, but these are of secondary importance when considering the look-elsewhere effect and we drop these from the notation.

Using equation 7.8 with $\sum w^l = 1$ implies that the Bayesian evidence for hypothesis H is given by

$$p(\boldsymbol{x}|H) = \int d\boldsymbol{z}\ p(\boldsymbol{x}|\boldsymbol{z}, H)p(\boldsymbol{z}|H) = \sum_l m^l, \tag{7.11}$$

where the m^l correspond to the masses under hypothesis H. Hence, each mode contributes its mass to the evidence. It follows that the mass of mode l corresponds to the Laplace approximation of the evidence integral in equation 7.11, integrated over the region of the mode. Because the null hypothesis does not depend on $\boldsymbol{z}_{>1}$, the evidence for the null hypothesis is given by the likelihood evaluated at $z_1 = 0$, that is $p(x|H_0) = p(x|z_1 = 0) \equiv p_0(x)$. Together with equation 7.11 this gives the Bayes factor

$$B \equiv \frac{p(\boldsymbol{x}|H)}{p(\boldsymbol{x}|H_0)} = \frac{1}{p_0(\boldsymbol{x})} \sum_l m^l \equiv \sum_l b^l, \tag{7.12}$$

where b^l is defined as the contribution of mode l to the Bayes factor. Using equation 7.8 gives

$$b^l = \frac{p(\boldsymbol{x}|\boldsymbol{\mu}^l)}{p_0(\boldsymbol{x})} p(\boldsymbol{\mu}^l)(2\pi)^{M/2}\sqrt{\det \boldsymbol{\Sigma}^l} = \frac{p(\boldsymbol{x}|\boldsymbol{\mu}^l)}{p_0(\boldsymbol{x})} \frac{V_{\text{posterior}}(\boldsymbol{\mu}^l)}{V_{\text{prior}}(\boldsymbol{\mu}^l)}, \tag{7.13}$$

where we have introduced the effective volume of the prior at $\boldsymbol{\mu}^l$ as,

$$V^{-1}_{\text{prior}}(\boldsymbol{\mu}^l) \equiv p(\boldsymbol{\mu}^l), \tag{7.14}$$

appropriate for the case of a narrow posterior relative to the prior. In the remainder of this paper we will drop the $\boldsymbol{\mu}^l$ dependence of the prior volume, as appropriate for a flat prior on $\boldsymbol{z}$.

Intuitively, one can think of each b^l as the Bayes factor one would get if mode l were the only mode in the posterior. If the maximum b^l is sufficiently large, it alone can provide a useful approximation to the Bayes factor, meaning the MPM mode dominates the Bayes factor. The first ratio on the right hand side of equation 7.13 corresponds to the likelihood ratio of the signal hypothesis to the null hypothesis, evaluated at the location of the peak, $\boldsymbol{z} = \boldsymbol{\mu}^l$. This is greater than or equal to 1 since adding parameters to the null hypothesis can only improve the fit. The second ratio gives the ratio of the posterior volume to the prior volume at the peak, which is always less than 1. This acts as a penalty to the likelihood ratio, often referred to as the Occam's razor penalty [251], or model complexity penalty, and compensates for the look-elsewhere effect in the case of a multimodal posterior. The higher the prior-to-posterior volume ratio, the higher the chance that peaks with a high likelihood will occur because of statistical fluctuations, thus the larger the penalty required to compensate.

Just as q_L is the estimator associated with MAP, we can define $q_b \equiv 2 \ln b$ as the estimator associated with MPM, such that

$$q_b = q_L - 2 \ln \frac{V_{\text{prior}}}{V_{\text{posterior}}}. \tag{7.15}$$

The MPM mode corresponds to the mode with maximum q_b. This illustrates how the MAP estimator ignores the look-elsewhere penalty by effectively considering the posterior and prior to be overlapping delta functions, which presumes a priori knowledge of the parameters and gives a prior-to-posterior volume ratio of unity.

An interesting question to consider is whether one can relate q_b to the look-elsewhere corrected statistical significance in a frequentist sense. In the absence of the look-elsewhere effect, the significance is given by $\sqrt{q_L}$, but simply taking $\sqrt{q_b}$ as the look-elsewhere corrected significance would not be correct. In the next section we turn to a frequentist description of the look-elsewhere effect to motivate a new estimator which applies a small modification to q_b and has a simple interpretation in terms of the significance, or p-value.

Before ending this section we discuss the choice of priors appropriate for a look-elsewhere analysis. If one has no prior knowledge regarding the location of an anomaly, then a uniform prior for the $\boldsymbol{z}_{>1}$ parameters is appropriate. If the prior is wide and posterior narrow this induces a large look-elsewhere effect. This choice of prior is not controversial. On the other hand, the choice of prior for the amplitude parameter z_1 is less clear. If one has no prior knowledge of the signal amplitude, then one should be open to a signal of any size, however one does not want the amplitude prior to induce a look-elsewhere penalty. In Bayesian hypothesis testing the amplitude parameter is treated analogously to the other parameters, thus if one uses too broad an amplitude prior it will induce an unwanted look-elsewhere penalty, whereas if one chooses too narrow an amplitude prior one risks discounting a large signal. Based on this we rewrite b in the following form, explicitly separating the marginalization over $\boldsymbol{z}_{>1}$ and z_1,

$$b = e^{q_L/2} \frac{V_{>1,\text{posterior}}}{V_{>1,\text{prior}}} \frac{V_{1,\text{posterior}}}{V_{1,\text{prior}}}. \tag{7.16}$$

The posterior volume terms are given by the covariance matrix, as in equation 7.9, and $V_{>1,\text{prior}}$ is given by the choice of prior on $\boldsymbol{z}_{>1}$. It thus remains to justify a choice of $V_{1,\text{prior}}$, which we will do by turning to a frequentist description of the look-elsewhere effect in the next section.

7.3 From Bayesian to frequentist hypothesis testing

Standard statistics literature states that Bayesian and frequentist hypothesis testing follow different methodologies and may give very different results. One famous illustration of this is the Jeffreys-Lindley "paradox" [243], however, there is much debate as to whether this is indeed a paradox and how relevant it is for scientific discourse (see [123] for a review in the context of high energy physics). While Bayesian statistics uses the Bayes factor for hypothesis testing, frequentist statistics uses the maximum likelihood ratio, or $\hat{q}_L$. One of the most important aspects of frequentist methodology is the computation of the false positive rate using the p-value, which quantifies how often a test statistic, for example $\hat{q}_L$, will take a specific value or larger under the assumptions of the null hypothesis. This has an intuitive interpretation as it directly relates to the false positive rate of the test statistic. On the other hand, Bayesian methodology rejects the p-value. The basis for this rejection is the likelihood principle, which states that any inference about the parameters $\boldsymbol{z}$ from the data $\boldsymbol{x}$ can only be made via the likelihood $p(\boldsymbol{x}|\boldsymbol{z})$. When the likelihood principle is applied to testing a hypothesis with parameters $\boldsymbol{z}$ one must use the marginal likelihood by integrating out these parameters – as in the Bayesian evidence of equation 7.3 – thus Bayesian methodology explicitly satisfies the likelihood principle. It is commonly argued that p-values violate the likelihood principle, because they rely on the frequentist properties of a distribution that go beyond the likelihood principle. However, the Bayes factor provides a less reliable tool for model comparison, as it is often interpreted in terms of arbitrary, model-independent scales [272], unlike the p-value which directly relates to the false positive rate.

We seek to elucidate how the answers of the two schools of statistics relate to one and other when it comes to the hypothesis testing. Both schools of statistics should give a similar, or at least related answer, when the question is phrased similarly. For uncertainty quantification it is often argued that the two schools do not answer the same question, since the Bayesian school treats data as fixed and varies the models, while the frequentist school varies the data at a fixed model. However, when it comes to hypothesis testing the distinction is less prominent: for example, when comparing two discrete hypotheses without any marginalizations, the answer in both cases gives the likelihood ratio as the optimal statistic (assuming equal prior for the two hypotheses). For continuous hypotheses it is often argued this is not possible. Here we will show that the two answers, the p-value and the Bayes factor, can be related with a specific choice of prior. It is important to emphasize that we are not claiming to equate the Bayesian and frequentist methodologies, but rather motivate a connection.

In this work we define the p-value as the probability under the null hypothesis, H_0, of

a random variable, Q, to be observed to have a value equal to or more extreme than the value observed, q. We thus use the notation $P(Q \geq q)$ for the p-value. To compute the p-value of a test statistic, one must consider how the test statistic is distributed under the null hypothesis. For the example of $\hat{q}_L$ this distribution is not universal: scanning over continuous variables, as in the look-elsewhere effect, will modify this distribution in a model dependent manner. Moreover, increasing the model complexity in other ways, for example by including extra degrees of freedom, will further modify the distribution. To account for extra degrees of freedom, Wilks' theorem [389] provides the asymptotic distribution of $\hat{q}_L$ for a hypothesis test where H has ν more degrees of freedom than H_0. However, Wilks' theorem relies on technical conditions, such as the observed value not being at the edge of the interval, and does not consider the look-elsewhere effect. Generalization of Wilks' theorem for the look-elsewhere effect have been considered in [133, 134] and have been translated into a practical procedure in [174]. As a result, a frequentist approach consists of a series of considerations to determine the change in the distribution of $\hat{q}_L$ due to different sources of model complexity. This is unlike the Bayesian methodology where all forms of model complexity are accounted for in the same way, as they are encoded into the Bayes factor. By connecting the two methodologies, we will present a test statistic whose distribution is universal, regardless of the model complexity and look-elsewhere effect.

Maximum Posterior Significance (MPS) estimation

We start by considering the typical case of one degree of freedom, corresponding to a single signal with amplitude z_1 and features described by $\boldsymbol{z}_{>1}$. We denote q_L maximized over the amplitude parameter only as $\check{q}_L(\boldsymbol{z}_{>1}) \equiv \max_{z_1} q_L(\boldsymbol{z})$, not to be confused with $\hat{q}_L \equiv \max_{\boldsymbol{z}} q_L(\boldsymbol{z})$ which is q_L maximized over all parameters. For a t-tailed test (where t is equal 1 or 2), Wilks' theorem gives the asymptotic p-value of $\check{q}_L$, at any position $\boldsymbol{z}_{>1}$, as

$$P(\check{Q}_L \geq \check{q}_L) = \frac{t}{2}\tilde{F}_1(\check{q}_L) \xrightarrow{\check{q}_L \to \infty} \frac{t}{\sqrt{2\pi\check{q}_L}} e^{-\check{q}_L/2}, \tag{7.17}$$

where $\tilde{F}_\nu$ is the complementary cumulative distribution function (CCDF) of a chi-squared random variable with ν degrees of freedom. This maximization over z_1 at a fixed choice of $\boldsymbol{z}_{>1}$ corresponds to the p-value in the absence of the look-elsewhere effect, referred to as the *local* p-value. Further maximizing over $\boldsymbol{z}_{>1}$ introduces the look-elsewhere effect, which can be parameterized by multiplying by the trials factor N such that

$$P(\hat{Q}_L \geq \hat{q}_L) = N\frac{t}{\sqrt{2\pi\hat{q}_L}} e^{-\hat{q}_L/2}. \tag{7.18}$$

This is referred to as the *global* p-value. It is this form that encapsulates the Bonferroni correction [77] which effectively multiplies the local p-value by N to account for the look-elsewhere effect. For discrete problems the trials factor equals the number of trials performed.

However, in the continuous case it is ill-defined, but it quantifies how the probability of finding a spurious peak increases as one looks elsewhere in the space spanned by $\boldsymbol{z}_{>1}$. Accounting for the look-elsewhere effect thus requires an expression for the trials factor.

It follows from equation 7.18 that one can define a test statistic,

$$q_S = q_L - 2\ln N + \ln 2\pi q_L - 2\ln t \tag{7.19}$$

such that the global p-value tends to

$$P(\hat{Q}_S \geq \hat{q}_S) \to e^{-\hat{q}_S/2}, \tag{7.20}$$

as either $N \to \infty$ or $\hat{q}_S \to \infty$, so this also applies for $N = 1$. See the appendix for this chapter (Section 7.8) for a derivation. Unlike $\hat{q}_L$, $\hat{q}_S$ has a distribution that is independent of N – the look-elsewhere effect has been absorbed into the test statistic. Intuitively one can think of the $2\ln N$ term as a penalty to q_L to correct for the look-elsewhere effect, while the $\ln 2\pi q_L$ term removes q_L dependent bias, ensuring the p-value depends on $\hat{q}_S$ alone in the asymptotic limit. Thus to account for the look-elsewhere effect one need only compute $\hat{q}_S$ and use this equation to compute the p-value. Because the p-value is a monotonically decreasing function of $\hat{q}_S$, one can think of selecting the peak with maximum q_S as selecting the peak with minimum p-value or maximum statistical significance. We refer to the mode with maximum q_S as the MPS mode, deferring an explanation for this nomenclature until the end of the subsection. The similarity of q_S to q_b from equation 7.15 suggests a connection between the frequentist and Bayesian pictures, and we now invoke this connection to find an expression for N and in turn generalize the Bonferroni correction to continuous parameters.

Heuristically, the Bayes factor describes the probability of the alternative hypothesis relative to the null, determined by the likelihood (as measured by $\hat{q}_L$), while the p-value averages its inverse over all values larger than $\hat{q}_L$ and will be smaller than the likelihood. We expect that for higher $\hat{q}_L$ the effect is larger because we are further into the tail of the distribution. There is no unique relation between the two, but one simple option is that the p-value scales as $B^{-1}/\hat{q}_L \approx \hat{b}^{-1}/\hat{q}_L$, where hats now indicate quantities associated with the MPS mode. Because we have the freedom to choose the prior on z_1, we can *d*efine the relation between the Bayes factor and p-value as

$$\frac{\hat{b}^{-1}}{\hat{q}_L} \equiv P(\hat{Q}_L \geq \hat{q}_L). \tag{7.21}$$

Comparing equation 7.16 with equation 7.18 then gives

$$\frac{V_{>1,\text{prior}}}{\hat{V}_{>1,\text{posterior}}} \frac{V_{1,\text{prior}}}{\hat{V}_{1,\text{posterior}}} \frac{e^{-\hat{q}_L/2}}{\hat{q}_L} = N \frac{t}{\sqrt{2\pi\hat{q}_L}} e^{-\hat{q}_L/2}. \tag{7.22}$$

By requiring that this relation holds in the absence of the look-elsewhere effect, the trials factor can be identified as

$$N = \frac{V_{>1,\text{prior}}}{\hat{V}_{>1,\text{posterior}}}, \tag{7.23}$$

and the amplitude prior volume is given by

$$V_{1,\text{prior}} = t\sqrt{\hat{q}_L}\frac{\hat{V}_{1,\text{posterior}}}{\sqrt{2\pi}} = t\sqrt{\hat{q}_L}\hat{\sigma}_1 \approx t\hat{\mu}_1. \tag{7.24}$$

In the final steps we used $\hat{V}_{1,\text{posterior}} = \sqrt{2\pi}\hat{\sigma}_1$, where $\hat{\sigma}_1$ is the error on the amplitude parameter, $\hat{\mu}_1$, and that the signal-to-noise ratio obeys $\sqrt{\hat{q}_L} \approx \hat{\mu}_1/\hat{\sigma}_1$. Since the look-elsewhere effect leads to large $\hat{q}_L$, this prior volume will be larger than the posterior volume. This choice of amplitude prior volume ensures that there is no trials factor associated with the amplitude, as intuition would dictate. Substituting equations 7.23 and 7.24 into equation 7.19 yields

$$q_S = q_b + 2\ln q_L. \tag{7.25}$$

Hence, we have effectively applied a modification to the MPM estimator to give a combination of the MPM and MAP estimators, so that the asymptotic p-value is neatly given by $e^{-\hat{q}_S/2}$. In the context of the look-elsewhere effect, the mode with maximum q_b will typically also be the mode with maximum q_L, and thus maximum q_S. However, this equivalence of MAP and MPM may not always be the case, as shown in figure 7.1.

A pure Bayesian might argue that equation 7.24 is not a valid prior, since it depends on the a posteriori amplitude parameter $\hat{\mu}_1$; however, this prior does have an intuitive justification. If a scientist is willing to consider a signal of any amplitude, the prior cannot be zero at $\hat{\mu}_1$, as it would not make sense to discard the signal. On the other hand, making the prior significantly broader than $\hat{\mu}_1$ implies the scientist has some additional information on the nature of the amplitude. When there is no justification for broadening the prior, the narrowest possible prior still consistent with the measured value can be more reasonable than arbitrarily fixing the size of the prior a priori. This choice of amplitude prior is simply designed to allow for a signal with any amplitude, without inducing an unwanted look-elsewhere penalty.

Note that the explicit dependence on $\hat{q}_L$ and the marginal likelihood, via $\hat{b}$, in equation 7.21 is what makes the p-value inconsistent with the likelihood principle. One could instead consider equating $\hat{b}^{-1}$ directly with the p-value, making it consistent with the likelihood principle. This would require an amplitude prior of $V_{1,\text{prior}} = t\hat{\sigma}_1^2/\hat{\mu}_1$, which we deem unreasonable as it is smaller than the posterior volume $\hat{V}_{1,\text{posterior}}$. We emphasize that the equality of $\hat{b}^{-1}/\hat{q}_L$ to the p-value is not required for our approach to the look-elsewhere effect, but provides insight on the Bayesian-frequentist connection. At its core, our method considers the test statistic $\hat{q}_S$, from equation 7.19, and replaces the trials factor N with the prior-to-posterior volume of the non-amplitude parameters $\boldsymbol{z}_{>1}$. Intuitively one can think of the number of trials as the number of posterior volumes that fit within the prior volume, and this intuition suggests $\hat{b}^{-1}$ scales linearly with the p-value.

Because the asymptotic p-value scales linearly with the prior volume, the non-asymptotic form of the p-value can be derived by dividing the prior volume into $K \gg 1$ regions and

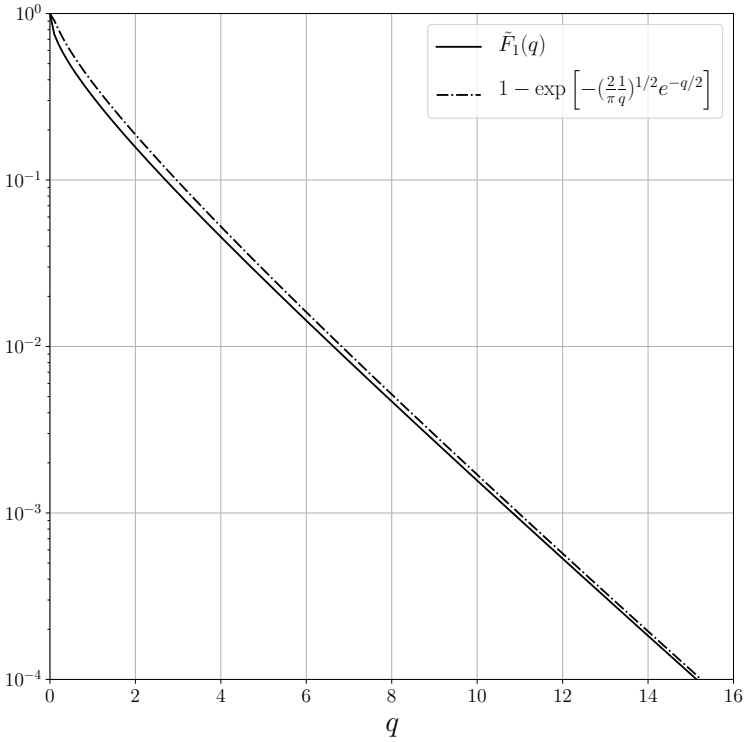

Figure 7.2: Equation 7.27 is a good approximation to $\tilde{F}_1(q)$ over the entire range of q. This suggests that MPS is still accurate in the absence of the look-elsewhere effect for a two-tailed test, even non-asymptotically.

evaluating the p-value for each. Assuming independence between these regions, the product of the p-values for each region can be used to obtain p-value of the full volume. Further assuming that the asymptotic regime still applies, this gives

$$P(\hat{Q}_S \geq \hat{q}_S) = \lim_{K\to\infty} \left[1 - \left(1 - \frac{e^{-\hat{q}_S/2}}{K}\right)^K\right] = 1 - \exp\left(-e^{-\hat{q}_S/2}\right). \tag{7.26}$$

Just as equation 7.20 is a generalization of the Bonferroni correction, equation 7.26 is a generalization of the Šidák correction [338] to continuous variables. This expression generalizes the p-value into the non-asymptotic regime.

For $N \gg 1$ every realization will have a positive peak, hence even the one-tailed p-value will approach 1 for sufficiently low $\hat{q}_L$, which equation 7.26 predicts to be for $\hat{q}_S < 0$. In the absence of the look-elsewhere effect ($N = 1$) a one-tailed test should approach a p-value of 0.5, while equation 7.26 approaches 1 as $\hat{q}_S \to -\infty$. Thus, the non-asymptotic agreement breaks down for $t = 1$ and $N = 1$. On the other hand, if $t = 2$ and $N = 1$, substituting

$q_S = q_L + \ln 2\pi q_L - 2\ln 2$ into equation 7.26 gives

$$P(\hat{Q}_L \geq \hat{q}_L)_{N=1,t=2} = 1 - \exp\left[-\left(\frac{2}{\pi}\frac{1}{\hat{q}_L}\right)^{1/2} e^{-\hat{q}_L/2}\right]. \tag{7.27}$$

The term in the square brackets can be identified as the asymptotic expansion of $\tilde{F}_1(\hat{q}_L)$. We show the non-asymptotic agreement of this equation with $\tilde{F}_1(\hat{q}_L)$, the true two-tailed p-value for $N = 1$, in figure 7.2. This illustrates the ability of the generalized Šidák correction to produce correct non-asymptotic results, even in the absence of the look-elsewhere effect. Hence, although we have applied asymptotic approximations throughout the above calculations, we have obtained a result that is valid even in the non-asymptotic limit. Inverting equation 7.27 gives the significance, or number of sigma, S, as

$$S^2 \approx \hat{q}_S - \ln 2\pi\hat{q}_S + 2\ln t, \tag{7.28}$$

with corrections of order $\mathcal{O}(\hat{q}_S^{-1})$. In the limit of $\hat{q}_S \to \infty$, the significance can be interpreted as $\sqrt{\hat{q}_S}$, in an analogous way to $\sqrt{\hat{q}_L}$ in the absence of the look-elsewhere effect. This motivates the name maximum posterior significance (MPS) as q_S depends on the posterior via the trails factor N, and is monotonically related to the significance S.

In summary, by considering a frequentist description of the look-elsewhere effect we introduced $\hat{q}_S$ as a natural test statistic to use, such that the asymptotic p-value is given by $e^{-\hat{q}_S/2}$. We derived a general expression for the p-value which also applies in the non-asymptotic regime, and when there's no look-elsewhere effect. Adopting the prior of equation 7.24, we showed that one can write the p-value in terms of Bayes factor as $\hat{b}^{-1}/\hat{q}_L$. This intrinsically accounts for the look-elsewhere effect by identifying the trials factor as the prior-to-posterior volume ratio of $\boldsymbol{z}_{>1}$ at the MPS mode. While one can compute the Bayes factor using a variety of methods, we will use the Laplace approximation to evaluate the posterior volume of each mode, as in section 7.2. To outline the step-by-step approach:

Maximum Posterior Significance (MPS) estimation:

1. Scan over the space of non-amplitude parameters, $\boldsymbol{z}_{>1}$, locating peaks in the posterior with any amplitude, z_1. Often only the highest few peaks are needed.
2. Compute q_L and the posterior volume, using equation 7.9, for each peak.
3. Compute q_b for each peak using equation 7.15 with the amplitude prior of equation 7.24.
4. Compute $q_S = q_b + 2\ln q_L$ for each peak.
5. Find the peak with maximum q_S.
6. Compute the (global) p-value using equation 7.26 and significance using 7.28.

Multiple degrees of freedom

For models with multiple degrees of freedom, the frequentist approach is to apply Wilks' theorem [389]. This is valid in the asymptotic limit, where, for a two-tailed test, the local p-value is given by

$$P_\nu(\check{Q}_L \geq \check{q}_L) = \tilde{F}_\nu(\check{q}_L) \xrightarrow{\check{q}_L \to \infty} \frac{1}{\Gamma(\nu/2)} \left(\frac{\check{q}_L}{2}\right)^{\nu/2-1} e^{-\check{q}_L/2}, \tag{7.29}$$

for a model with ν degrees of freedom. Note that the limit assumes $q \gg \nu$, but for $\nu = 2$ it is exact for any q. Wilks' theorem can address the model complexity problem of having multiple (ν) continuous amplitude parameters. A specific example from particle physics is a decay process with ν decay channels, each with amplitude A_i ($0 \leq i \leq \nu$). In such a case $\max_{\{A_i\}} q_L(\{A_i\}, ...) \sim \tilde{F}_\nu$. Wilks' theorem is not sufficiently general: it fails if the parameters are at the edge of their distribution, and it does not naturally handle the model complexity of the look-elsewhere effect, where one scans over a wide range of values for one or more parameters. Upon introduction of the look-elsewhere effect a frequentist would typically consider single trials distributed as $\sim \tilde{F}_\nu$, and then use a ν-dependent trials factor [174]. Thus in a frequentist approach extra degrees of freedom and the look-elsewhere effect are treated separately. On the other hand, a Bayesian approach accounts for both in the same way.

To apply the Bayesian methodology, we first reparameterize the model so that there is only a single amplitude parameter by introducing branching ratios α_i, such that each amplitude parameter is $A_i = \alpha_i z_1$, where z_1 is the total amplitude parameter and $\sum_{i=1}^{\nu} \alpha_i^2 = 1$. To remove the constraint we adopt rotation angles: for example, for $\nu = 2$ we can work with a phase angle ϕ, such that $\alpha_1 = \cos\phi$ and $\alpha_2 = \sin\phi$. Thus, instead of working with A_1 and A_2 and considering $\max_{A_1,A_2} q_L(A_1, A_2, ...) \sim \tilde{F}_2$, we consider $\max_{z_1} q_L(z_1, \phi, ...) \sim \tilde{F}_1$ with $\boldsymbol{z}_{>1} = (\phi, ...)$. We can then directly apply the MPS prescription for $\nu = 1$, as in the previous subsection, by additionally marginalizing over ϕ to account for the model complexity with an additional prior-to-posterior volume penalty.

To be agnostic, one would choose a prior volume for ϕ of $V_{\phi,\text{prior}} = \pi$ (in practice a more complex prior may be appropriate, but it will typically be $\mathcal{O}(1)$). Furthermore, the average error on ϕ is typically equal to the relative error on the amplitude, thus $\sigma_\phi \approx \sigma_1/\mu_1 \approx q_L^{-1/2}$. This gives a model complexity correction of

$$\frac{V_{\phi,\text{prior}}}{\hat{V}_{\phi,\text{posterior}}} = \frac{\pi}{\sqrt{2\pi}\hat{\sigma}_\phi} = \sqrt{\pi}\left(\frac{\hat{q}_L}{2}\right)^{1/2} = \frac{\tilde{F}_2(\hat{q}_L)}{\tilde{F}_1(\hat{q}_L)}. \tag{7.30}$$

This shows that increasing the model complexity with an extra degree of freedom is accounted for in the Bayesian framework by marginalizing over ϕ. Thus, the Bayesian answer to an increase in model complexity, whether it be due to including extra degrees of freedom, or looking elsewhere, is identical: marginalization over the non-amplitude parameters $\boldsymbol{z}_{>1}$. The ν dependence of the local p-value in equation 7.29 can be interpreted as a Bayesian model

complexity penalty: a fixed p-value corresponds to a larger $\hat{q}_L$ as ν increases. Thus, MPS intrinsically generalizes Wilks' theorem by relating the trials factor to the prior-to-posterior volume.

7.4 Example I: Resonance searches

To test the theory of section 7.3 we first consider a resonance search example. These appear in many different areas of physics, including astroparticle and high energy physics. We consider a search for a new particle whose mass and cross-section are unknown. The data $\boldsymbol{x}$ could correspond to measurements of the invariant mass in the case of collider searches, or the energy flux in astroparticle searches. The probability density for a single measurement, x^i, is given by

$$p(x^i|f, x_*, \sigma_*) = f p_s(x^i|x_*, \sigma_*) + (1-f)p_b(x^i), \tag{7.31}$$

where p_s and p_b are the normalized signal and background distributions respectively, and f is the fraction of events belonging to the signal. We assume that the form of the signal and background are known; we take the signal to be a normal distribution $p_s(x^i|x_*, \sigma_*) = N(x^i|x_*, \sigma_*)$, and the background to be a power law. Thus the resonance has position x_* and width σ_*. Given data $\boldsymbol{x} = \{x^i\}_{i=1}^{N_d}$, the likelihood is given by the product of the individual probability densities over the data. Using equation 7.31 this gives the likelihood as

$$p(\boldsymbol{x}|f, x_*, \sigma_*) = \prod_{i=1}^{N_d} \left[f p_s(x^i|x_*, \sigma_*) + (1-f)p_b(x^i) \right]. \tag{7.32}$$

Note that the Bayesian evidence under the null hypothesis is independent of the parameters, namely

$$p_0(\boldsymbol{x}) \equiv p(\boldsymbol{x}|f=0) = \prod_{i=1}^{N_d} p_b(x^i). \tag{7.33}$$

While the likelihood depends on the number of data N_d, quantities such as the p-value will have converged provided N_d is sufficiently large to resolve the resonance. Throughout this section we fix $N_d = 10 V_{x_*,\text{prior}}/\sigma_*$ to ensure sufficient convergence. We note that more complex models might consider drawing N_d from a Poisson distribution, however this is unnecessary for our proof of concept.

We first consider a uniform prior on x_*, with range $(0, 10^3)$, i.e. a prior volume of $V_{x_*,\text{prior}} = 10^3$. We do not fit for σ_* and fix it to $\sigma_* = 0.5$ a priori, corresponding to the narrow-width approximation. In this case the posterior is only multimodal in the x_* dimension, thus to find peaks we split the parameter space along the x_* dimension into narrow bins of size Δx_* and compute the maximum likelihood of equation 7.32 within each bin. Ensuring Δx_* is sufficiently small, we determine the location of all peaks in the posterior, $\boldsymbol{\mu}^l$, by comparing adjacent bins. The Hessian at each peak is then computed using finite differencing, and inverted to give $\boldsymbol{\Sigma}^l$. Note, in this example we have an analytical form for the likelihood,

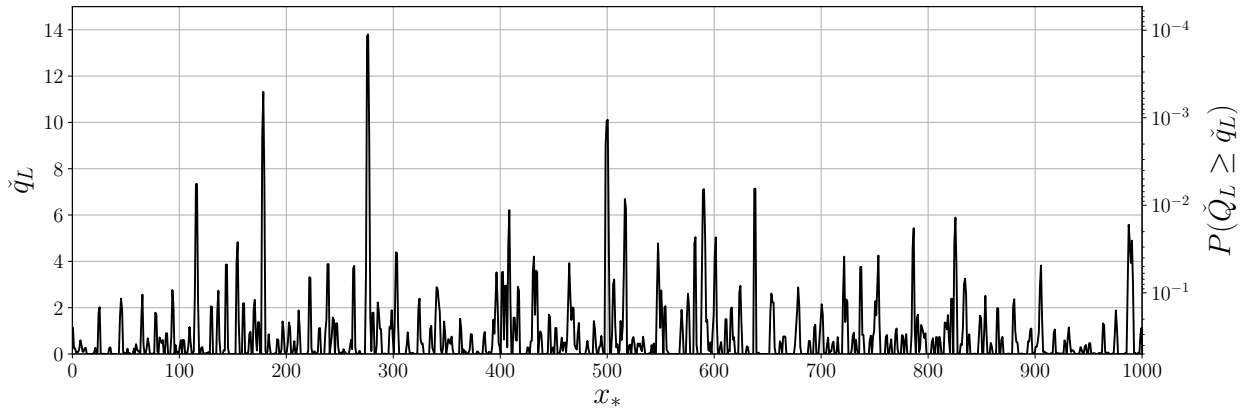

Figure 7.3: The local chi-squared (left axis) and local p-value (right axis) for an example data realization with true amplitude $f = 5 \times 10^{-3}$, position $x_* = 500$, and width $\sigma_* = 0.5$. While there is a peak with $\check{q}_L \approx 10$ at the correct position, the look-elsewhere effect leads to other, sometimes larger, peaks at random positions.

enabling verification of the numerical computation with analytical results. The value of q_b at each peak is then computed using equation 7.15, in turn giving $\hat{q}_S$.

Figure 7.3 shows the local chi-squared and local p-value as a function of x_* for an example data realization. We use true parameters $f = 5 \times 10^{-3}$ and $x_* = 500$. Recall from equation 7.17 that the local chi-squared and p-value correspond to the values obtained by maximizing over f at fixed x_*, i.e. they correspond to the values obtained without having corrected for the look-elsewhere effect. The local chi-squared $\check{q}_L$ can also be thought of as the projection of q_L onto the x_* axis. It can be seen that although there is a peak with $q_L \approx 10$ at the correct position, there are also multiple spurious peaks throughout the parameter space, with $\hat{q}_L \approx 14$ in this example. This illustrates the look-elsewhere effect: peaks with a local p-value of $\sim 10^{-4}$ are produced by noise, meaning a signal with such a local p-value should not be considered as significant as its local p-value naively suggests.

We now consider 10^5 different data realizations without a signal ($f = 0$) to study the distributions of $\hat{q}_L$ and $\hat{q}_S$ under the null hypothesis. The plots in figure 7.4 show the global p-value in terms of $\hat{q}_L$ and $\hat{q}_S$ for a variety of scenarios. One can think of the vertical axes as corresponding to the false positive rate (FPR) of a hypothesis test using threshold q.

We first compare three different prior volumes on x_*, $V_{x_*,\text{prior}} = 10^3, 10^2, 10^1$, to show the effectiveness of our method for large and small N. The top left plot of figure 7.4 shows that the p-value of $\hat{q}_L$ has a considerable prior volume dependence. This is the look-elsewhere effect: a larger prior volume leads to a larger trials factor and thus an increased probability of finding a higher maximum likelihood. On the other hand we see that $\hat{q}_S$ shows no prior dependence and is in good agreement with equation 7.26, even in the non-asymptotic regime.

We also investigate the variation of the p-value with the value of the width of the signal σ_*. This is shown in the top right plot of figure 7.4 where we consider $\sigma_* = 0.1, 0.5, 1.0$.

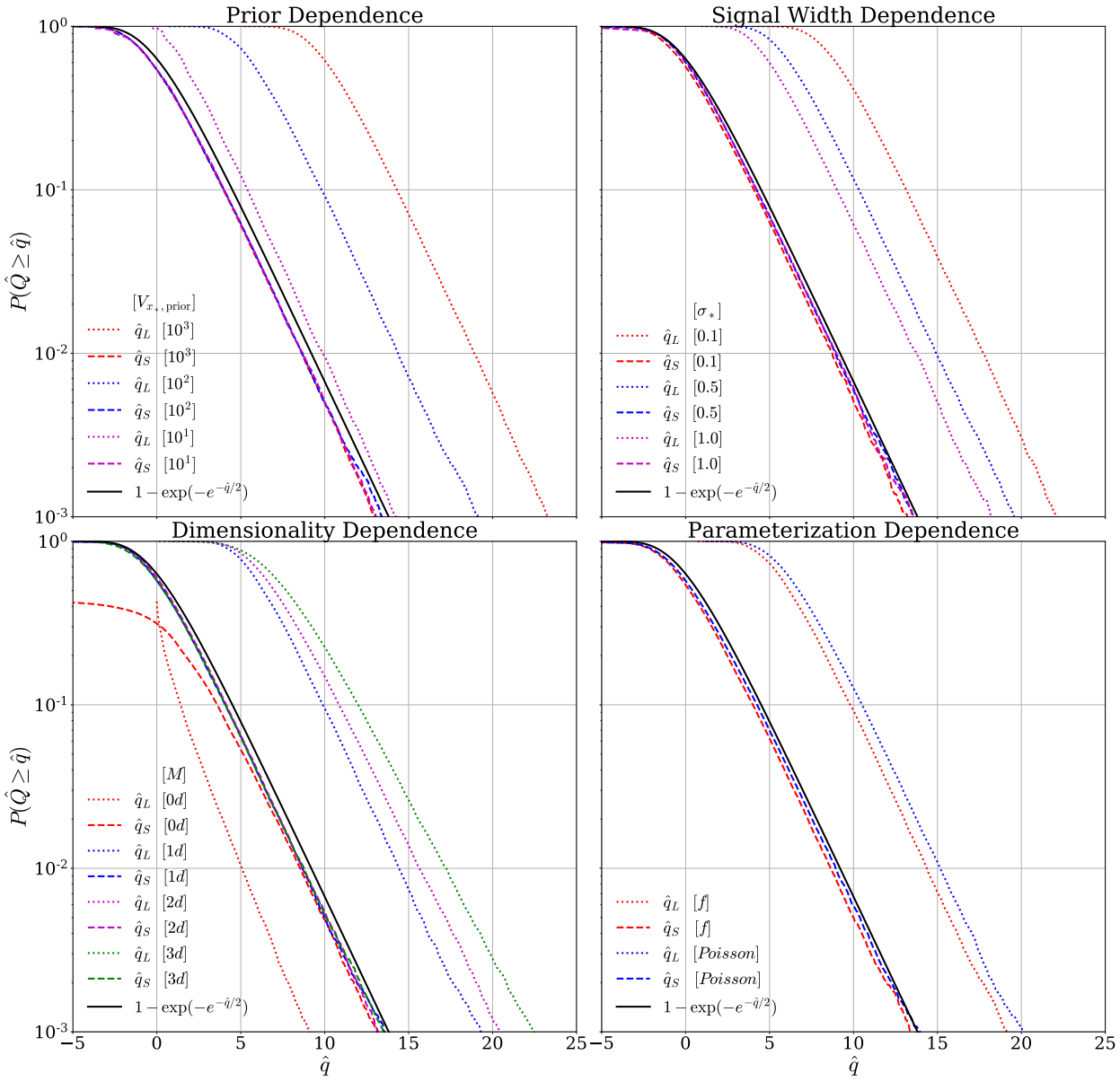

Figure 7.4: CCDFs of $\hat{q}_L$ (dotted) and $\hat{q}_S$ (dashed), computed using 10^5 simulations with no signal ($f = 0$). (Top Left) compares three prior volumes: 10^3 (red), 10^2 (blue), and 10^1 (magenta). (Top Right) compares different values of signal width σ_*: 0.1 (red), 0.5 (blue) and 1.0 (magenta). (Bottom Left) compares the dimensionality of $\boldsymbol{x}_*$: $0d$ (red), $1d$ (blue), $2d$ (magenta), and $3d$ (green). (Bottom Right) compares the un-binned f-parameterization (red) against a binned Poisson parameterization (blue). In all cases the p-value of $\hat{q}_L$ has large variation, whereas $\hat{q}_S$ does not. Furthermore, $\hat{q}_S$ closely follows the predictions of equation 7.26 (black).

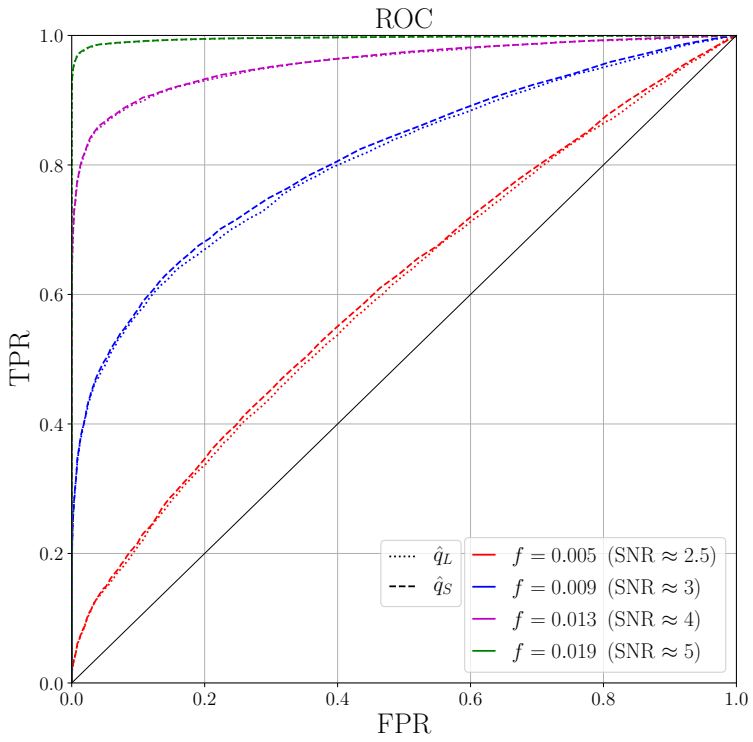

Figure 7.5: ROC curve: comparing the true positive rate (TPR), for a variety of f, with the false positive rate (FPR) for $\hat{q}_L$ (dotted) and $\hat{q}_S$ (dashed). The signal-to-noise ratio (SNR) corresponds to the average $\sqrt{\hat{q}_L}$ over all data realizations.

Smaller σ_* leads to a smaller posterior volume and thus a larger trials factor. Much like the discussion above for prior volume variation, $\hat{q}_L$ has a large σ_* dependence, unlike $\hat{q}_S$.

Next, we investigate the variation of the p-value with the dimensionality of the look-elsewhere effect. To do this we extended the model to consider a signal at vector position $\boldsymbol{x}_*$. Each data point now corresponds to a vector $\boldsymbol{x}^i$, and we extend the signal and background in a symmetric fashion across each dimension, keeping the total prior volume fixed. Within the context of collider searches, the components of $\boldsymbol{x}_*$ might correspond to a collection of invariant mass and jet properties. For astroparticle searches, the multiple dimensions might correspond to different directions in the sky. The bottom left plot of figure 7.4 shows the variation of the test statistics for dimensionality of 1, 2, and 3, for a constant prior volume of 100. It can be seen that, while the p-value of $\hat{q}_L$ is dependent on the dimensionality, the p-value of $\hat{q}_S$ is not. This justifies the naturally arising $(2\pi)^{M/2}$ prefactor in the posterior volume in equation 7.9. We also plot the $0d$ case, corresponding to only fitting for A with fixed x_*. Even though there is no look-elsewhere effect in this case, asymptotic agreement with equation 7.26 is still achieved. This shows our approach is still reliable in the $N \to 1$ limit, justifying its applicability for arbitrary N. As discussed in section 7.3, non-asymptotic

agreement is not expected for a one-tailed test in the absence of the look-elsewhere effect, as the p-value tends to 0.5 as $\hat{q}_L \to 0$; on the other hand, a two-tailed test would give non-asymptotic agreement as shown in figure 7.2.

The above discussion concerns an un-binned model, parameterized by the signal fraction f. Often in particle physics, one performs a binned analysis with the number of events in each bin modelled as a Poisson distribution [124]. We find similar results when using this Poisson parameterization, as pictured in the bottom right of figure 7.4. The Poisson line agrees with the black line slightly better than the f line does, likely because the Laplace approximation is more accurate in the Poisson case.

When it comes to hypothesis testing, the relation between the true positive rate (TPR) and the false positive rate (FPR) determines the predictive power of a test statistic. In order to compare the relative power of the test statistics we consider an ROC plot for a variety of true f values, shown in figure 7.5. We also quote the (local) signal-to-noise ratio (SNR), which we define as the average $\sqrt{\hat{q}_L}$ across 10^4 realizations for the given f. It can be seen that $\hat{q}_S$ and $\hat{q}_L$ have approximately equivalent ROC lines, suggesting MAP and MPS have equal predictive power. This is expected as the relation between the test statistics is approximately monotonic, as seen in equation 7.19. This agrees with the findings of [164] which considered a different Bayesian-inspired test statistic, and showed it to have an approximately equivalent ROC curve to the p-value of the likelihood. Also, it can be seen that the predictive power increases with true f – as expected a larger true signal is more likely to be correctly detected.

7.5 Example II: White noise

While we could continue the discussion in the context of resonance searches, we now consider a white noise time series example to illustrate the application of MPS to different models. This can be thought of as a toy model of a gravitational wave search. In this section we show how MPS handles additional model complexity as theorized in section 7.3. We consider a time series $y(x)$ comprising of measurements at N_d times, $\boldsymbol{x} = \{x^i\}_{i=1}^{N_d}$, with spacing $x^{i+1} - x^i = 1$. In the absence of a signal, each data point $y^i \equiv y(x^i)$ is assumed to be a standard normal random variable, i.e. we assume white noise. We consider a model with 2 degrees of freedom (dofs), with signal given by

$$p_s(x|A_1, A_2, x_*, \Delta, \sigma_*) = A_1 N(x|x_*, \sigma_*) + A_2 N(x|x_* + \Delta, \sigma_*) \tag{7.34}$$

where $A_{1,2} > 0$ are the amplitudes of each dof, x_* and $x_* + \Delta$ are the positions of the dofs, and σ_* is the common width.

As motivated in section 7.3, we reparameterize so that there's a single amplitude parameter, $z_1 = A$, and other parameters describing the properties of the single degree of freedom, $\boldsymbol{z}_{>1}$. We thus transform variables using $A_1 = A\cos\phi$ and $A_2 = A\sin\phi$, with $A > 0$ and $0 \leq \phi \leq \pi/2$ for a one-tailed test. By substituting the transformations into equation 7.34, the signal in the new parameterization is given by

$$p_s(x|A, \phi, x_*, \Delta, \sigma_*) = A\left[\cos\phi N(x|x_*, \sigma_*) + \sin\phi N(x|x_* + \Delta, \sigma_*)\right]. \tag{7.35}$$

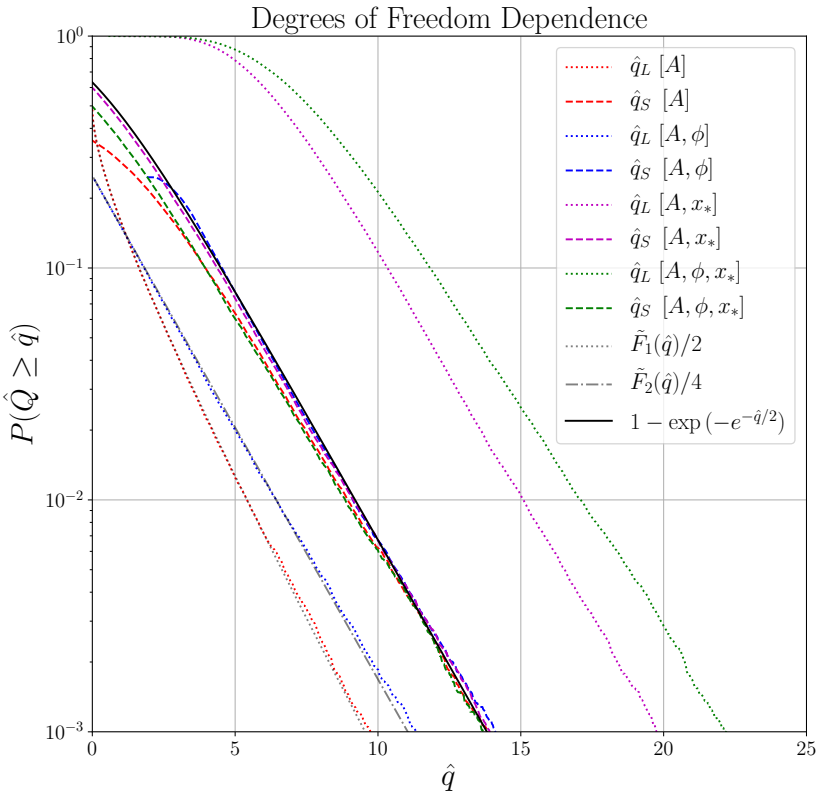

Figure 7.6: CCDFs of $\hat{q}_L$ and $\hat{q}_S$ averaged over 10^5 simulations with no signal ($A = 0$). The parameters in the square brackets are those being maximized, with other parameters being held fixed (as discussed in the text). The p-value of $\hat{q}_L$ varies depending on the model complexity, whereas $\hat{q}_S$ consistently follows the prediction of equation 7.26 (solid black).

The corresponding chi-squared difference between the data and the null hypothesis, equal to two times the log-likelihood-ratio, is given by

$$q_L(\boldsymbol{x}|A, \phi, x_*, \Delta, \sigma_*) = \sum_{i=1}^{N_d} \left[y^i\right]^2 - \left[y^i - p_s(x^i|A, \phi, x_*, \Delta, \sigma_*)\right]^2 . \tag{7.36}$$

We consider a uniform prior on x_* with range $(0, 100)$, i.e. a prior volume of $V_{x_*,\text{prior}} = 100$, and $N_d = 100$. We do not fit for σ_* or Δ and fix them to $\sigma_* = 0.5$ and $\Delta = 10$. The application of MPS is identical to the previous section, so we will not repeat the methodology here.

Considering 10^5 data realizations with no signal, figure 7.6 shows how $\hat{q}_L$ and $\hat{q}_S$ are distributed for different levels of model complexity. First we maximize over A, while holding all other parameters fixed. In this case $\hat{q}_L \sim \tilde{F}_1(\hat{q}_L)/2$ (red dotted line) as expected for a one-tailed test with one degree of freedom. Additionally maximizing over ϕ allows for 2 dofs, and gives $\hat{q}_L \sim \tilde{F}_2(\hat{q}_L)/4$ (blue dotted line). This is expected because there are 4

permutations of each dof having positive or negative amplitude, and $A_{1,2} > 0$ considers 1 of these 4. For both of these cases, $\hat{q}_S$ follows the same asymptotic distribution as predicted by equation 7.26. This verifies that the Bayesian picture of marginalizing over ϕ successfully reduces a model with 2 dofs to the same scale as 1 dof, in other words Wilks' Theorem has been replaced by marginalizing over ϕ. There is some discrepancy in the non-asymptotic regime for the maximization over A only (red dashed line), as discussed in section 7.3 for a one-tailed test.

We now introduce the look-elsewhere effect by allowing x_* to vary. First we maximize over A and x_* for fixed $\phi = 0$, as shown by the magenta lines. This is equivalent to a model with 1 dof because $\phi = 0$ corresponds to $A_2 = 0$. We see that the distribution of $\hat{q}_L$ (magenta dotted line) is shifted to the right compared to the red and blue dotted lines due to the look-elsewhere effect. However, the distribution of $\hat{q}_S$ (magenta dashed line) continues to follow the line predicted by equation 7.26. Finally, when maximizing over A, ϕ and x_*, i.e. a model with 2 dofs in the presence of the look-elsewhere effect, $\hat{q}_L$ (green dotted line) is further right-shifted, whereas $\hat{q}_S$ (green dashed line) again agrees with equation 7.26. The slight discrepancy in the A, ϕ, x_* maximization case is due to using too large a prior volume: there is a slight preference to having two well fitted peaks compared to one very well fitted peak, thus the distribution of ϕ is clustered towards $\phi = \pi/4$. Using a more appropriate prior for ϕ would improve agreement.

In summary, while the distribution of $\hat{q}_L$ is highly dependent on the model complexity, via the extra degrees of freedom and look-elsewhere effect, $\hat{q}_S$ has a universal distribution.

7.6 Example III: Non-Gaussian models of cosmological inflation

There is much interest in detecting non-Gaussian models of inflation via the cosmological power spectrum [253, 125, 219, 326, 291, 115]. A specific type of such a feature model adds the following oscillatory perturbation to the ΛCDM power spectrum,

$$P(k) = P_0(k)[1 + A\sin(2\omega k + \phi)], \tag{7.37}$$

where $P_0(k)$ is the featureless (ΛCDM) power spectrum and A, ω, and ϕ are the amplitude, frequency, and phase of the oscillatory perturbation. Such models are searched for using Planck 2013 data in [156] using the frequentist look-elsewhere analysis technique of [155]. In this section we seek to reproduce the conclusions of these papers using MPS.

Equation 7.37 can be written in the form $P(k) = P_0(k) + \Delta P(k)$ with

$$\begin{aligned} \Delta P(k; A, \omega, \phi) &= AP_0(k)[\cos\phi\sin(2\omega k) + \sin\phi\cos(2\omega k)] \\ &\equiv A\cos\phi P_s(k;\omega) + A\sin\phi P_c(k;\omega), \end{aligned} \tag{7.38}$$

where in the last line we explicitly separate terms with A and ϕ, as only ω couples to k. Assuming a linear relation, one can write $C_\ell = C_{\ell,0} + \Delta C_\ell$, with

$$\Delta C_\ell(A, \omega, \phi) = A\cos\phi C_{\ell,s}(\omega) + A\sin\phi C_{\ell,c}(\omega), \tag{7.39}$$

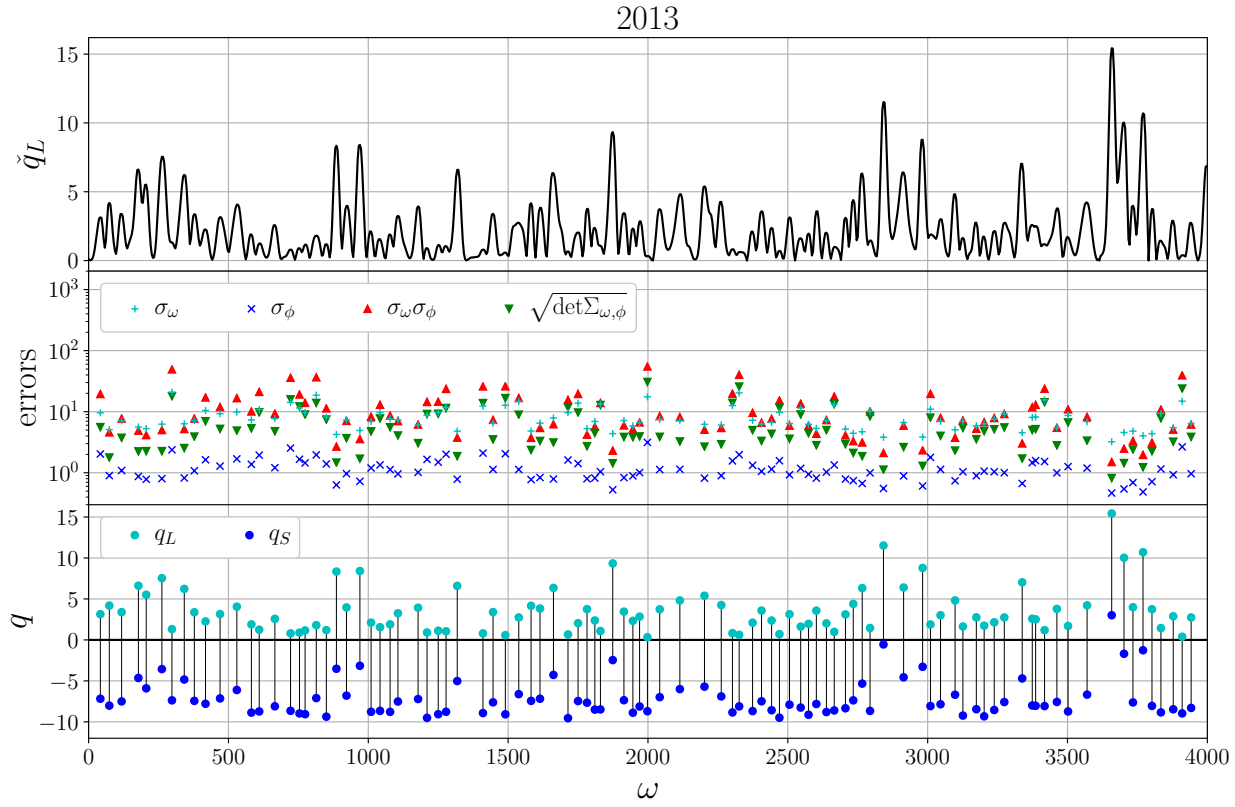

Figure 7.7: Planck results. Top: Plot of $\check{q}_L$, the projection of q_L onto the ω axis; this corresponds to q_L evaluated at the A and ϕ that maximize q_L at each ω. Middle: The errors obtained for the parameters, as well as a comparison with the determinant of the covariance matrix having removed the amplitude parameter, $\Sigma_{\omega,\phi}$. Bottom: A plot of q_L (blue) and q_S (cyan) for each peak, with the look-elsewhere correction depicted by the vertical black lines.

where $C_{\ell,s}$ and $C_{\ell,c}$ are the angular power spectra corresponding to P_s and P_c respectively. The Planck Likelihood [290] is given by

$$-2\log L(\hat{C}_\ell|A,\omega,\phi) = [\hat{C}_{\ell_1} - C_{\ell_1}(A,\omega,\phi)]\Delta_{\ell_1\ell_2}[\hat{C}_{\ell_2} - C_{\ell_2}(A,\omega,\phi)], \tag{7.40}$$

where $\hat{C}_\ell$ are the PCL estimates, and $\Delta_{\ell_1\ell_2} = \langle\Delta\hat{C}_{\ell_1}\Delta\hat{C}_{\ell_2}\rangle$ is the PCL covariance matrix. In order to compute the likelihood for the null hypothesis, CosmoMC [238] was used to find the best fit values for the cosmological and nuisance parameters. When computing the likelihood for the signal hypothesis, the cosmological parameters were held fixed at these values; while they should really be re-fitted for the signal hypothesis, this is found to have little effect in [156]. The C_ℓ are evaluated using CAMB [239] with a sufficiently high accuracy setting to ensure resolution of the rapid oscillations. To speed up the evaluation of the likelihood over parameter space, $C_{\ell,s}(\omega)$ and $C_{\ell,c}(\omega)$ were computed over a discrete range of ω between 0 and 4000 with step-size $\Delta\omega = 5$, with intermediate values computed via spline interpolation.

A flat prior was chosen for ω and ϕ. The rest of the analysis is analogous to the previous examples: we find all the local maxima of the posterior, compute the Hessian using finite differencing, compute the covariance matrix, and use this to find $\hat{q}_S$. Unlike the previous examples, we note that ω and ϕ are correlated, as illustrated in the middle plot of figure 7.7, so it is important to use the determinant of the full covariance matrix and not just its diagonal components. It is also interesting to note that higher peaks have smaller errors.

The results obtained using the CAMspec component of the 2013 Planck likelihood[2] are pictured in figure 7.7. The maximum occurs at $\omega \approx 3660$ with $\hat{q}_L = 15.4$, giving a naive significance of $\sqrt{\hat{q}_L} \approx 4$ sigma. However, we find that $\hat{q}_S = 3.0$, giving a global p-value of $1 - \exp(1 - e^{-3/2}) = 0.20$ using equation 7.26, and significance of $S = 1.3$ sigma. Thus the signal is in fact far less significant in light of the look-elsewhere effect. The prescription of [156] gives a p-value of 0.13, which is in reasonable agreement. Note that our likelihood profile does not match [156] exactly due to our approximate approach, hence the p-value quoted here is the value one would obtain by applying the prescription of [156] to our likelihood profile. We applied the same analysis to the 2015 `plik_lite` likelihood [8] and found a p-value of approximately 1, suggesting no evidence for such models of non-Gaussianity.

7.7 Conclusions

This work has employed Bayesian and frequentist thinking to provide a fast method to account for the look-elsewhere effect when scanning over a large parameter space. We started by considering the Bayesian approach, and explained how maximizing the posterior mass, as in MPM, is a more appropriate choice than maximizing the posterior density, as in MAP. Bayesian methodology naturally considers model complexity and the look-elsewhere effect by marginalization, which penalizes the likelihood by the prior-to-posterior volume ratio. (Under the Laplace approximation, the posterior volume is simply related to the determinant of the covariance matrix.) We then considered the frequentist approach by writing the global p-value as the local p-value multiplied by the trials factor. By drawing an analogy between the two approaches we identified the trials factor as the prior-to-posterior volume ratio of the parameters being scanned over, in turn generalizing the Bonferroni correction to continuous problems. We introduced q_S and in turn MPS, a hybrid of MPM and MAP, which considers the mode with maximum q_S. Finally, we generalized the Šidák correction to continuous problems, providing a universal way to assign the global p-value in both the asymptotic and non-asymptotic regimes.

We illustrated the effectiveness of MPS by considering several examples from (astro)particle physics and cosmology, showing it to have equal predictive power to MAP while naturally accounting for the look-elsewhere effect. MPS effectively shifts the hypothesis testing threshold of the maximum likelihood ratio to a generic scale: while the maximum likelihood ratio, or equivalently the best fit chi-squared $\chi^2 = \hat{q}_L$, depends on the model

[2] One should sum the different components of the likelihood, but this is unnecessary for our proof of concept.

complexity and extent of the look-elsewhere effect, $\hat{q}_S$ does not. In other words, instead of considering fixed $\hat{q}_L$ thresholds, one should consider fixed $\hat{q}_S$ thresholds.

Unlike current methods that rely on performing numerous simulations, MPS accounts for the look-elsewhere effect by using information from the data alone, as one need only compute the likelihood and the posterior volume to evaluate q_S. This provides a more efficient way to quantify statistical significance as it does not require expensive simulations. In a typical situation one would focus on the most promising anomalies only, with $\hat{q}_S$ providing a scale that gives good guidance on what false positive rate one should expect. Subsequently, one would obtain additional information to verify the veracity of an anomaly when possible.

For our proof of concept it was sufficient to only consider simple physical examples in this paper, but there are many real-world applications where our methods can be employed. Examples include searches for new particles in astroparticle and particle data, searches for gravitational wave signals in LIGO/VIRGO data, searches for exoplanets in transit and radial velocity data, as well as many more. In some of these cases the look-elsewhere penalty can be considerably large, reaching beyond 6 sigma. The problem is very general, as almost every search for unknown objects, events, new physics, or other phenomena whose existence is unknown, has to deal with the look-elsewhere effect. We note that while this work only explored the look-elsewhere effect from scanning over parameter space, we expect that similar methods can be applied to other manifestations of the look-elsewhere effect, for example scanning over models when fitting gravitational wave templates.

The goal of a data analyst searching for anomalies is to report the most promising anomalies in terms of having a small p-value, or a high Bayes factor. By clarifying the origins of the look-elsewhere effect and model complexity penalty for continuous parameters we hope to open the way to refinements in anomaly searches that can improve the overall success rate of a detection. This should be a common goal of any experimental analysis regardless of which school of statistics one belongs to.

Acknowledgements

We thank Benjamin Nachman for insightful comments on the manuscript, and Benjamin Wallisch for valuable discussion regarding example III. This research made use of the Cori supercomputer at the National Energy Research Scientific Computing Center (NERSC), a U.S. Department of Energy Office of Science User Facility operated under Contract No. DE-AC02-05CH11231. This material is based upon work supported by the National Science Foundation under Grant Numbers 1814370 and NSF 1839217, and by NASA under Grant Number 80NSSC18K1274.

7.8 Appendix A: Derivation of the CCDF of $\hat{q}_S$

The asymptotic (large q_L) CCDF of the global maximum of q_L is for a one-tail test is given in equation 7.18 as

$$P_{Q_L}(Q_L \geq q_L) = N\frac{1}{2}\tilde{F}_1(q_L) \tag{7.41}$$

$$= N\frac{1}{\sqrt{2\pi q_L}}e^{-q_L/2} + Nq_L^{-1/2}e^{-q_L/2}\mathcal{O}\left(q_L^{-1}\right), \tag{7.42}$$

where here we include the leading order correction, and drop hats and take $t = 1$ for convenience. Consider the transformation of variables to q_S, defined by

$$q_S \equiv g(q_L) \equiv q_L - 2\ln N + \ln 2\pi q_L. \tag{7.43}$$

It can be shown that the inverse of g is given by

$$q_L = g^{-1}(q_S) = W_0\left(\frac{N^2 e^{q_S}}{2\pi}\right) \tag{7.44}$$

$$= q_S + \ln\frac{N^2}{2\pi} - \ln\left(q_S + \ln\frac{N^2}{2\pi}\right) + \mathcal{O}\left(\frac{L_2}{L_1}\right), \tag{7.45}$$

where $W_0(z)$ is the principal branch of the Lambert W function. The asymptotic expansion has been performed in the final line, with the shorthand $L_i \equiv \ln^i \frac{N^2 e^{q_S}}{2\pi}$. Assuming N is constant to study the limiting behaviour, the CCDF of q_S is thus

$$P_{Q_S}\left(Q_S \geq q_S\right) = P_{Q_L}\left[Q_L \geq g^{-1}(q_S)\right] \tag{7.46}$$

$$= e^{-q_S/2}e^{-\mathcal{O}(L_2/L_1)}\left(1 - \frac{\ln\left(q_S + \ln\frac{N^2}{2\pi}\right) + \mathcal{O}\left(\frac{L_2}{L_1}\right)}{q_S + \ln\frac{N^2}{2\pi}}\right)^{-1/2} + \mathcal{O}\left(\frac{e^{-q_S/2}}{q_S + \ln\frac{N^2}{2\pi}}\right) \tag{7.47}$$

$$\to e^{-q_S/2}, \tag{7.48}$$

where the limit corresponds to either $N \to \infty$ or $q_S \to \infty$. This means the result still applies asymptotically in the absence of the look-elsewhere effect ($N = 1$).

Chapter 8

Self-calibrating the look-elsewhere effect: Fast evaluation of the statistical significance using peak heights

The contents of this chapter was originally published in [61],

> Self-Calibrating the Look-Elsewhere Effect: Fast Evaluation of the Statistical Significance Using Peak Heights
> *Bayer A.E., Seljak U., Robnik J.* (arXiv:2108.06333)
> MNRAS 508-1 (Nov 2021) 1346–1357

In experiments where one searches a large parameter space for an anomaly, one often finds many spurious noise-induced peaks in the likelihood. This is known as the look-elsewhere effect, and must be corrected for when performing statistical analysis. This chapter introduces a method to calibrate the false alarm probability (FAP), or p-value, for a given dataset by considering the heights of the highest peaks in the likelihood. Specifically, we derive an equation relating the global p-value to the rank and height of local maxima. In the simplest form of self-calibration, the look-elsewhere-corrected χ^2 of a physical peak is approximated by the χ^2 of the peak minus the χ^2 of the highest noise-induced peak, with accuracy improved by considering lower peaks. In contrast to alternative methods, this approach has negligible computational cost as peaks in the likelihood are a byproduct of every peak-search analysis. We apply to examples from astronomy, including planet detection, periodograms, and cosmology.

8.1 Introduction

When searching a large parameter space for a signal, with limited a priori knowledge of the signal's location, the likelihood distribution is often multimodal with the vast majority of peaks corresponding to spurious noise-induced events. This is known as the look-elsewhere effect, or problem of multiple comparisons, and must be accounted for when performing a hypothesis test to avoid reporting a false detection [263, 331].

This effect is particularly prevalent in astronomy, with numerous examples including searching for gravitational waves and exoplanets. In the gravity wave example, one searches for a signal of many different possible known shapes and unknown time, which can lead to a large look-elsewhere effect, modifying the p-value by many orders of magnitude (see e.g. [88, 4, 261]). A similarly large effect occurs when searching for exoplanetary transits in stellar photometry experiments such as Kepler [79], where the period, phase, and other properties of the transit are unknown [37, 43, 44, 140]). Additionally, the look-elsewhere effect can occur in wavelet analysis: this has many astronomical applications, one of which is detecting asteroid families in the main belt [see e.g. 38, 45, 40, 39, and references therein].

The look-elsewhere effect is also prominent in searches for new particles, where the mass of the particle is unknown: this issue gained much attention when the LHC detected the Higgs boson [1, 102]. Just like collider searches, astroparticle experiments also suffer from the look-elsewhere effect, with examples including: constraining the dark matter self-annihilation cross-section via gamma ray emission from galaxy clusters [23], searching for WIMPs via charged cosmic rays [302], searching for non-baryonic dark matter via X-ray emission from the Milky Way [324], explaining the source of high energy astrophysical neutrinos [2, 151], and in the spectral analysis of solar neutrinos [301]. The look-elsewhere effect additionally appears in inflationary cosmology when searching for anomalies in the primordial power spectrum [155, 156, 201], and when detecting planar structures in the satellite systems of galaxies [98].

The look-elsewhere effect is also relevant in numerous areas outside of physics. In biology, modern DNA sampling techniques can be used to perform genetic association to find links between genotypes and phenotypes [32, 348]. When large DNA sequences are used there is a high probability of obtaining spurious signals and thus a large look-elsewhere effect. Another medical example is the process of testing the effectiveness of drugs in clinical trials [298]. Furthermore, it is important to consider when attempting to find hidden prophecies in ancient religious texts [259]. The look-elsewhere effect is ubiquitous in physics and beyond and there is thus much motivation for a fast method to account for it.

When performing a hypothesis test, frequentists typically consider the p-value, whereas Bayesians consider the Bayes factor. The p-value is often referred to as the false positive rate (FPR) or false alarm probability (FAP), as it quantifies how often a given test statistic is expected to take on a particular, or more extreme, value under the assumptions of the null hypothesis. Hence, the smaller the p-value, the less likely the null hypothesis and the larger the statistical significance of the alternative hypothesis. Typically one considers the p-value of the likelihood ratio between the alternative and the null hypotheses, in which case

the look-elsewhere effect causes an increase in the p-value at a fixed value of the likelihood: discretely speaking, if one performs N trials, the probability of a spurious event increases by a factor of N. Conversely, the likelihood required to achieve a given p-value is increased by the look-elsewhere effect, meaning one needs to find peaks with a larger likelihood to achieve a given statistical significance.

A brute force method to account for the change in the p-value is to perform simulations of the null hypothesis and determine the p-value numerically. This, however, is extremely computationally expensive: for example, to determine the 5-sigma level, which corresponds to a p-value of $\sim 10^{-7}$, one would need to perform $\sim 10^7$ simulations. To achieve a balance between efficiency and accuracy, scientists have often applied the theory of [133, 134] to find an upper bound for the p-value. [37, 43, 44] applied this to the Lomb-Scargle periodogram [249, 314], and various generalizations, by performing intricate analytical calculations. However, these calculations depend on the type of periodogram considered and are thus not applicable to more general situations. For example, in the task considered by [41] it was found that analytic approximations are either inaccurate or slow to compute. Furthermore, [174] applied the theory of [133, 134] in the context of particle searches by using the expected number of upcrossings to approximate the asymptotic p-value via the Taylor approximation. This method still requires multiple simulations, although fewer than the brute force approach.

An alternative approach to account for the look-elsewhere effect, which requires neither simulations nor model-specific calculation, has been recently developed by [60]. Drawing a connection with Bayesian methodology, it associates the look-elsewhere effect *trials factor* with the prior-to-posterior volume ratio. This method has been shown to be effective for a variety of models when evaluating the Bayes factor using the Laplace approximation and the posterior volume using the Hessian matrix. This makes the computation very fast, however, when considering models of increased complexity where the Laplace approximation is not valid, one may have to employ Monte Carlo (MC) methods (e.g. [106]) to accurately evaluate the Bayes factor, which would increase the computational time.

In this paper we work towards a general approach to account for the look-elsewhere effect, which requires neither simulation, nor model-specific calculations, nor explicit evaluation of the Bayes factor. To achieve this, we consider the distribution of peak heights in the multimodal likelihood computed from the data. We show that, given this information alone, one can estimate the p-value directly from the likelihood of the data in examples of varying complexity. Since the peak heights are a byproduct of a peak search, this information is readily available and provides a very fast way to estimate the p-value. Because this approach accounts for the look-elsewhere effect by using information from the data alone, we name the method *self-calibration*.

The paper is organized as follows. Section 8.2 reviews the look-elsewhere effect and defines quantities such as the trials factor. Section 8.3 then motivates the method to self-calibrate the p-value and trials factor. This is then illustrated in Section 8.4 for a variety of examples related to planet detection (spots in exoplanetary transit light curves, the LS periodogram, and a Kepler exoplanet search) and an example from cosmology (searching for oscillatory features in the primordial power spectrum). Finally, conclusions are given in

Section 8.5.

8.2 Background

This section briefly summarizes the look-elsewhere effect (see [60] for a more thorough discussion). We consider a model with M parameters, $\boldsymbol{z}$, such that z_1 is the amplitude of a signal, and $\boldsymbol{z}_{>1}$ describes the properties of the signal that one is scanning over, or fitting for. For example, searching for the signal of a planet transit with amplitude z_1, orbital period z_2, and phase z_3.

From a frequentist perspective, one is interested in comparing the data $\boldsymbol{x}$ against the null hypothesis H_0 of there being no signal ($z_1 = 0$). Writing the likelihood as $p(\boldsymbol{x}|\boldsymbol{z})$, a common test statistic to consider is related to the likelihood ratio,

$$q_L(\boldsymbol{z}) \equiv 2 \ln \frac{p(\boldsymbol{x}|\boldsymbol{z})}{p(\boldsymbol{x}|\boldsymbol{z}_0)}, \tag{8.1}$$

where $\boldsymbol{z}_0$ represents the values of the parameters under H_0. To assess the significance one typically considers the maximum value of q_L, denoted $\hat{q}_L$, but below we will generalize this concept. For a Gaussian likelihood, q_L is equal to the difference in χ^2 between the null and signal hypotheses, thus we will often simply refer to this as the chi-squared. In such a case, and in the absence of the look-elsewhere effect, $\sqrt{\hat{q}_L}$ gives the number-of-sigma significance. This can then be related to the p-value, often referred to as the false positive rate (FPR) or false alarm probability (FAP), depending on the problem in question: for example in the case of a chi-squared random variable with s degrees of freedom, the p-value is given by the complementary cumulative distribution function of a chi-squared random variable with s degrees of freedom.The p-value in the absence of the look-elsewhere effected is referred to as the *local* p-value.

In the presence of the look-elsewhere effect the p-value must be corrected, sometimes by many orders of magnitude. This is often done by introducing the trials factor, N, such that the look-elsewhere-corrected p-value is parameterized as

$$P(\hat{Q}_L > \hat{q}_L) = N P_{\text{local}}(\hat{Q}_L > \hat{q}_L), \tag{8.2}$$

where P_{local} is the local p-value, and $\hat{Q}_L$ is the random variable associated with $\hat{q}_L$. The look-elsewhere-corrected p-value is often referred to as the *global* p-value as it considers the probability that the global maximum of the likelihood occurs above a specific threshold. The trials factor thus quantifies the extent of the look-elsewhere effect, with $N = 1$ corresponding to no look-elsewhere effect, and progressively larger values corresponding to a more severe look-elsewhere effect. Computing the trials factor is the main challenge in accounting for the look-elsewhere effect, and is typically done either by (i) performing model-dependent analytical calculations [37, 43, 44], (ii) running many numerical simulations [174], or (iii) by evaluating the Bayes factor [60]. In the next section we present self-calibration as a fast method to estimate N and the p-value directly from the data, without any expensive computation.

8.3 The Self-calibration Method

In this section we introduce our proposal for self-calibrating the p-value using the heights of peaks in the likelihood. We present a Bayesian derivation in the appendix for this chapter (Section 8.6), but give a more concise frequentist motivation in this section.

In the presence of a large look-elsewhere effect ($N \gg 1$), the parameter space will contain many peaks. Various works have derived the number of upcrossings (or the number of local maxima in the case of multidimensional fields) that breach a particular threshold, τ, of the chi-squared [see e.g. 303, 7, 133, 134, 31]. One can write the expected number of upcrossings as

$$\langle n_{\rm up}(\tau)\rangle = C\tau^{\alpha}e^{-\tau/2}, \tag{8.3}$$

where α and C are model dependent quantities, and we have ignored boundary effects which are negligible for large trials factor N. Note that the trials factor is absorbed into C.

By evaluating the expected number of upcrossings from Eq. 8.3 at the maximum chi-squared value, denoted $\hat{q}_L$, one obtains an asymptotic (large $\hat{q}_L$) approximation of the p-value as

$$P(\hat{Q}_L \geq \hat{q}_L) \simeq \langle n_{\rm up}(\hat{q}_L)\rangle = C\hat{q}_L^{\alpha}e^{-\hat{q}_L/2}. \tag{8.4}$$

In the case of chi-squared random variable with s degrees of freedom $\alpha = (s-1)/2$ [134]. Many problems of physical interest obey this form, for example a periodogram with 1 harmonic signal corresponds to $s = 2$ [37], or $s = 2h$ in the case of h harmonic signals [42], while particle physicists hunting for a mass resonance typically consider $s = 1$ [124]. We will assume this form of α in the remainder of the text.[1]

On the other hand, the coefficient C can be complicated to compute as it is sensitive to the look-elsewhere effect and has a high degree of model dependence. Efforts have been made to compute it for many different scenarios [37, 42, 43, 44]; however, this requires a specific calculation for each example considered, and for many problems can become analytically intractable [41].

To avoid having to compute C, we eliminate it by combining Eqs. 8.3 and 8.4. This gives

$$P(\hat{Q}_L > \hat{q}_L) \simeq \langle n_{\rm up}(\tau)\rangle \left(\frac{\hat{q}_L}{\tau}\right)^{(s-1)/2} e^{-(\hat{q}_L-\tau)/2} \tag{8.5}$$

$$= e^{-\frac{1}{2}\left[\hat{q}_L-\tau-2\ln\langle n_{\rm up}(\tau)\rangle-(s-1)\ln\frac{\hat{q}_L}{\tau}\right]}. \tag{8.6}$$

In essence, Eq. 8.6 is calibrating the expected number of upcrossings at the maximum value of the chi-squared, $\hat{q}_L$, by using the number of upcrossings at a lower value from the chi-squared, τ.

This relation in Eq. 8.6 was also found by [174] using a slightly different argument. In order to apply their method, [174] suggests to evaluate $\langle n_{\rm up}(\tau)\rangle$ by performing thousands

[1] One can consider non-linear models for which τ^{α} is replaced by a model-dependent polynomial function of τ [see e.g. 43], however we will demonstrate that this is often unnecessary to a good approximation in §8.4.

of simulations and computing the numerical average. While this requires fewer simulations than evaluating the p-value directly it can still be computationally expensive [17].

In self-calibration we bypass such simulations, and quickly obtain the p-value directly from a single dataset. We do this by firstly noting that the number of peaks is a good approximation to the number of upcrossings in the asymptotic $\hat{q}_L$ limit. Using peaks instead of upcrossings is beneficial because peaks are a byproduct of any peak-search analysis, making them readily available. Secondly, we use the fact that, by definition, there are n peaks with q_L larger than or equal to the n^{th} highest peak, $q_L^{(n)}$. We thus take $\tau \approx q_L^{(n)}$ and $\langle n_{\text{up}}(\tau)\rangle \approx n$. Substituting this into Eq. 8.6 enables application to a single dataset, negating the need for simulations. Finally, we apply the Šidák correction [338] to Eq. 8.6 to improve the non-asymptotic behaviour. The Šidák correction replaces the asymptotic p-value such that $P \to 1 - e^{-P}$, to provide better agreement in the non-asymptotic (large P) regime [see e.g. 60].

In this paper we will consider models with 1 amplitude parameter and s other parameters, i.e. s degrees of freedom, thus the total number of parameters is given by $M = s + 1$. Using bars to denote self-calibrated values, the self-calibrated p-value is thus given by

$$\bar{P}(\hat{Q}_L > \hat{q}_L) \equiv 1 - \exp\left(-e^{-\frac{1}{2}\left[\hat{q}_L - \tau_n - 2\ln n - (M-2)\ln\frac{\hat{q}_L}{\tau_n}\right]}\right), \tag{8.7}$$

where τ_n is known as the *threshold*. We provide discussion on choices of τ_n in the appendix for this chapter (Section 8.6), but $\tau_n \approx q_L^{(n)}$ is a suitable approximation which we will often employ. Hence, Eq. 8.7 relates the p-value to an exponential function of the difference in peak height between the highest peak and the n^{th} highest peak (via the $\hat{q}_L - \tau_n \approx \hat{q}_L - q_L^{(n)}$ term), and some logarithmic correction terms that depend on the choice of n and the dimensionality of the model M. We will discuss the choice of n in Section 8.4.

The form of the self-calibrated calibrated p-value can be more succinctly written as

$$\bar{P}(\hat{Q}_S > \hat{q}_S) \equiv 1 - \exp\left(-e^{-\bar{\hat{q}}_S/2}\right), \tag{8.8}$$

where

$$\bar{\hat{q}}_S \equiv \hat{q}_L - \tau_n - 2\ln n - (M-2)\ln\frac{\hat{q}_L}{\tau_n}. \tag{8.9}$$

It is beneficial to work with the q_S test statistic instead of q_L as the look-elsewhere dependence is absorbed into q_S, making the p-value relation in Eq. 8.8 independent of the look-elsewhere effect [60]. In the case of a one-tailed test (i.e. positive amplitude), it follows from [60] that the self-calibrated statistical significance, or number of sigma, $\bar{S}$, is given by

$$\bar{S} = \sqrt{\bar{\hat{q}}_S - \ln 2\pi\bar{\hat{q}}_S}, \tag{8.10}$$

for large $\bar{\hat{q}}_S$. Note for sufficiently large $\bar{\hat{q}}_S$, $\bar{S} \approx \sqrt{\bar{\hat{q}}_S}$. Moreover, the self-calibrated trials factor, $\bar{N}$, is given by

$$2\ln\bar{N} = \tau_n + 2\ln n + \ln 2\pi\hat{q}_L + (M-2)\ln\frac{\hat{q}_L}{\tau_n}. \tag{8.11}$$

The above discussion concerns the distribution of peak heights for pure noise. In practice, a dataset might contain one, or more, physical peaks; applying Eq. 8.7 to such a dataset would give the significance of a signal under the null hypothesis, i.e. assuming all peaks are noise. The existence of physical peaks will cause an overestimation of τ_n and in turn an overestimation of the p-value, or an underestimation of the significance. This could introduce false negatives, but not false positives, making this a conservative approach. However, this overestimation of the p-value will be small for n larger than a few due to the slowly varying logarithmic term $\ln n$. In cases where there are multiple peaks introduced by physical sources, one can reduce this effect by iteratively removing non-maximal physical signals and appropriately relabeling the peaks in terms of n. More generally, Eq. 8.9 shows that to self-calibrate $\hat{q}_S$ for $M = 2$ one must correct $\hat{q}_L$ by $\tau_n + 2\ln n$. Plotting this correction as a function of n and comparing to the expected variance of $\tau_n + 2\ln n$ will indicate if the data is consistent with noise and thus whether the result of self-calibration is reliable.

To perform self-calibration there is thus one parameter to be chosen, the index of the peak n. We will explore this choice in depth in Section 8.4. Meanwhile, it is instructive to consider Eq. 8.9 for $n = 2$. In this case, the self-calibrated $\hat{q}_S$ is simply given by the χ^2 difference between the highest and second highest peaks (using $\tau_n \approx q_L^{(n)}$), apart from small logarithmic corrections. If the highest peak is known to be physical, and the remaining peaks noise, this would give the χ^2 difference between the physical peak and the highest noise-induced peak. So, in its simplest form, self-calibration corresponds to computing the look-elsewhere-corrected p-value from this difference in χ^2. Furthermore, for sufficiently large $\hat{q}_S$, the look-elsewhere-corrected chi-squared, S^2, approximately equals this difference in χ^2.

It is sometimes the case that the χ^2 is only known up to a constant factor: for example in periodograms used for radial-velocity exoplanet searches, unknown jitter effects mean the measurement errors are only known up to a constant factor [see e.g. 37]. In such cases, it is common practice to consider different normalizations of the χ^2 to cancel out this factor, resulting in different analytical formulae for the p-value for each choice of normalization. These formulae can be complicated and have dependence on the number of data points, N_d. Instead, we can use self-calibration to analyse such cases in a general manner. If we only know the chi-squared up some constant factor, k, we have $q_L = kq'_L$ and $\tau_n = k\tau'_n$. We can thus self-calibrate k by applying equation 8.9 twice using the n^{th} and m^{th} peaks, resulting in the following simultaneous equation:

$$\bar{\hat{q}}_S = k(\hat{q}'_L - \tau'_n) - 2\ln n - (M-2)\ln\frac{\hat{q}'_L}{\tau'_n} \tag{8.12}$$

$$= k(\hat{q}'_L - \tau'_m) - 2\ln m - (M-2)\ln\frac{\hat{q}'_L}{\tau'_m}. \tag{8.13}$$

Solving for k gives the self-calibrated k as

$$\bar{k} = \frac{2\ln\frac{n}{m} + (M-2)\ln\frac{\tau'_m}{\tau'_n}}{\tau'_m - \tau'_n}. \tag{8.14}$$

This can then be substituted into Eq. 8.12 to compute the self-calibrated $\hat{q}_S$, and in turn the significance using Eq. 8.10.

8.4 Results

In this section we apply self-calibration of the p-value to various astrophysical examples. We start with a search for a Gaussian peak in white noise in subsection 8.4. This is a very common example when searching for a single event in the data: it could for example correspond to detecting a single transient in an exoplanetary lightcurve (e.g. a starspot). We then consider the Lomb-Scargle (LS) periodogram in subsection 8.4, where we allow for data with non-fixed spacing to investigate the effects of aliasing on self-calibration. We then study a more complex (non-linear) planet transit model in subsection 8.4, where we apply it to a Kepler exoplanet search. Finally, in subsection 8.4 we consider an example from cosmology, namely searching for oscillatory features in the primordial power spectrum using data from Planck [289].

Single Transient

We now apply self-calibration to an example of a search for a Gaussian peak in a white noise time series. Physically speaking, this could correspond to searching for Gaussian-like transients in lightcurves (e.g. starspots), Gaussian-like peaks in spatial maps, as well as numerous other examples. We consider a time series of data measurements $y(x)$ comprising of N_d data points, $\boldsymbol{x} = \{x^i\}_{i=1}^{N_d}$, with spacing $x^{i+1} - x^i = 1$. Each $y^i \equiv y(x^i)$ measurement has normally distributed noise with zero mean and unit variance. We seek a signal of the form $AN(x|x_*, \sigma_*)$ where A is the amplitude, and $N(x|x_*, \sigma_*)$ is a normal distribution with mean x_* and width σ_*. For these standard normal measurements, q_L from Eq. 8.1 equals the difference in chi-squared between the null and signal hypotheses:

$$q_L(\boldsymbol{x}|A, x_*, \sigma_*) = \sum_{i=1}^{N_d} \left[y^i\right]^2 - \left[y^i - AN(x|x_*, \sigma_*)\right]^2 . \tag{8.15}$$

To test self-calibration we first consider the case of pure noise. To study different choices of n we consider 10^3 pure noise realizations. We use a signal width of $\sigma_* = 2$ such that the width of the peak is larger than the spacing of data points. Hence, this is a 2 dimensional problem, $M = 2$, with parameters A and x_*. We consider 6 prior volumes in the range 250 to 10^4. Note that $V_{x_*,\text{prior}} = 10^4$ contains approximately 700 peaks.

Fig. 8.1 shows the complementary cumulative distribution of $\bar{\hat{q}}_S$ (Eq. 8.9) for different choices of n. There is excellent agreement with Eq. 8.8 for all choices of n for the full range of p-values, suggesting the self-calibrated $\hat{q}_S$ is distributed as the true $\hat{q}_S$. To investigate the correspondence between the self-calibrated values and the values computed using the prior-to-posterior volume, we consider the bias and standard deviation of $\bar{P}/P - 1$ and $\bar{S} - S$ on the

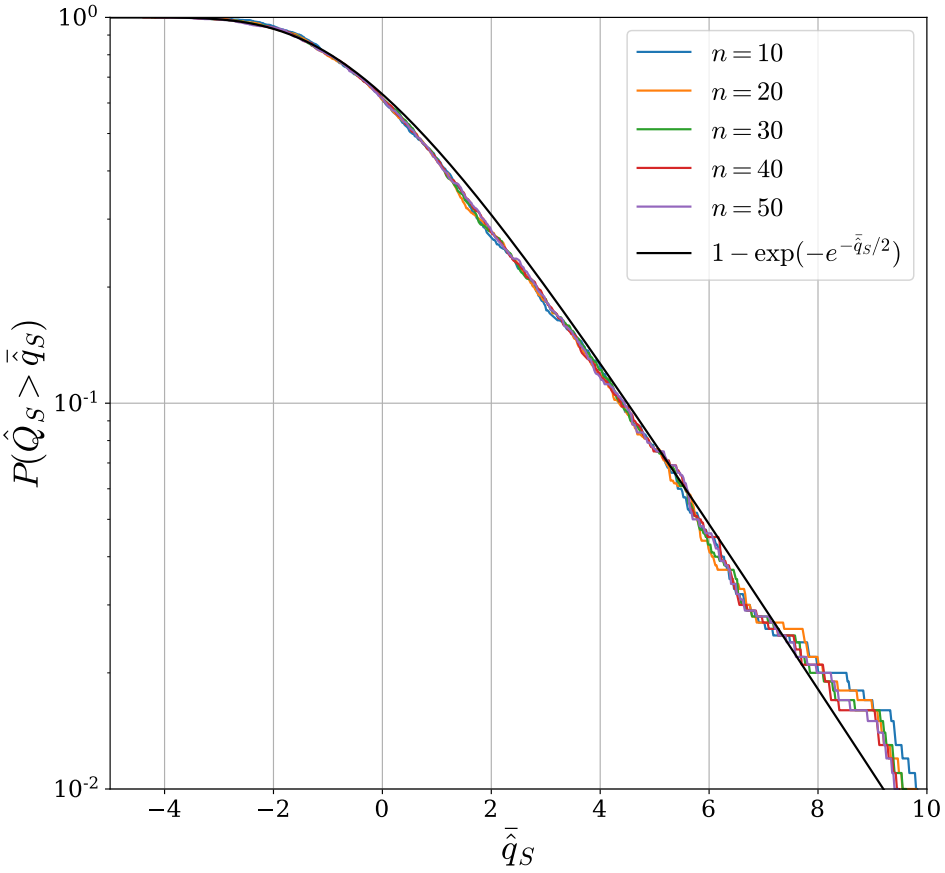

Figure 8.1: Complementary cumulative distribution of $\bar{\hat{q}}_S$ averaged over 10^3 simulations with no signal ($A = 0$). Self-calibration is performed with a variety of choices of n, all of which agree with the theoretical expectation of Eq. 8.8 (black line). The noise at high $\bar{\hat{q}}_S$ is due to the finite number of simulations used.

left and right of Fig. 8.2 respectively. It can be seen that the bias is close to 0, with variance decreasing with n and roughly levelling off at around $n \sim 20 - 40$. The plots also show independence of prior volume, provided n is sufficiently small. As V_{prior} is decreased, the number of peaks in the data decreases, and thus the maximum possible n that can be used decreases. It can be seen that the smaller the prior volume, the earlier in n self-calibration picks up a bias. This bias continues to grow with n, however we cut the lines for the sake of neatness. Thus the bias remains constant up to a particular value of n which grows with $V_{x_*,\text{prior}}$, i.e. the number of peaks in the data.

The variance of the p-value levels of at around $n = 20$. At this n the rms of the fractional error of the p-value is around 0.2, meaning the self-calibrated p-value will have an error of $\pm 20\%$, corresponding to a error in S of ± 0.25. We see that even for the case of $V_{x_*,\text{prior}} = 250$,

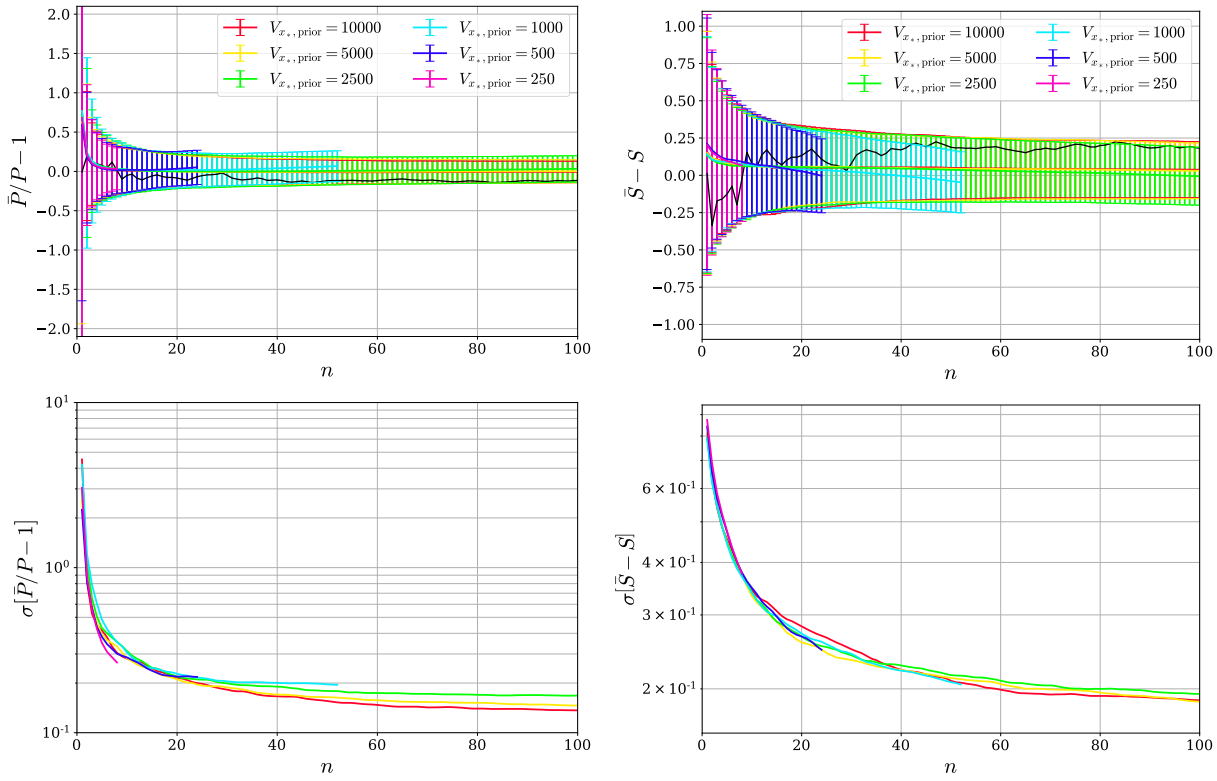

Figure 8.2: Analysis of the bias and variance of the self-calibrated p-value and significance as a function of n. (Top left) The mean over 10^3 realizations of $\bar{P}/P - 1$ as a function of n. Error bars represent the standard deviation. (Bottom left) The standard deviation of $\bar{P}/P - 1$ as a function of n. Similarly, (top right) is the mean of $\bar{S} - S$ as a function of n, with (bottom right) the standard deviation. In all cases 6 prior volumes are considered. An example of a single realization for $V_{x_*,\text{prior}} = 10^4$ is shown by the black line.

which contains approximately 20 peaks per realization, one can self-calibrate the p-value to $\pm 30\%$ using $n = 8$. This level of accuracy suffices for many applications where one wishes to quickly quantify the p-value. A final feature of Fig. 8.2 are the black lines in the upper plots, which show the bias for a single realization, where it can be seen that there is correlation between different values of n. One might consider taking the average (e.g. the median) value of the self-calibrated significance computed over a range of n to reduce the effect of random fluctuations in peak heights from a single realization; however, this correlation between peaks means that such averaging is not guaranteed to perfectly remove the effect.

As discussed earlier in Section 8.3, the rms of $\tau_n + 2\ln n$ as a function of n can be used as a diagnostic to verify whether the data is consistent with pure noise, and thus whether the result of self-calibration is reliable. This is plotted in Fig. 8.3. It can be seen that the variance is prior independent up to a value of n that grows with $V_{x_*,\text{prior}}$. The line for

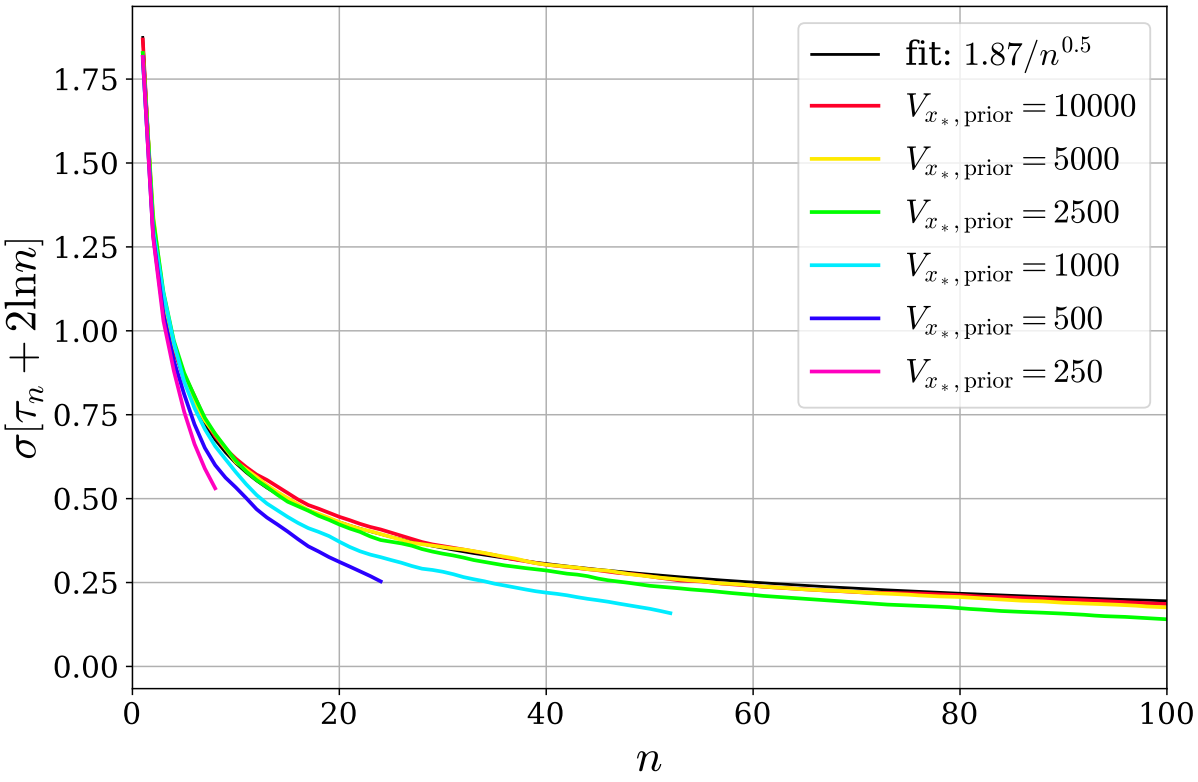

Figure 8.3: The standard deviation of $\tau_n + 2\ln n$ as a function of n for a variety of prior volumes. The black line shows a fit to the $V_{\text{prior}} = 10,000$ line given by Eq. 8.16.

$V_{x_*,\text{prior}} = 10^4$ is suitably converged in the range $1 \leq n \leq 100$; a fitting formula for this line is given by

$$\sigma\left[\tau_n + 2\ln n\right] = \frac{1.87}{\sqrt{n}}. \tag{8.16}$$

While this fitting formula was derived for the $M = 2$ white noise model considered in this section, we expect it to still be a good indicator for other models. Thus when performing self-calibration, we advise plotting $\tau_n + 2\ln n$ against n for the dataset, and comparing this with the error envelope obtained from the above fitting formula. If the line falls within the error envelope one can be confident that self-calibration was performed reliably. If not, it could indicate the presence of physical peaks in the data, in which case one must identify and remove these physical signals from the data, and then repeat the self-calibration process, to avoid an overly conservative estimate of the p-value.

Having studied self-calibration in the context of pure noise, we now illustrate the method applied to data with a physical signal. Fig. 8.4 shows a typical example of a peak at $x_* = 300$ with true significance $S \approx 5.4$. The left plot shows the distribution of q_L projected onto the x_* dimension. It can be seen that there are numerous peaks, including a physically injected

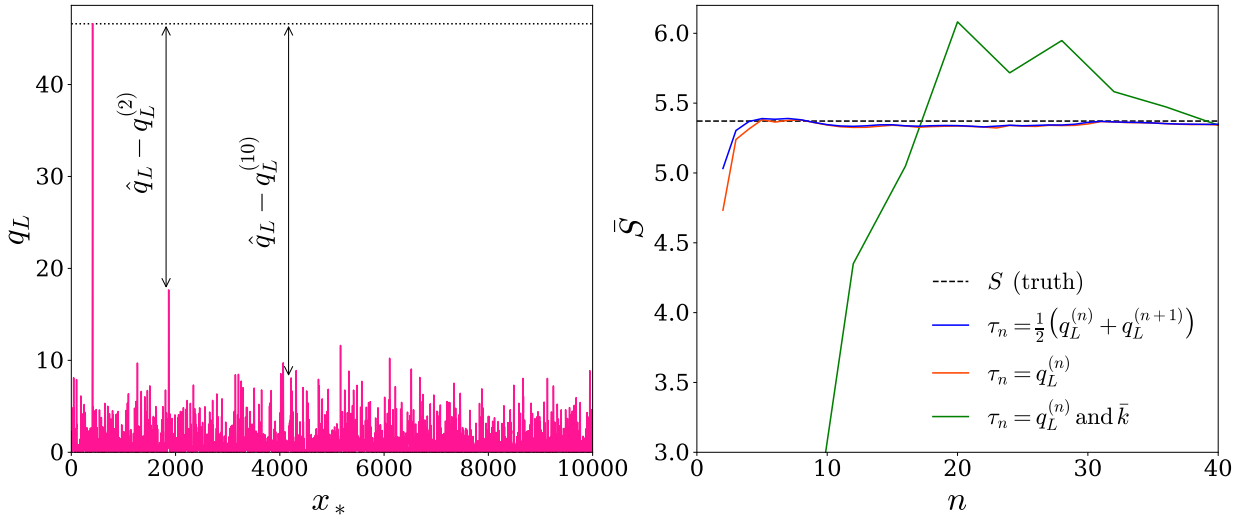

Figure 8.4: An example of self-calibration for a dataset with a true signal at $x_* = 300$ and true significance $S = 5.38$. *Left panel*: Distribution of q_L projected onto the x_* axis, i.e. having maximized over A for each x_*. The maximum q_L is defined as $\hat{q}_L \equiv q_L^{(1)} \approx 47$, corresponding to a 6.8 sigma signal. Two arrows illustrate $\hat{q}_L - q_L^{(n)}$ for $n = 2, 10$, with $q_L^{(2)} \approx 18$ and $q_L^{(10)} \approx 9$. *Right panel*: The self-calibrated value of the statistical significance $\bar{S}$ of the peak to be a true peak, as a function of n with two options for τ_n (blue and orange). We see these lines are well converged for $n \gtrsim 5$. We also plot the results of self-calibration when one does not know the normalization of q_L and must additionally self-calibrate its normalization, denoted by k (green). In this case results are noisier — due to the error in the estimate of k — but are still within 1-sigma for $n \gtrsim 15$ and within 0.1-sigma for $n \gtrsim 40$.

SELF-CALIBRATING THE *P*-VALUE
(1) By scanning over $\boldsymbol{z}_{>1}$ compute the likelihood ratio, and in turn q_L, of all high peaks. Denote the highest as $\hat{q}_L$.
(2) Choose n (as discussed above).
(3) Compute τ_n, the average q_L of the n^{th} and $(n+1)^{\text{th}}$ highest noise peaks.
(4) Evaluate the p-value of the highest peak as $1 - \exp\left(-e^{-\bar{\hat{q}}_S/2}\right)$, where $\bar{\hat{q}}_S = \hat{q}_L - \tau_n - 2\ln n - (M-2)\ln\frac{\hat{q}_L}{\tau_n}$.
(5) Evaluate the statistical significance, or number of sigma, approximately as $\bar{S} = \sqrt{\bar{\hat{q}}_S - \ln 2\pi\bar{\hat{q}}_S + 2\ln t}$.
(6) Plot $\tau_n + 2\ln n$ vs n together with the theoretical error envelope to verify the peaks are consistent with noise. If not, remove additional physical peaks and repeat. One can then also repeat the process to evaluate the significance of each physical peak.

Table 8.1: Algorithm for self-calibration of the p-value and statistical significance.

peak at $x_* \approx 300$ with $\hat{q}_L \equiv q_L^{(1)} = 46.6$. The right of Fig. 8.4 shows the self-calibrated significance compared to the true significance, as a function of n and for different choices of threshold τ_n given in Eq. 8.41. The simplest form of self-calibration is to use the difference in height of the highest two peaks, i.e. $\tau_n = q_L^{(n)}$ with $n = 2$. Using Eq. 8.9 this gives $\bar{\hat{q}}_S = \hat{q}_L - q_L^{(2)} - 2\ln 2$. It can be seen that such an approach leads to under-predicting the number of sigma S by 0.6, so even at $n = 2$ we obtain a useful diagnostic. Nevertheless, to achieve a better estimate of S one can use higher n. By considering the blue line it can be seen that the average of the 1st and 2nd noise peaks produces a better estimate than the 1st noise peak alone, however taking such an average becomes unnecessary for larger n, and convergence is achieved by around $n = 5$. This is typical for peaks of significance 5 sigma and above, thus for cases of physical interest one can often use very low values of n.

We also show the self-calibrated number of sigma for when the normalization of q_L is unknown. While this is not a typical example, this is often the case when the measurement error is not known in radial-velocity exoplanet searches [see e.g. 37]. To tackle such problem with self-calibration, one can self-calibrate using 2 peaks to estimate the normalization of q_L, k, using Eq. 8.14. While one is free to use any choice of peaks, indexed by n and m, we choose $m = n/3$ for this plot — it is also typically good practice to choose $|m - n| \gtrsim 5$ to avoid correlation between adjacent peak heights. The self-calibrated $\bar{S}$ in this case is depicted with the green line in Fig. 8.4. It can be seen that there is now more noise in the estimated number of sigma, due to the noise introduced by having to estimate k, however results are still within 1-sigma of the correct value for $n \gtrsim 15$ and within 0.1-sigma for $n \gtrsim 40$.

In summary, we have shown that one can accurately compute the look-elsewhere corrected p-value by considering the heights of likelihood peaks, without needing to evaluate the posterior volume or performing simulations. This is true even in the non-asymptotic limit. One can reliably self-calibrate the p-value to order 10% accuracy, or equivalently the number-of-sigma significance to ± 0.25, using a single dataset. The self-calibration algorithm for the p-value is summarized in Table 8.1.

Lomb-Scargle Periodogram

The Lomb-Scargle (LS) periodogram [249, 314] (see [372] for a review) corresponds to a search for a sine wave in a white noise background. Thus instead of the Gaussian signal considered in the previous subsection, we now consider a signal of the form $A\sin(\omega t + \phi)$, where t is time, $\omega = 2\pi/f$ is the angular frequency of the orbit, and ϕ is the phase. For standard normal measurements, y^i, q_L from Eq. 8.1 equals the difference in chi-squared between the null and signal hypotheses:

$$q_L(\boldsymbol{t}|A,\omega,\phi) = \sum_{i=1}^{N_d} \left[y^i\right]^2 - \left[y^i - A\sin(\omega t + \phi)\right]^2 . \tag{8.17}$$

Physically this could correspond to the radial velocity of a star caused by a planet orbiting it [37], to the spectral analysis of solar neutrinos [301], or to many other examples.

One of the key features of the LS periodogram over the classical periodogram is its consideration of non-uniform time measurements. For a time series with span T, we thus consider N_d time measurements $\boldsymbol{t} = \{t^i\}_{i=1}^{N_d}$ between $0 \leq t < T$, with both fixed-uniform spacing and randomly distributed measurements. For data with fixed-uniform spacing, there is no information beyond the Nyquist frequency, $f_{\rm Nyq} = 1/(2\Delta t)$, where $\Delta t = (N_d - 1)/T$ is the uniform data spacing. All frequency features beyond the Nyquist are indistinguishable from their aliases in the $f < f_{\rm Nyq}$ region, so in the case of fixed-uniformly space data, the aliasing follows a periodic pattern with period $2f_{\rm Nyq}$.[2] Moreover, for uniform sampling we only need to consider the discrete frequencies $f_i = i/(2N_d\Delta t)$, and since Fourier modes are orthogonal on this uniform basis the amplitudes of the different frequencies are independent of each other. In this situation we can fit for each frequency separately and our self-calibration can be applied without error.

However, data without fixed-uniform spacing can extract frequency features beyond the Nyquist frequency. As such, the periodic signal associated with a periodogram can introduce aliasing, whereby a physical signal at a particular frequency will give rise to peaks at both the true frequency and various other frequencies. Moreover, the peaks become correlated, and a proper analysis requires a joint fit of all the peaks [163]. Instead, we can try to analyze the peaks individually without the joint analysis, ignoring the correlations: we will show that this still provides reasonably accurate results. We refer the reader to the work of [37] for a detailed mathematical description of the effects of correlations/aliasing on the p-value in the context of the LS periodogram.

We consider an LS periodogram with time of observation $T = 1$, and $N_d = 100$ measurements. The Nyquist frequency is thus $f_{\rm Nyq} = 50$. We additionally apply a low frequency cutoff of $f_{\rm min} = 0.5$ corresponding to the minimum frequency detectable for $T = 1$. We use 10^6 realizations to numerically compute the distribution of the self-calibrated $\hat{q}_S$ and in turn the p-value.

The left panel of Fig. 8.5 shows the distribution of self-calibrated $\hat{q}_S$ for this LS periodogram. We consider self-calibration using $n = 10$, but have checked results are robust to this choice. The solid blue line corresponds to fixed-uniform data spacing with $f_{\rm max} = f_{\rm Nyq}$. It can be seen that there is good agreement with Eq. 8.8: this is to be expected as the maximum of the LS periodogram is known to correspond to a chi-square with 2 degrees of freedom at any fixed phase [314], and so obeys Eq. 8.2.

The dashed and dotted blue lines of Fig. 8.5 correspond to the p-value for random uniformly spaced data (i.e. the data points t^i are drawn from a uniform distribution). In this case we consider maximum frequencies of both 50 and 500, as non-uniformly space data produces information beyond the Nyquist frequency. It can be seen that the results of self-calibration for $f_{\rm max} = f_{\rm Nyq} = 50$ shows slightly worse agreement than the fixed-spacing case, and that the agreement worsens as we increase to $f_{\rm max} = 10 f_{\rm Nyq} = 500$. The reason for this worsening is that considering non-uniformly spaced data, and frequencies above the Nyquist,

[2] Note the LS periodogram contains an additional symmetry $P(f) = P(-f)$, thus $P(f) = P(2f_{Nyq} - f)$, and there is no information beyond $f_{\rm Nyq}$.

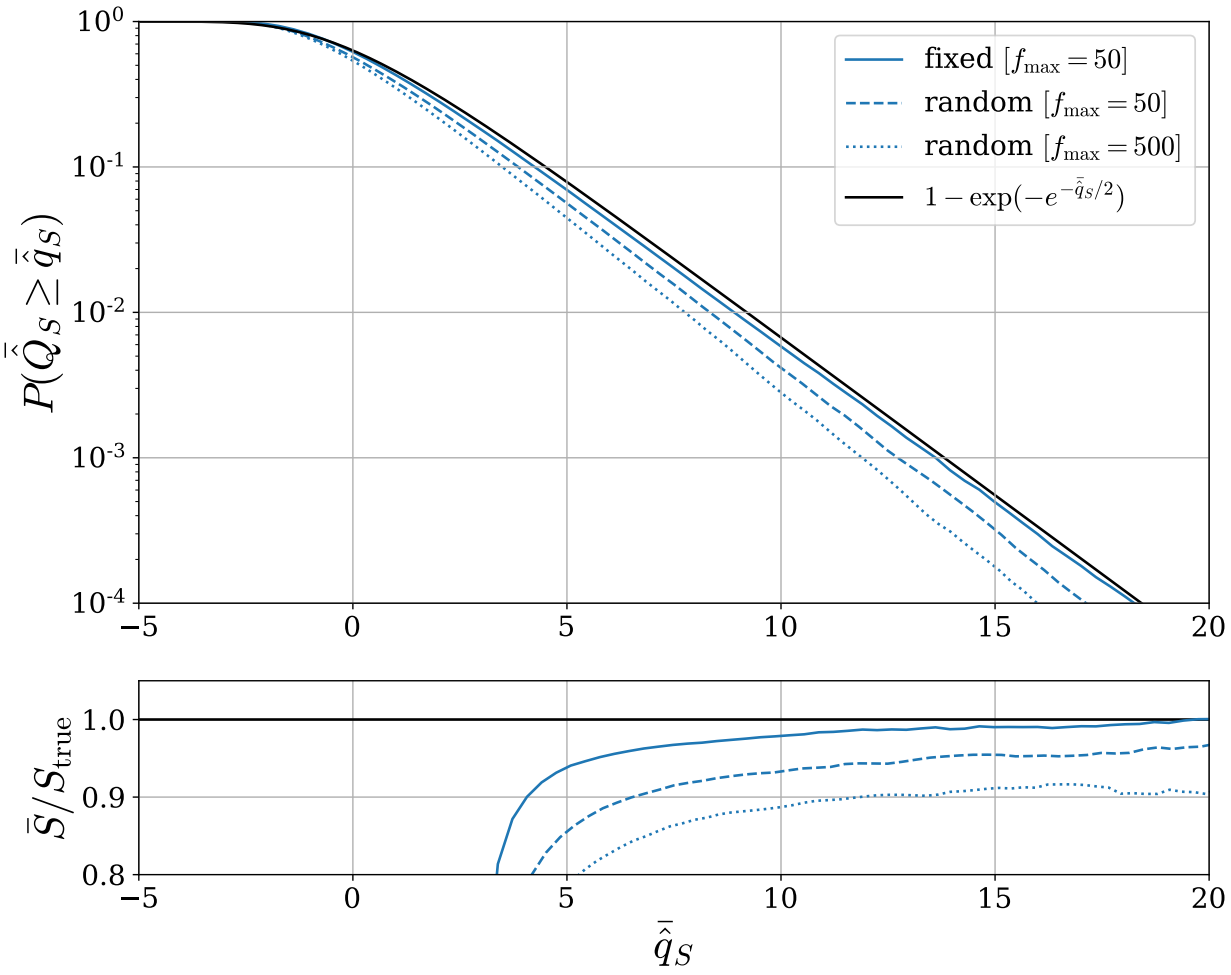

Figure 8.5: *Top panel*: Distribution of self-calibrated $\hat{q}_S$ for the LS periodogram with $f_{\rm Nyq} = 50$. We consider self-calibration using $n = 10$, but have checked results are robust to this choice. We plot lines for fixed data spacing with $f_{\max} = f_{\rm Nyq}$ (solid), random data spacing with $f_{\max} = f_{\rm Nyq}$ (dashed), and random data spacing with $f_{\max} = 10 f_{\rm Nyq}$ (dotted). The fixed-spacing case agrees well with Eq. 8.8, while the agreement slightly deteriorates for randomly spaced data, and further as $f_{\max}$ is increased beyond the Nyquist. This discrepancy is introduced by aliasing effects, however the *bottom panel* shows the asymptotic fractional error on the number of sigma, S, is at most 10%.

introduces additional peaks to the likelihood, which are correlated with each other. For this example the self-calibrated number of sigma S is still correct to within 10% in the asymptotic limit, as shown in the bottom panel of Fig. 8.5. Furthermore, since self-calibration underestimates the statistical significance it could lead to false negatives, but not false positives, so it is a conservative estimate. Given the difference between the local and global p-value is often many orders of magnitude, self-calibration assuming independent peaks still provides a useful fast way to approximate the p-value in cases of data without fixed-uniform spacing.

To illustrate why a joint analysis would be needed, consider fixed-uniform data spacing, for which the likelihood peaks in the range $(0, f_{\rm Nyq}]$ will repeat themselves in each $(if_{\rm Nyq}, (i+1)f_{\rm Nyq}]$, $i \in \mathbb{Z}$ region due to aliasing. This means that if one were to consider $f_{\rm max} > f_{\rm Nyq}$ there would be multiple repeated peaks of the same height and using self-calibration on individual peaks assuming they are uncorrelated would break down. One can instead do a joint fit of all the peaks. In this specific example this is equivalent to removing the signal associated with the maximum peak from the data before computing $q_L^{(2)}$, and then iteratively removing the signal of each peak to get to higher n. This will remove all the peaks caused by aliasing, and generalizes self-calibration to any frequency range in the case of fixed-uniform data spacing. While this example is not of physical interest, as there is no extra information beyond the Nyquist frequency for fixed-uniform data spacing, it motivates the solution to aliasing in cases of non-uniform data spacing. In our experiments we found this procedure of removing the peaks tends to remove too much signal from the higher order peaks when the peaks are correlated, so that self-calibration is not very accurate. For this reason we argue the correct procedure is to fit for all the peaks jointly, which is computationally expensive and beyond the scope of this work.

Kepler Exoplanet Search

Here we apply self-calibration in the context of exoplanet searches in the Kepler Space Telescope data [217]. The look-elsewhere effect is particularly prominent in these applications: in a pure-noise simulation a typical highest noise peak will have signal-to-noise ratio of 5.5, i.e. $q_L \approx 30$, corresponding to multiplicity, or trials factor, in excess of 10^7.

We have some time series data of fluxes $\{y_i\}_{i=0}^{N_d}$, measured at discrete time points t_i, which are evenly spaced $t_i = i\Delta$. The flux is composed of a signal and noise,

$$y_i = s(t_i) + \mathcal{N}_i. \tag{8.18}$$

We consider two noise scenarios: (i) when $\mathcal{N}_i$ is assumed to be a normally distributed random variable with zero mean and unit variance which is not correlated with other data points (i.e. white noise), and (ii) adding realistic Kepler 90 stellar variability. Here we consider a signal $s(t)$ comprised of a periodic train of $\mathcal{T}$ transits with period P. The signal has $M = 3$ parameters, $z = (A, P, \phi)$: amplitude, period and phase, respectively. The form of the signal is given by

$$s(t|A, P, \phi) = A \sum_{r=1}^{\mathcal{T}} U\left(\frac{t - (r + \phi)P}{\tau_K(P)}\right), \tag{8.19}$$

where $U(x)$ is a U-shaped transit template which is nonzero in the region (-1/2, 1/2). We use Kepler's third law to give the duration of each transit event as $\tau_K \propto P^{1/3}$, effectively assuming that planet's orbits are circular and perfectly aligned with the line of sight.

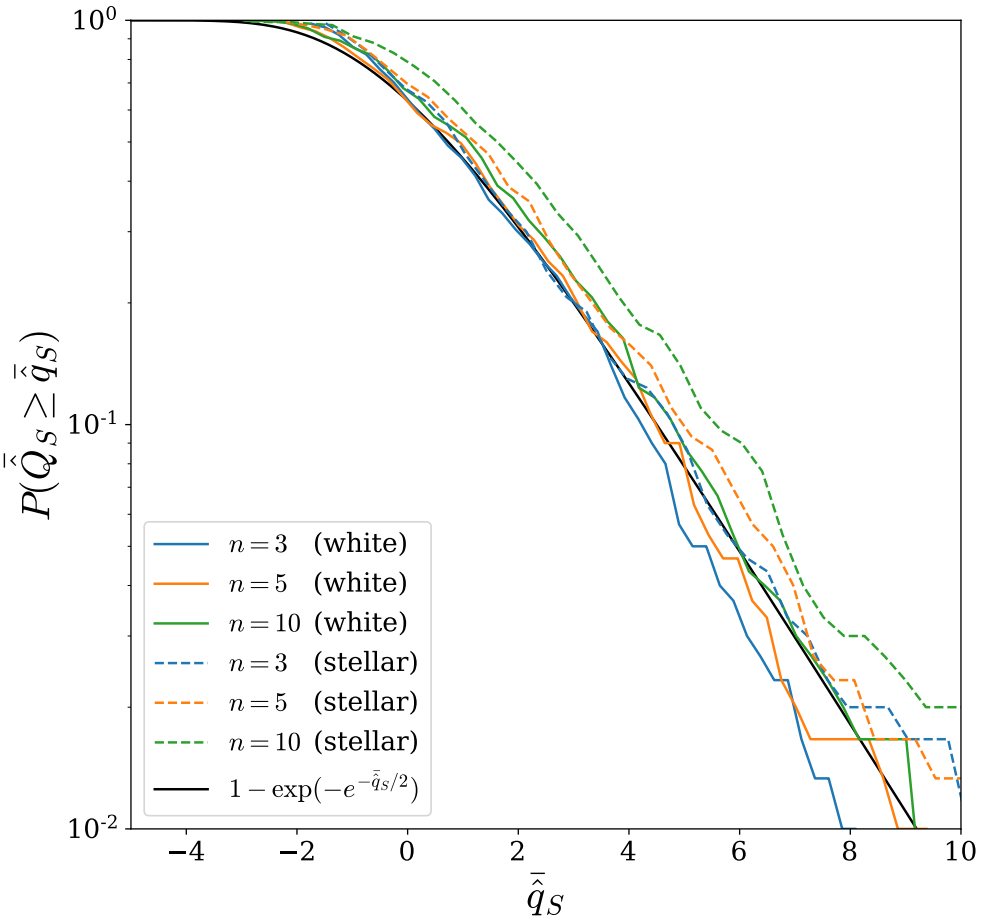

Figure 8.6: Distribution of self-calibrated $\hat{q}_S$ for exoplanet searches in Kepler-like data. We consider both white noise (solid) and realistic Kepler 90 stellar variability (dashed). We self-calibrate using $n = 3, 5, 10$, and compare with Eq. 8.8.

Following the analysis of [308, 307], matched filtering the data y with the template $s_0(t|z)$ gives the signal-to-noise, $\sqrt{q_L}$, as

$$q_L(\phi|P)^{1/2} = \mathcal{F}^{-1}\left\{\frac{\mathcal{F}\{y\}^*\ \mathcal{F}\{s_0\}}{\mathcal{P}}\right\}, \tag{8.20}$$

$$\text{SNR} = \mathcal{F}^{-1}\left\{\frac{\mathcal{F}\{d\}^*\ \mathcal{F}\{s\}}{\mathcal{P}}\right\}, \tag{8.21}$$

where $\mathcal{F}\{\cdot\}$ is the discrete Fourier transform and $\mathcal{P}$ is the noise power spectrum (which for white noise is a constant equal to the number of data points N_d). The template is normalized such that $\sum_{i=0}^{N_d} |\mathcal{F}\{s_0\}_i|^2/\mathcal{P}_i = 1$.

We consider a star like Kepler 90, where data spanning 1465.6 days of observations with $\Delta = 29.4$ minute intervals is available. To test self-calibration for such a model, we consider

300 noise-only simulations. We simulate time series, apply matched filtering, and search over periods in the range 3–300 days and over all phases.

Fig. 8.6 shows the distribution of the self-calibrated $\hat{q}_S$ over these noise realizations for a few choices of peak index, n. The solid lines correspond to the case of white noise, for which it can be seen that there is good agreement with Eq. 8.8, and thus self-calibration produces accurate results for all n considered.

Next we add realistic Kepler 90 stellar variability to the model, as shown by the dashed lines of Fig. 8.6. In this case the model decides whether to fit for the stellar variability or the exoplanet, or both. The null hypothesis is now noise and stellar variability, on top of which we are looking for signatures of exoplanets. We model stellar variability as a Gaussian process, measuring first its power spectrum from the data directly [307], and then fitting for all of the Fourier components of the stellar variability (approximately 70,000 components). It has been shown in [308] that the results of this joint fit are equivalent to the matched filter analysis, where we use the power spectrum for $\mathcal{P}$ in Eq. 8.20. We assume that the different planet peaks do not interact with each other. This makes such analysis feasible, unlike for the periodogram case where a joint fit of multiple peaks would be very expensive. Fig. 8.6 shows good agreement between self-calibration and simulations in the case of stellar variability, although there is a slight discrepancy for large n, so for an optimal analysis it suffices to use $n = 5$. For Kepler data we can scramble the data mixing up different time intervals, which destroys exoplanet periodicity, and guarantees that we have pure noise peaks in the scrambled data, so we do not need to worry about presence of real planets in the lower amplitude peaks.

Searching for oscillatory features in the primordial power spectrum

Here we apply self-calibration to real data, in the context of a search for oscillatory features in the primordial power spectrum. This is an example of a non-Gaussian model of cosmological inflation considered by [155, 156]. The model considered adds an oscillatory perturbation to the ΛCDM power spectrum as follows,

$$P(k) = P_0(k)[1 + A\sin(2\omega k + \phi)], \tag{8.22}$$

where $P_0(k)$ is the featureless (ΛCDM) power spectrum and A, ω, and ϕ are the amplitude, frequency, and phase of the oscillatory perturbation. This is thus an example of an $M = 3$ model. In this case one is uncertain what the frequency or phase of the oscillation is, and one scans over a large range of frequencies to seek a fit to theory – introducing a large look-elsewhere effect. Using the Planck 2013 likelihood [290], we follow the analysis detailed in §6 of [60] to produce the likelihood distribution on the left of Fig. 8.7. It can be seen to be highly multimodal with many spurious peaks. The maximum occurs as $\omega \approx 3660$ with $\hat{q}_L = 15.4$. Note there are three additional peaks with $\hat{q}_L \geq 10$, and several more between 8 and 10.

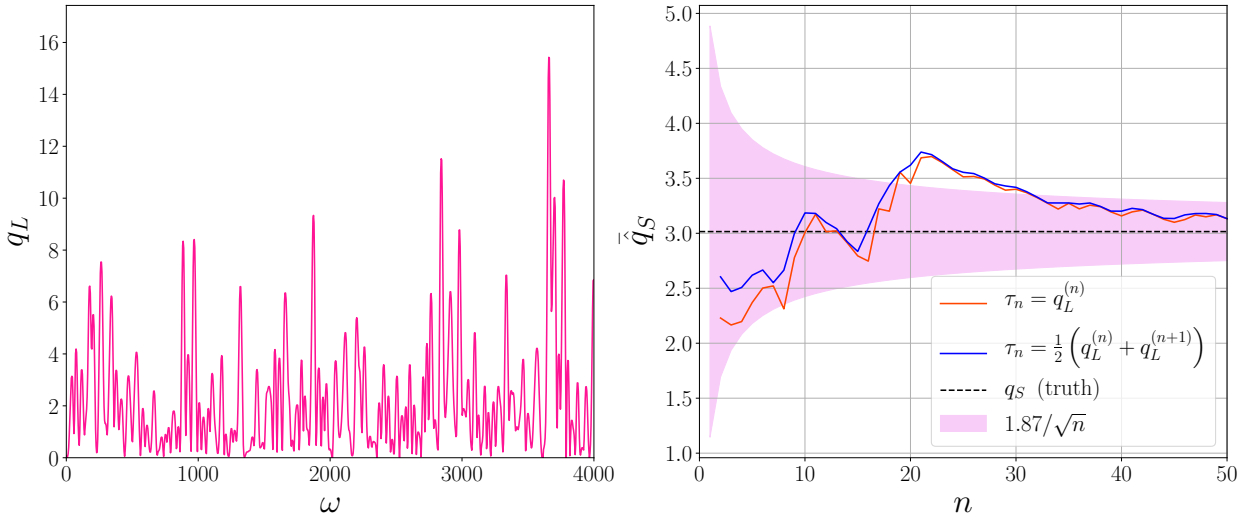

Figure 8.7: Example of self-calibration when searching the primordial power spectrum for oscillatory features. *Left panel*: Distribution of q_L projected onto the ω axis, i.e. having maximized over A and ϕ for each ω. The highest peak is at $\omega \approx 3660$, with $\hat{q}_L = 15.4$, giving an uncorrected significance of $\sqrt{15.4} \approx 4$ sigma. *Right panel*: The self-calibrated value of $\hat{q}_S$ for the highest peak as a function of n, with two options for τ_n from Eq. 8.41 and theoretical error envelope from Eq. 8.16. The true value of $\hat{q}_S$ from using the posterior volume is $\hat{q}_S = 3.0$, and self-calibration shows good agreement for all n. Using Eq. 8.8 with $\hat{q}_S = 3$ gives the p-value as 0.20, giving a significance of $S = 1.3$; this suggests that uncorrected 4-sigma peaks, such as this one, arise relatively commonly from noise fluctuations.

The true value of $\hat{q}_S$, obtained using the posterior volume, was found to be $\hat{q}_S = 3$. The right of Fig. 8.7 illustrates self-calibration, as detailed in Table 8.1, showing that a sufficiently accurate approximation of $\hat{q}_S$ is achieved for all choices of n. Furthermore, the pink shaded region represents the fitting formula for the error envelope presented in Eq. 8.16, and it can be seen that it encloses the data well. This suggests that self-calibration is reliable and in general this can be used as a diagnostic, as one typically will not have the true value of $\hat{q}_S$ to compare with. It is useful to note that even though Eq. 8.16 was obtained in the context of an $M = 2$ white noise example, it still provides a useful diagnostic for different models.

8.5 Conclusions

This paper presents a new method, *self-calibration*, to compute statistical significance in the presence of the look-elsewhere effect by considering only the heights of peaks in the data likelihood distribution. These peaks are a byproduct of any peak-hunting data analysis, so there is negligible computational cost in this approach. In contrast, existing methods rely on simulations, model-dependent analytical calculations, or explicit evaluation of the Bayes

factor, all of which can be time consuming.

In its simplest form, self-calibration subtracts the χ^2 of the highest noise peak, typically assumed to be the second peak, from the χ^2 of the highest peak, to approximate the look-elsewhere-corrected χ^2 of the highest peak. Accuracy is improved by considering lower peaks, typically at a negligible computational cost since they are also a byproduct of the analysis. This approach assumes that these lower peaks are dominated by noise; when this is not the case, i.e. there are multiple physical peaks, one can iteratively remove physical peaks from the data after verifying that their p-value is small. One can then also self-calibrate the significance of each physical peak. This is another reason to favor low amplitude peaks for the subtraction: one may not be certain if the highest peaks are physical or not, but one is often certain that the lower peaks are generated by noise. An alternative approach to noise-only peaks is to use some form of scrambled data where we know the signal has been eliminated. For example, in the exoplanet detections from transits this could either be an inverted or scrambled time series which eliminates the periodicity of the planet transits.

We showed that self-calibration gives an accurate estimate of the FAP, or p-value, of the highest peak(s) in various astrophysical examples, including planet detection, periodograms, and cosmology. We also developed a version of self-calibration which can be applied when the noise and likelihood are not known, where one must also self-calibrate the normalization. Our approach is general, but approximate: there are situations where fitting for individual peaks is inaccurate, and a joint fit accounting for the effects of correlated peaks is required. An example is periodogram analysis in the case of non-uniform data spacing. This is not conceptually any different in the sense that if peaks are correlated then a joint fit is required, but it is computationally difficult, and for specific situations such as periodograms methods have been developed where one can account for these effects without doing a joint fit [37]. However, we have demonstrated that even without correcting for these effects, self-calibration provides a simple method to quickly determine whether a significant detection has been made, and is thus complementary to the more specialized methods that apply to specific situations.

Acknowledgements

We are grateful to the anonymous referee for providing many useful comments that helped cultivate the paper into its final form. AEB thanks Omer Ronen for insightful discussion on the manuscript. This research made use of the Cori supercomputer at the National Energy Research Scientific Computing Center (NERSC), a U.S. Department of Energy Office of Science User Facility operated under Contract No. DE-AC02-05CH11231. We acknowledge the use of CAMB [239] and CosmoMC [238] for the analysis in Section 8.4. This material is based upon work supported by the National Science Foundation under Grant Numbers 1814370 and NSF 1839217, and by NASA under Grant Number 80NSSC18K1274.

8.6 Appendix A: Bayesian derivation of self-calibration

Here we provide a Bayesian derivation of the self-calibration equations presented in Section 8.3, based on the work of [60].

From a Bayesian perspective one considers the Bayes factor, which considers the evidence ratio for the hypothesis that there is a signal H to the null hypothesis H_0. The Bayes factor is the ratio of Bayesian evidence under each hypothesis:

$$p(\boldsymbol{x}|H) = \int d\boldsymbol{z}\; p(\boldsymbol{z}|H)p(\boldsymbol{x}|\boldsymbol{z}, H), \tag{8.23}$$

where $p(\boldsymbol{z}|H)$ is the prior under hypothesis H and $p(\boldsymbol{x}|\boldsymbol{z}, H)$ is the likelihood under hypothesis H. Thus the Bayesian evidence is equal to the prior-weighted average of the likelihood. For the examples considered in this paper we consider a null hypothesis for which the parameters are fixed, thus the evidence for H_0 is simply given by the null likelihood $p(\boldsymbol{x}|\boldsymbol{z}_0)$.

For a multimodal likelihood, the Bayes factor can be approximated by performing a local integration at each peak. This leads to a sum over contributions from each posterior mode. If the location of the ℓ^{th} highest mode is $\boldsymbol{z} = \boldsymbol{\mu}^\ell$, this gives Bayes factor as

$$B \equiv \frac{p(\boldsymbol{x}|H)}{p(\boldsymbol{x}|H_0)} \approx \sum_\ell b^\ell, \tag{8.24}$$

where each b^ℓ is the contribution of mode ℓ to the Bayes factor and can be parameterized as

$$b^\ell = \frac{p(\boldsymbol{x}|\boldsymbol{\mu}^\ell)}{p(\boldsymbol{x}|\boldsymbol{z}_0)} \frac{V^\ell_{\text{posterior}}}{V^\ell_{\text{prior}}} = e^{q_L^\ell/2} \frac{V^\ell_{\text{posterior}}}{V^\ell_{\text{prior}}}, \tag{8.25}$$

where V^ℓ_{prior} ($V^\ell_{\text{posterior}}$) is the prior (posterior) volume associate with mode ℓ. A common approximation for these volumes is the Laplace approximation [230], in which case the prior volume is given by

$$V^\ell_{\text{prior}} \simeq 1/p(\boldsymbol{\mu}^\ell), \tag{8.26}$$

and the posterior volume is given by

$$V^\ell_{\text{posterior}} \simeq (2\pi)^{M/2} \sqrt{\det \boldsymbol{\Sigma}^\ell}, \tag{8.27}$$

where $\boldsymbol{\Sigma}$ is the covariance matrix. We note that in principle one can compute Bayes factor exactly, and the Laplace approximation is only a simple, but often effective, approximation [60].

One can combine the frequentist and Bayesian perspectives to define a new test statistic

$$q_S \equiv q_L - 2\ln N + \ln 2\pi q_L - 2\ln t, \tag{8.28}$$

where the trials factor is taken as the prior-to-posterior volume ratio for the parameters $\boldsymbol{z}_{>1}$ at the maximum peak,

$$N = \frac{\hat{V}_{>1,\text{prior}}}{\hat{V}_{>1,\text{posterior}}}, \tag{8.29}$$

and $t = 1, 2$ for a one, two-tailed test. Note that while this is the prior-to-posterior volume ratio for $\boldsymbol{z}_{>1}$, we will often simply refer to it as the prior-to-posterior volume. The p-value is then given by

$$P(\hat{Q}_S > \hat{q}_S) \simeq 1 - \exp\left(-e^{-\hat{q}_S/2}\right), \tag{8.30}$$

and applies both asymptotically and non-asymptotically. The key difference between the p-value expressions of equations 8.2 and 8.30 is that the latter has no explicit N dependence, meaning the p-value in terms of $\hat{q}_S$ is unaffected by the look-elsewhere effect. This makes $\hat{q}_S$ a more useful statistic to use. The statistical significance, or the number of sigma, can be approximated as

$$S \approx \sqrt{\hat{q}_S - \ln 2\pi\hat{q}_S + 2\ln t}, \tag{8.31}$$

with corrections of order $\mathcal{O}(\hat{q}_S^{-3/2})$. The look-elsewhere-corrected chi-squared is S^2. Note for sufficiently large $\hat{q}_S$, $S \approx \sqrt{\hat{q}_S}$.

Thus all one needs to evaluate the p-value is $\hat{q}_S$, which itself depends on $\hat{q}_L$ and N. Computing N requires the evaluation of the posterior volume over $\boldsymbol{z}_{>1}$, which can be evaluated using the Laplace approximation, Variational Inference or Monte Carlo Markov Chain methods. However, we seek a faster alternative.

The asymptotic scaling of the p-value with N, and thus the prior volume, in Eq. 8.30 offers a way to evaluate the trials factor from the distribution of q_L across the peaks in a dataset. One can evaluate the p-value in subvolumes of the data by counting the number of peaks above some threshold, and then rescale this to give the p-value for the entire volume. Calibrating N in this way is cheaper than evaluating the posterior volume directly, for example by using Monte Carlo methods, because peaks in the likelihood are a byproduct of the peak-search analysis. Moreover, provided the q_L peaks are dominated by noise, one can perform this calibration on the data directly without needing to run simulations. We thus call this method self-calibration, as one is calibrating N using the peaks belonging to the measured data itself.

We start by splitting the prior volume of $\boldsymbol{z}_{>1}$ into K bins, such that the prior volume of one of the bins is

$$V'_{>1,\text{prior}} = V_{>1,\text{prior}}/K. \tag{8.32}$$

We note that for non-uniform priors, one may need to include correction terms in this splitting of the prior volume to obtain greater accuracy (see e.g. [44] for discussion of splitting complex prior volumes).

Thinking of the p-value in terms of the false-positive rate (FPR), one can approximate the p-value of a bin as the fraction of bins containing a peak with $q_L > \tau$, for some threshold τ. Smaller peaks typically have a larger error on $\boldsymbol{z}_{>1}$, which scales approximately as the noise-to-signal, i.e. $\hat{q}_L^{-(M-1)/2}$. Thus the average posterior volume in one of the bins is related to

the posterior volume of the full volume by

$$V'_{>1,\text{posteriror}} \simeq (\hat{q}_L/\tau)^{(M-1)/2} V_{>1,\text{posteriror}}. \tag{8.33}$$

Substituting equations 8.32 and 8.33 into Eq. 8.29 gives the trials factor for a single bin as

$$N' \simeq \frac{N}{K}\left(\frac{\tau}{\hat{q}_L}\right)^{(M-1)/2}. \tag{8.34}$$

Denoting the number of bins containing at least one peak with $q_L > \tau$ as $n_{\text{bins}}(\tau)$, the FPR is given by the fraction of bins satisfying this condition. Equating the FPR to the p-value from Eq. 8.30 gives

$$\frac{n_{\text{bins}}(\tau)}{K} = 1 - \exp\left(-e^{-\frac{1}{2}[\tau - 2\ln N' + \ln 2\pi\tau - 2\ln t]}\right) \tag{8.35}$$

$$= 1 - \exp\left(-e^{-\frac{1}{2}\left[\tau - 2\ln N + 2\ln K + \ln 2\pi\hat{q}_L + (M-2)\ln\frac{\hat{q}_L}{\tau} - 2\ln t\right]}\right). \tag{8.36}$$

Rearranging and taking the $K \to \infty$ limit gives an expression for the trials factor

$$2\ln N \to \tau + 2\ln n_{\text{peaks}}(\tau) + \ln 2\pi\hat{q}_L + (M-2)\ln\frac{\hat{q}_L}{\tau} - 2\ln t, \tag{8.37}$$

where the number of bins with a q_L peak larger than τ tends to the number of peaks with q_L larger than τ in the full volume, i.e. $n_{\text{bins}}(\tau) \to n_{\text{peaks}}(\tau)$ as $K \to \infty$. This provides an estimate of N that relies purely on the number and height of the peaks, without needing to evaluate their posterior volumes.[3] Combining with Eq. 8.28 this gives $\hat{q}_S$ as

$$\hat{q}_S = \hat{q}_L - \tau - 2\ln n_{\text{peaks}}(\tau) - (M-2)\ln\frac{\hat{q}_L}{\tau}, \tag{8.38}$$

and Eq. 8.30 gives the p-value as

$$P(\hat{Q}_L > \hat{q}_L) = 1 - \exp\left(-e^{-\frac{1}{2}\left[\hat{q}_L - \tau - 2\ln n_{\text{peaks}}(\tau) - (M-2)\ln\frac{\hat{q}_L}{\tau}\right]}\right). \tag{8.39}$$

In the above formulae, τ is a continuous variable, however a single dataset consists of a discrete set of peaks. To self-calibrate the p-value of a particular dataset we thus evaluate the above formulae for the n^{th} peak such that

$$n_{\text{peaks}}(\tau_n) = n, \tag{8.40}$$

with

$$\tau = \tau_n \equiv \frac{q_L^{(n)} + q_L^{(n+1)}}{2} \approx q_L^{(n)}, \tag{8.41}$$

[3]Note we have ignored boundary effects (see e.g. [43]), however these will be negligible in the case of large N, i.e. when the look-elsewhere effect is considerable.

where $q_L^{(n)}$ is the q_L value of the n^{th} highest peak. We choose τ as the average of $q_L^{(n)}$ and $q_L^{(n+1)}$ as a simple way to account for the discreteness of the data: for example, if $q_L^{(n)} = 10$ and $q_L^{(n+1)} = 6$, one can only conclude that n corresponds to τ in the range $6 < \tau \leq 10$. Having said that, in many cases one can choose $\tau_n = q_L^{(n)}$ and achieve sufficient accuracy. We explore both options of τ_n in the main paper.

Substituting equations 8.40 and 8.41 into equations 8.37, 8.38, and 8.39, gives the self-calibrated estimate of $2 \ln N$ as

$$2 \ln \bar{N} \equiv \tau_n + 2 \ln n + \ln 2\pi \hat{q}_L + (M-2) \ln \frac{\hat{q}_L}{\tau_n} - 2 \ln t, \tag{8.42}$$

the self-calibrated estimate of $\hat{q}_S$ as

$$\bar{\hat{q}}_S \equiv \hat{q}_L - \tau_n - 2 \ln n - (M-2) \ln \frac{\hat{q}_L}{\tau_n}, \tag{8.43}$$

and the self-calibrated estimate of the p-value as

$$\bar{P}(\hat{Q}_L > \hat{q}_L) \equiv 1 - \exp\left(-e^{-\frac{1}{2}\left[\hat{q}_L - \tau_n - 2\ln n - (M-2)\ln\frac{\hat{q}_L}{\tau_n}\right]}\right), \tag{8.44}$$

where bars are used to indicate these expressions are the self-calibrated approximations. Finally, Eq. 8.31 can be used to compute the self-calibrated significance as

$$\bar{S} = \sqrt{\bar{\hat{q}}_S - \ln 2\pi \bar{\hat{q}}_S + 2 \ln t}, \tag{8.45}$$

for large $\bar{\hat{q}}_S$. Note for sufficiently large $\bar{\hat{q}}_S$, $\bar{S} \approx \sqrt{\bar{\hat{q}}_S}$. Eqs. 8.42, 8.43, 8.44, and 8.45 correspond to Eqs. 8.11, 8.9, 8.7, and 8.10, quoted in the main text.

Chapter 9

Conclusions

High-precision measurements of large-scale cosmic structure are expected to revolutionize our understanding of fundamental physics, for example by the quantifying neutrino mass and elucidating the nature of dark energy. To fully realize the potential of these measurements, this dissertation has tackled various of the theoretical and numerical challenges that must be addressed.

We started in Chapter 1 by introducing the standard model of cosmology. We motivated that measurements of cosmic structure on small, nonlinear, scales can provide a wealth of cosmological information. In particular, small scale information can help measure the neutrino mass as massive neutrino suppress the growth of structure on these scales. To analyse the information regrading neutrino mass in cosmic structure, we modelled massive neutrinos by incorporating them into a particle-mesh simulation, FastPM, in Chapter 2. By applying an innovative method we were able to bypass the problem of shot noise that typically plagues neutrino simulations, enabling modeling the effects of massive neutrinos at a more reasonable computational cost than traditional methods.

In Chapter 3 we then used simulations to investigate the amount of information in the 3d matter field, δ_m. In particular we considered the power spectrum, halo mass function, and void size function. We found that these three statistics show very different degeneracies in the M_ν–σ_8 plane, meaning that tight constraints are achieved when combining all three. For the particular setup considered, we find that one can get constraints of 0.018eV using scales up to $k_{\rm max} = 0.5\,h/{\rm Mpc}$.

However, the 3d matter field is not observable in modern cosmological surveys. In Chapter 4 we then explored how much information regarding neutrino mass there will be in the observable galaxy and weak lensing fields. Additionally, instead of looking at specific types of higher-order statistics, we considered the total information in the field by looking at the complex phases. By fixing the linear physics between two N-body simulations, we found that there is actually very little information regarding neutrino mass in the galaxy and weak lensing fields up to scales of $k = 1\,h/{\rm Mpc}$. We do however find that there is some signal from velocities, which trace the matter field via $\delta_v \sim f\delta_m$. This motivates using peculiar velocity information, together with accurate modelling of redshift-space distortions, to opti-

mize constraints on neutrino mass. Moreover, combining large-scale structure with cosmic microwave background measurements can help break degeneracies, in turn unlocking further information.

To move towards obtaining optimal constraints, in Chapter 5 we then employ Bayesian forward modeling with field-level inference to extract all the information in cosmic fields. We developed a framework to perform a combined analysis of galaxy clustering information with galaxy peculiar velocity information to reconstruct the initial conditions of the Universe. We achieve this using differentiable forward modeling. In the future, this approach can be extended to obtain constraints on cosmological parameters, including for neutrino mass, and can be combined with analysis from weak lensing and the CMB.

In Chapter 6 we then focus in on a particular challenge in doing inference for cosmological analyses, the super-sample covariance effect. The is a nonlinear effect on the covariance matrix which arises due to neglect of modes that are larger than the size of the simulation. The effect of these missing large modes can cause a significant difference in the covariance matrix on small nonlinear scales. We created a set of FastPM simulations which we used to quantify this effect for the power spectrum, bispectrum, halo mass function, and void size function. These simulations can be used to quantify the effect for other statistics in the future.

Finally, performing Bayesian inference in high dimensions, or for nonlinear models, can result in non-trivial volume effects which make it difficult to reliably perform inference. One such problem is known as the look-elsewhere effect which occurs when scanning a large parameter space for evidence of a new model or a signal. In such cases it becomes difficult to quantify the statistical significance of a discovery, moreover, the posterior might be multimodal making it difficult to determine which peak is the most significant. In Chapter 7 we analyse the look-elsewhere effect from a frequentist and Bayesian perspective, ultimately marrying the two to produce a unified method that quantifies the effect orders of magnitude faster than traditional methods. Then in Chapter 8 we presented an even faster method of *self-calibrating* the effect by simply computing the significance from the heights and ranks of peaks in the likelihood. This approach has negligible computational cost as peaks in the likelihood are a byproduct of every peak-search analysis.

All put together, this dissertation has provided various key ingredients to enable the cosmology community to move closer to an optimal detection of neutrino mass and beyond. Much work lies ahead, both in terms of creating accurate forward models which include effects such as redshift-space distortions, and in terms of developing numerical tools to make high-dimensional inference more tractable, such as more efficient samplers and optimizers. Over the coming decades, by combining accurate forward modeling of large-scale structure on nonlinear scales with innovative methods in Bayesian inference, we can hope to uncover a wealth of information about our Universe.

Bibliography

[1] Georges Aad et al. "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC". In: *Phys. Lett.* B716 (2012), pp. 1–29. DOI: 10.1016/j.physletb.2012.08.020. arXiv: 1207.7214 [hep-ex].

[2] M. G. Aartsen et al. "Observation of High-Energy Astrophysical Neutrinos in Three Years of IceCube Data". In: *Physical Review Letters* 113.10 (Sept. 2014). ISSN: 1079-7114. DOI: 10.1103/physrevlett.113.101101. URL: http://dx.doi.org/10.1103/PhysRevLett.113.101101.

[3] K. N. Abazajian et al. "CMB-S4 Science Book, First Edition". In: *ArXiv e-prints* (Oct. 2016). arXiv: 1610.02743.

[4] B. P. Abbott et al. "GW150914: First results from the search for binary black hole coalescence with Advanced LIGO". In: *Physical Review D* 93.12 (June 2016). ISSN: 2470-0029. DOI: 10.1103/physrevd.93.122003. URL: http://dx.doi.org/10.1103/PhysRevD.93.122003.

[5] T. M. C. Abbott et al. "Dark Energy Survey year 1 results: Cosmological constraints from galaxy clustering and weak lensing". In: *Phys. Rev. D* 98 (4 Aug. 2018), p. 043526. DOI: 10.1103/PhysRevD.98.043526. URL: https://link.aps.org/doi/10.1103/PhysRevD.98.043526.

[6] Julian Adamek, Ruth Durrer, and Martin Kunz. "Relativistic N-body simulations with massive neutrinos". In: *Journal of Cosmology and Astroparticle Physics* 2017.11 (Nov. 2017), pp. 004–004. ISSN: 1475-7516. DOI: 10.1088/1475-7516/2017/11/004. URL: http://dx.doi.org/10.1088/1475-7516/2017/11/004.

[7] R. J. Adler. *The Geometry of Random Fields*. Springer-Verlag New York, 1981.

[8] N. Aghanim et al. "Planck 2015 results. XI. CMB power spectra, likelihoods, and robustness of parameters". In: *Astron. Astrophys.* 594 (2016), A11. DOI: 10.1051/0004-6361/201526926. arXiv: 1507.02704 [astro-ph.CO].

[9] Q. R. Ahmad et al. "Direct evidence for neutrino flavor transformation from neutral current interactions in the Sudbury Neutrino Observatory". In: *Phys. Rev. Lett.* 89 (2002), p. 011301. DOI: 10.1103/PhysRevLett.89.011301. arXiv: nucl-ex/0204008 [nucl-ex].

[10] M. H. Ahn et al. "Measurement of neutrino oscillation by the K2K experiment". In: *Phys. Rev. D* 74 (7 Oct. 2006), p. 072003. DOI: 10.1103/PhysRevD.74.072003. URL: https://link.aps.org/doi/10.1103/PhysRevD.74.072003.

[11] Virginia Ajani et al. *Constraining neutrino masses with weak-lensing starlet peak counts*. 2020. arXiv: 2001.10993 [astro-ph.CO].

[12] M. Aker et al. "Improved Upper Limit on the Neutrino Mass from a Direct Kinematic Method by KATRIN". In: *Physical Review Letters* 123.22 (Nov. 2019). ISSN: 1079-7114. DOI: 10.1103/physrevlett.123.221802. URL: http://dx.doi.org/10.1103/PhysRevLett.123.221802.

[13] Kazuyuki Akitsu, Yin Li, and Teppei Okumura. "Cosmological simulation in tides: power spectra, halo shape responses, and shape assembly bias". In: *Journal of Cosmology and Astroparticle Physics* 2021.4, 041 (Apr. 2021), p. 041. DOI: 10.1088/1475-7516/2021/04/041. arXiv: 2011.06584 [astro-ph.CO].

[14] Kazuyuki Akitsu and Masahiro Takada. "Impact of large-scale tides on cosmological distortions via redshift-space power spectrum". In: *Phys. Rev. D* 97.6, 063527 (Mar. 2018), p. 063527. DOI: 10.1103/PhysRevD.97.063527. arXiv: 1711.00012 [astro-ph.CO].

[15] Kazuyuki Akitsu, Masahiro Takada, and Yin Li. "Large-scale tidal effect on redshift-space power spectrum in a finite-volume survey". In: *Phys. Rev. D* 95.8, 083522 (Apr. 2017), p. 083522. DOI: 10.1103/PhysRevD.95.083522. arXiv: 1611.04723 [astro-ph.CO].

[16] Shadab Alam et al. "The clustering of galaxies in the completed SDSS-III Baryon Oscillation Spectroscopic Survey: cosmological analysis of the DR12 galaxy sample". In: *Monthly Notices of the Royal Astronomical Society* 470.3 (Mar. 2017), pp. 2617–2652. ISSN: 1365-2966. DOI: 10.1093/mnras/stx721. URL: http://dx.doi.org/10.1093/mnras/stx721.

[17] S. Algeri et al. "On methods for correcting for the look-elsewhere effect in searches for new physics". In: *Journal of Instrumentation* 11.12 (Dec. 2016), P12010–P12010. ISSN: 1748-0221. DOI: 10.1088/1748-0221/11/12/p12010. URL: http://dx.doi.org/10.1088/1748-0221/11/12/P12010.

[18] Yacine Ali-Haïmoud and Simeon Bird. "An efficient implementation of massive neutrinos in non-linear structure formation simulations". In: *Monthly Notices of the Royal Astronomical Society* 428.4 (Nov. 2012), pp. 3375–3389. ISSN: 1365-2966. DOI: 10.1093/mnras/sts286. URL: http://dx.doi.org/10.1093/mnras/sts286.

[19] R. Allison et al. "Towards a cosmological neutrino mass detection". In: *Phys. Rev. D* 92 (12 Dec. 2015), p. 123535. DOI: 10.1103/PhysRevD.92.123535. URL: https://link.aps.org/doi/10.1103/PhysRevD.92.123535.

[20] E. Allys et al. "New interpretable statistics for large-scale structure analysis and generation". In: *Phys. Rev. D* 102.10, 103506 (Nov. 2020), p. 103506. DOI: 10.1103/PhysRevD.102.103506. arXiv: 2006.06298 [astro-ph.CO].

[21] Luca Amendola et al. "Cosmology and fundamental physics with the Euclid satellite". In: *Living Reviews in Relativity* 21.1 (2018), p. 2. DOI: 10.1007/s41114-017-0010-3. URL: https://doi.org/10.1007/s41114-017-0010-3.

[22] F. P. An et al. "Observation of Electron-Antineutrino Disappearance at Daya Bay". In: *Phys. Rev. Lett.* 108 (17 Apr. 2012), p. 171803. DOI: 10.1103/PhysRevLett.108.171803. URL: https://link.aps.org/doi/10.1103/PhysRevLett.108.171803.

[23] B. Anderson et al. "Search for gamma-ray lines towards galaxy clusters with the Fermi-LAT". In: *Journal of Cosmology and Astroparticle Physics* 2016.02 (Feb. 2016), pp. 026–026. ISSN: 1475-7516. DOI: 10.1088/1475-7516/2016/02/026. URL: http://dx.doi.org/10.1088/1475-7516/2016/02/026.

[24] Neha Anil Kumar et al. "Primordial trispectrum from kinetic Sunyaev-Zel'dovich tomography". In: *Phys. Rev. D* 106.6 (2022), p. 063533. DOI: 10.1103/PhysRevD.106.063533. arXiv: 2205.03423 [astro-ph.CO].

[25] T. Araki et al. "Measurement of neutrino oscillation with KamLAND: Evidence of spectral distortion". In: *Phys. Rev. Lett.* 94 (2005), p. 081801. DOI: 10.1103/PhysRevLett.94.081801. arXiv: hep-ex/0406035 [hep-ex].

[26] Maria Archidiacono and Steen Hannestad. "Efficient calculation of cosmological neutrino clustering in the non-linear regime". In: *Journal of Cosmology and Astroparticle Physics* 2016.06 (June 2016), pp. 018–018. ISSN: 1475-7516. DOI: 10.1088/1475-7516/2016/06/018. URL: http://dx.doi.org/10.1088/1475-7516/2016/06/018.

[27] Maria Archidiacono, Steen Hannestad, and Julien Lesgourgues. *What will it take to measure individual neutrino mass states using cosmology?* 2020. arXiv: 2003.03354 [astro-ph.CO].

[28] Maria Archidiacono et al. "Physical effects involved in the measurements of neutrino masses with future cosmological data". In: *Journal of Cosmology and Astroparticle Physics* 2017.02 (Feb. 2017), pp. 052–052. DOI: 10.1088/1475-7516/2017/02/052. URL: https://doi.org/10.1088/1475-7516/2017/02/052.

[29] Metin Ata et al. "Predicted future fate of COSMOS galaxy protoclusters over 11 Gyr with constrained simulations". In: *Nature Astronomy* 6 (June 2022), pp. 857–865. DOI: 10.1038/s41550-022-01693-0. arXiv: 2206.01115 [astro-ph.CO].

[30] Alejandro Aviles and Arka Banerjee. "A Lagrangian perturbation theory in the presence of massive neutrinos". In: *Journal of Cosmology and Astroparticle Physics* 2020.10 (Oct. 2020), pp. 034–034. ISSN: 1475-7516. DOI: 10.1088/1475-7516/2020/10/034. URL: http://dx.doi.org/10.1088/1475-7516/2020/10/034.

[31] JEAN-MARC Azaïs and CÉLINE Delmas. "Asymptotic Expansions for the Distribution of the Maximum of Gaussian Random Fields". In: *Extremes* 5.2 (2002), pp. 181–212. DOI: 10.1023/A:1022123321967. URL: https://doi.org/10.1023/A:1022123321967.

[32] Paul I W de Bakker et al. "Efficiency and power in genetic association studies". In: *Nature Genetics* 37.11 (2005), pp. 1217–1223. DOI: 10.1038/ng1669. URL: https://doi.org/10.1038/ng1669.

[33] Tobias Baldauf, Vincent Desjacques, and Uroš Seljak. "Velocity bias in the distribution of dark matter halos". In: *Phys. Rev. D* 92 (2015), p. 123507. DOI: 10.1103/PhysRevD.92.123507. arXiv: 1405.5885 [astro-ph.CO].

[34] Tobias Baldauf et al. "Galaxy bias and non-linear structure formation in general relativity". In: *Journal of Cosmology and Astroparticle Physics* 2011.10, 031 (Oct. 2011), p. 031. DOI: 10.1088/1475-7516/2011/10/031. arXiv: 1106.5507 [astro-ph.CO].

[35] Tobias Baldauf et al. "Halo stochasticity from exclusion and nonlinear clustering". In: *Phys. Rev. D* 88.8 (2013), p. 083507. DOI: 10.1103/PhysRevD.88.083507. arXiv: 1305.2917 [astro-ph.CO].

[36] Tobias Baldauf et al. "Linear response to long wavelength fluctuations using curvature simulations". In: *Journal of Cosmology and Astroparticle Physics* 2016.9, 007 (Sept. 2016), p. 007. DOI: 10.1088/1475-7516/2016/09/007. arXiv: 1511.01465 [astro-ph.CO].

[37] R. V. Baluev. "Assessing the statistical significance of periodogram peaks". In: *Monthly Notices of the Royal Astronomical Society* 385.3 (Apr. 2008), pp. 1279–1285. ISSN: 1365-2966. DOI: 10.1111/j.1365-2966.2008.12689.x. URL: http://dx.doi.org/10.1111/j.1365-2966.2008.12689.x.

[38] R. V. Baluev. "Statistical detection of patterns in unidimensional distributions by continuous wavelet transforms". In: *Astronomy and Computing* 23, 151 (Apr. 2018), p. 151. DOI: 10.1016/j.ascom.2018.04.004. arXiv: 1711.07820 [astro-ph.IM].

[39] R. V. Baluev and E. I. Rodionov. "Analysing the Main Belt asteroid distributions by wavelets". In: *Celestial Mechanics and Dynamical Astronomy* 132.6-7, 34 (July 2020), p. 34. DOI: 10.1007/s10569-020-09976-2. arXiv: 2001.02990 [astro-ph.EP].

[40] R. V. Baluev, V. Sh. Shaidulin, and A. V. Veselova. "High-Velocity Moving Groups in the Solar Neighborhood in Gaia DR2". In: *Acta Astronomica* 70.2 (June 2020), pp. 141–168. DOI: 10.32023/0001-5237/70.2.4. arXiv: 2011.09812 [astro-ph.SR].

[41] R. V. Baluev et al. "Massive Search for Spot- and Facula-Crossing Events in 1598 Exoplanetary Transit Light Curves". In: *Acta Astronomica* 71.1 (Mar. 2021), pp. 25–53. DOI: 10.32023/0001-5237/71.1.2. arXiv: 2105.01704 [astro-ph.EP].

[42] Roman V. Baluev. "Detecting non-sinusoidal periodicities in observational data using multiharmonic periodograms". In: *Monthly Notices of the Royal Astronomical Society* 395.3 (May 2009), pp. 1541–1548. ISSN: 1365-2966. DOI: 10.1111/j.1365-2966.2009.14634.x. URL: http://dx.doi.org/10.1111/j.1365-2966.2009.14634.x.

[43] Roman V. Baluev. "Detecting non-sinusoidal periodicities in observational data: the von Mises periodogram for variable stars and exoplanetary transits". In: *Monthly Notices of the Royal Astronomical Society* 431.2 (Mar. 2013), pp. 1167–1179. ISSN: 0035-8711. DOI: 10.1093/mnras/stt238. eprint: https://academic.oup.com/mnras/article-pdf/431/2/1167/13761397/stt238.pdf. URL: https://doi.org/10.1093/mnras/stt238.

[44] Roman V. Baluev. "Keplerian periodogram for Doppler exoplanet detection: optimized computation and analytic significance thresholds". In: *Monthly Notices of the Royal Astronomical Society* 446.2 (Nov. 2014), pp. 1478–1492. ISSN: 0035-8711. DOI: 10.1093/mnras/stu2191. eprint: https://academic.oup.com/mnras/article-pdf/446/2/1478/13767828/stu2191.pdf. URL: https://doi.org/10.1093/mnras/stu2191.

[45] Roman V. Baluev and Vakhit S. Shaidulin. "Fine-resolution analysis of exoplanetary distributions by wavelets: hints of an overshooting iceline accumulation". In: *Astrophys. and Space Science* 363.9, 192 (Sept. 2018), p. 192. DOI: 10.1007/s10509-018-3416-9. arXiv: 1712.06374 [astro-ph.EP].

[46] Arka Banerjee and Tom Abel. "Nearest neighbour distributions: New statistical measures for cosmological clustering". In: *Monthly Notices of the Royal Astronomical Society* 500.4 (Jan. 2021), pp. 5479–5499. DOI: 10.1093/mnras/staa3604. arXiv: 2007.13342 [astro-ph.CO].

[47] Arka Banerjee and Neal Dalal. "Simulating nonlinear cosmological structure formation with massive neutrinos". In: *JCAP* 11 (2016), p. 015. DOI: 10.1088/1475-7516/2016/11/015. arXiv: 1606.06167 [astro-ph.CO].

[48] Arka Banerjee et al. "Reducing noise in cosmological N-body simulations with neutrinos". In: *Journal of Cosmology and Astroparticle Physics* 2018.09 (Sept. 2018), pp. 028–028. ISSN: 1475-7516. DOI: 10.1088/1475-7516/2018/09/028. URL: http://dx.doi.org/10.1088/1475-7516/2018/09/028.

[49] Josh Barnes and Piet Hut. "A hierarchical O(N log N) force-calculation algorithm". In: *Nature* 324.6096 (Dec. 1986), pp. 446–449. DOI: 10.1038/324446a0.

[50] Alexandre Barreira. "The squeezed matter bispectrum covariance with responses". In: *Journal of Cosmology and Astroparticle Physics* 2019.3, 008 (Mar. 2019), p. 008. DOI: 10.1088/1475-7516/2019/03/008. arXiv: 1901.01243 [astro-ph.CO].

[51] Alexandre Barreira, Elisabeth Krause, and Fabian Schmidt. "Accurate cosmic shear errors: do we need ensembles of simulations?" In: *JCAP* 10 (2018), p. 053. DOI: 10.1088/1475-7516/2018/10/053. arXiv: 1807.04266 [astro-ph.CO].

[52] Alexandre Barreira, Elisabeth Krause, and Fabian Schmidt. "Complete super-sample lensing covariance in the response approach". In: *Journal of Cosmology and Astroparticle Physics* 2018.6, 015 (June 2018), p. 015. DOI: 10.1088/1475-7516/2018/06/015. arXiv: 1711.07467 [astro-ph.CO].

[53] Alexandre Barreira and Fabian Schmidt. "Response approach to the matter power spectrum covariance". In: *Journal of Cosmology and Astroparticle Physics* 2017.11, 051 (Nov. 2017), p. 051. DOI: 10.1088/1475-7516/2017/11/051. arXiv: 1705.01092 [astro-ph.CO].

[54] Alexandre Barreira and Fabian Schmidt. "Responses in large-scale structure". In: *Journal of Cosmology and Astroparticle Physics* 2017.6, 053 (June 2017), p. 053. DOI: 10.1088/1475-7516/2017/06/053. arXiv: 1703.09212 [astro-ph.CO].

[55] Alexandre Barreira et al. "Galaxy bias and primordial non-Gaussianity: insights from galaxy formation simulations with IllustrisTNG". In: *JCAP* 12 (2020), p. 013. DOI: 10.1088/1475-7516/2020/12/013. arXiv: 2006.09368 [astro-ph.CO].

[56] Alexandre Barreira et al. "Separate Universe simulations with IllustrisTNG: baryonic effects on power spectrum responses and higher-order statistics". In: *Mon. Not. R. Astron. Soc.* 488.2 (Sept. 2019), pp. 2079–2092. DOI: 10.1093/mnras/stz1807. arXiv: 1904.02070 [astro-ph.CO].

[57] Adrian E. Bayer, Arka Banerjee, and Yu Feng. "A fast particle-mesh simulation of non-linear cosmological structure formation with massive neutrinos". In: *Journal of Cosmology and Astroparticle Physics* 2021.01 (Jan. 2021), pp. 016–016. DOI: 10.1088/1475-7516/2021/01/016. URL: https://doi.org/10.1088/1475-7516/2021/01/016.

[58] Adrian E. Bayer, Arka Banerjee, and Uroš Seljak. "Beware of fake ν 's: The effect of massive neutrinos on the nonlinear evolution of cosmic structure". In: *Phys. Rev. D* 105.12, 123510 (June 2022), p. 123510. DOI: 10.1103/PhysRevD.105.123510. arXiv: 2108.04215 [astro-ph.CO].

[59] Adrian E. Bayer, Chirag Modi, and Simone Ferraro. "Joint velocity and density reconstruction of the Universe with nonlinear differentiable forward modeling". In: (Oct. 2022). arXiv: 2210.15649 [astro-ph.CO].

[60] Adrian E. Bayer and Uroš Seljak. "The look-elsewhere effect from a unified Bayesian and frequentist perspective". In: *Journal of Cosmology and Astroparticle Physics* 2020.10, 009 (Oct. 2020), p. 009. DOI: 10.1088/1475-7516/2020/10/009. arXiv: 2007.13821 [physics.data-an].

[61] Adrian E. Bayer, Uroš Seljak, and Jakob Robnik. "Self-calibrating the look-elsewhere effect: fast evaluation of the statistical significance using peak heights". In: *Monthly Notices of the Royal Astronomical Society* 508.1 (Nov. 2021), pp. 1346–1357. DOI: 10.1093/mnras/stab2331. arXiv: 2108.06333 [astro-ph.IM].

[62] Adrian E. Bayer et al. "Detecting Neutrino Mass by Combining Matter Clustering, Halos, and Voids". In: *Astrophys. J.* 919.1, 24 (Sept. 2021), p. 24. DOI: 10.3847/1538-4357/ac0e91. arXiv: 2102.05049 [astro-ph.CO].

[63] Adrian E. Bayer et al. "Super-sample covariance of the power spectrum, bispectrum, halos, voids, and their cross-covariances". In: (Oct. 2022). arXiv: 2210.15647 [astro-ph.CO].

[64] Joel Bergé, Adam Amara, and Alexandre Réfrégier. "Optimal Capture of Non-Gaussianity in Weak-Lensing Surveys: Power Spectrum, Bispectrum, and Halo Counts". In: *The Astrophysical Journal* 712 (Apr. 2010), pp. 992–1002. DOI: 10.1088/0004-637X/712/2/992. arXiv: 0909.0529 [astro-ph.CO].

[65] Edmund Bertschinger and Avishai Dekel. "Recovering the Full Velocity and Density Fields from Large-Scale Redshift-Distance Samples". In: *Astrophys. J. Lett.* 336 (Jan. 1989), p. L5. DOI: 10.1086/185348.

[66] F. Beutler et al. "The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: signs of neutrino mass in current cosmological data sets". In: *Monthly Notices of the Royal Astronomical Society* 444.4 (Sept. 2014), pp. 3501–3516. ISSN: 1365-2966. DOI: 10.1093/mnras/stu1702. URL: http://dx.doi.org/10.1093/mnras/stu1702.

[67] Naren Bhandari et al. *Fisher Matrix Stability*. 2021. arXiv: 2101.00298 [astro-ph.CO].

[68] Simeon Bird, Matteo Viel, and Martin G. Haehnelt. "Massive neutrinos and the non-linear matter power spectrum". In: *Monthly Notices of the Royal Astronomical Society* 420.3 (Jan. 2012), pp. 2551–2561. ISSN: 0035-8711. DOI: 10.1111/j.1365-2966.2011.20222.x. URL: http://dx.doi.org/10.1111/j.1365-2966.2011.20222.x.

[69] Simeon Bird et al. "An efficient and accurate hybrid method for simulating non-linear neutrino structure". In: *Monthly Notices of the Royal Astronomical Society* 481.2 (Aug. 2018), pp. 1486–1500. ISSN: 0035-8711. DOI: 10.1093/mnras/sty2376. eprint: https://academic.oup.com/mnras/article-pdf/481/2/1486/25716087/sty2376.pdf. URL: https://doi.org/10.1093/mnras/sty2376.

[70] M. Birkinshaw and S. F. Gull. "A test for transverse motions of clusters of galaxies". In: *Nature* 302.5906 (Mar. 1983), pp. 315–317. DOI: 10.1038/302315a0.

[71] A. Blanchard et al. "Euclid preparation". In: *Astronomy & Astrophysics* 642 (Oct. 2020), A191. ISSN: 1432-0746. DOI: 10.1051/0004-6361/202038071. URL: http://dx.doi.org/10.1051/0004-6361/202038071.

[72] Michael R. Blanton et al. "Sloan Digital Sky Survey IV: Mapping the Milky Way, Nearby Galaxies, and the Distant Universe". In: *AJ* 154.1, 28 (July 2017), p. 28. DOI: 10.3847/1538-3881/aa7567. arXiv: 1703.00052 [astro-ph.GA].

[73] Diego Blas, Julien Lesgourgues, and Thomas Tram. "The Cosmic Linear Anisotropy Solving System (CLASS). Part II: Approximation schemes". In: *Journal of Cosmology and Astroparticle Physics* 2011.7, 034 (July 2011), p. 034. DOI: 10.1088/1475-7516/2011/07/034. arXiv: 1104.2933 [astro-ph.CO].

[74] L. Blot et al. "Non-linear matter power spectrum covariance matrix errors and cosmological parameter uncertainties". In: *Monthly Notices of the Royal Astronomical Society* 458 (June 2016), pp. 4462–4470. DOI: 10.1093/mnras/stw604. arXiv: 1512.05383 [astro-ph.CO].

[75] V. Böhm et al. "MADLens, a python package for fast and differentiable non-Gaussian lensing simulations". In: *Astronomy and Computing* 36, 100490 (July 2021), p. 100490. DOI: 10.1016/j.ascom.2021.100490. arXiv: 2012.07266 [astro-ph.CO].

[76] J. R. Bond, A. H. Jaffe, and L. Knox. "Estimating the power spectrum of the cosmic microwave background". In: *Phys. Rev. D* 57.4 (Feb. 1998), pp. 2117–2137. DOI: 10.1103/PhysRevD.57.2117. arXiv: astro-ph/9708203 [astro-ph].

[77] C. E. Bonferroni. "Teoria statistica delle classi e calcolo delle probabilità". In: *Pubblicazioni del R Istituto Superiore di Scienze Economiche e Commerciali di Firenze* (1936).

[78] M. Bonici et al. "Euclid: Forecasts from the void-lensing cross-correlation". In: *arXiv e-prints*, arXiv:2206.14211 (June 2022), arXiv:2206.14211. arXiv: 2206.14211 [astro-ph.CO].

[79] William J. Borucki et al. "Kepler Planet-Detection Mission: Introduction and First Results". In: *Science* 327.5968 (Feb. 2010), p. 977. DOI: 10.1126/science.1185402.

[80] E. G. Patrick Bos et al. "The darkness that shaped the void: dark energy and cosmic voids". In: *Monthly Notices of the Royal Astronomical Society* 426.1 (Sept. 2012), pp. 440–461. ISSN: 0035-8711. DOI: 10.1111/j.1365-2966.2012.21478.x. URL: http://dx.doi.org/10.1111/j.1365-2966.2012.21478.x.

[81] F. R. Bouchet et al. *Perturbative Lagrangian Approach to Gravitational Instability*. 1994. arXiv: astro-ph/9406013 [astro-ph].

[82] Jacob Brandbyge and Steen Hannestad. "Grid based linear neutrino perturbations in cosmologicalN-body simulations". In: *Journal of Cosmology and Astroparticle Physics* 2009.05 (May 2009), pp. 002–002. ISSN: 1475-7516. DOI: 10.1088/1475-7516/2009/05/002. URL: http://dx.doi.org/10.1088/1475-7516/2009/05/002.

[83] Jacob Brandbyge and Steen Hannestad. "Resolving cosmic neutrino structure: a hybrid neutrinoN-body scheme". In: *Journal of Cosmology and Astroparticle Physics* 2010.01 (Jan. 2010), pp. 021–021. ISSN: 1475-7516. DOI: 10.1088/1475-7516/2010/01/021. URL: http://dx.doi.org/10.1088/1475-7516/2010/01/021.

[84] Jacob Brandbyge, Steen Hannestad, and Thomas Tram. "Momentum space sampling of neutrinos in N-body simulations". In: *Journal of Cosmology and Astroparticle Physics* 2019.03 (Mar. 2019), pp. 047–047. ISSN: 1475-7516. DOI: 10.1088/1475-7516/2019/03/047. URL: http://dx.doi.org/10.1088/1475-7516/2019/03/047.

[85] Thejs Brinckmann et al. "The promising future of a robust cosmological neutrino mass measurement". In: *Journal of Cosmology and Astroparticle Physics* 2019.01 (Jan. 2019), pp. 059–059. ISSN: 1475-7516. DOI: 10.1088/1475-7516/2019/01/059. URL: http://dx.doi.org/10.1088/1475-7516/2019/01/059.

[86] Richard H. Byrd et al. "A Limited Memory Algorithm for Bound Constrained Optimization". In: *SIAM Journal on Scientific Computing* 16.5 (1995), pp. 1190–1208. DOI: 10.1137/0916069. eprint: https://doi.org/10.1137/0916069. URL: https://doi.org/10.1137/0916069.

[87] Lachlan A. Campbell et al. "The 6dF Galaxy Survey: Fundamental Plane data". In: *Mon. Not. R. Astron. Soc.* 443.2 (Sept. 2014), pp. 1231–1251. DOI: 10.1093/mnras/stu1198. arXiv: 1406.4867 [astro-ph.GA].

[88] Kipp Cannon, Chad Hanna, and Jacob Peoples. *Likelihood-Ratio Ranking Statistic for Compact Binary Coalescence Candidates with Rate Estimation*. 2015. arXiv: 1504.04632 [astro-ph.IM].

[89] Carmelita Carbone, Margarita Petkova, and Klaus Dolag. "DEMNUni: ISW, Rees-Sciama, and weak-lensing in the presence of massive neutrinos". In: *Journal of Cosmology and Astroparticle Physics* 2016.07 (July 2016), pp. 034–034. ISSN: 1475-7516. DOI: 10.1088/1475-7516/2016/07/034. URL: http://dx.doi.org/10.1088/1475-7516/2016/07/034.

[90] J. Carron. "On the assumption of Gaussianity for cosmological two-point statistics and parameter dependent covariance matrices". In: *Astron. Astrophys.* 551, A88 (Mar. 2013), A88. DOI: 10.1051/0004-6361/201220538. arXiv: 1204.4724 [astro-ph.CO].

[91] Julien Carron. "Information Escaping the Correlation Hierarchy of the Convergence Field in the Study of Cosmological Parameters". In: *Physical Review Letters* 108.7 (Feb. 2012). ISSN: 1079-7114. DOI: 10.1103/physrevlett.108.071301. URL: http://dx.doi.org/10.1103/PhysRevLett.108.071301.

[92] Julien Carron. "On the Incompleteness of the Moment and Correlation Function Hierarchy as Probes of the Lognormal Field". In: *The Astrophysical Journal* 738, 86 (Sept. 2011), p. 86. DOI: 10.1088/0004-637X/738/1/86. arXiv: 1105.4467 [astro-ph.CO].

[93] Emanuele Castorina and Azadeh Moradinezhad Dizgah. "Local Primordial Non-Gaussianities and super-sample variance". In: *Journal of Cosmology and Astroparticle Physics* 2020.10, 007 (Oct. 2020), p. 007. DOI: 10.1088/1475-7516/2020/10/007. arXiv: 2005.14677 [astro-ph.CO].

[94] Emanuele Castorina et al. "Cosmology with massive neutrinos II: on the universality of the halo mass function and bias". In: *Journal of Cosmology and Astroparticle Physics* 2014.02 (Feb. 2014), pp. 049–049. ISSN: 1475-7516. DOI: 10.1088/1475-7516/2014/02/049. URL: http://dx.doi.org/10.1088/1475-7516/2014/02/049.

[95] Emanuele Castorina et al. "DEMNUni: the clustering of large-scale structures in the presence of massive neutrinos". In: *Journal of Cosmology and Astroparticle Physics* 2015.07 (July 2015), pp. 043–043. ISSN: 1475-7516. DOI: 10.1088/1475-7516/2015/07/043. URL: http://dx.doi.org/10.1088/1475-7516/2015/07/043.

[96] M Cataneo et al. "On the road to percent accuracy – III. Non-linear reaction of the matter power spectrum to massive neutrinos". In: *Monthly Notices of the Royal Astronomical Society* 491.3 (Nov. 2019), pp. 3101–3107. ISSN: 1365-2966. DOI: 10.1093/mnras/stz3189. URL: http://dx.doi.org/10.1093/mnras/stz3189.

[97] M Cataneo et al. "On the road to percent accuracy: non-linear reaction of the matter power spectrum to dark energy and modified gravity". In: *Monthly Notices of the Royal Astronomical Society* 488.2 (July 2019), pp. 2121–2142. ISSN: 1365-2966. DOI: 10.1093/mnras/stz1836. URL: http://dx.doi.org/10.1093/mnras/stz1836.

[98] Marius Cautun et al. "Planes of satellite galaxies: when exceptions are the rule". In: *Monthly Notices of the Royal Astronomical Society* 452.4 (Aug. 2015), pp. 3838–3852. ISSN: 0035-8711. DOI: 10.1093/mnras/stv1557. eprint: https://academic.oup.com/mnras/article-pdf/452/4/3838/18240589/stv1557.pdf. URL: https://doi.org/10.1093/mnras/stv1557.

[99] Kwan Chuen Chan, Azadeh Moradinezhad Dizgah, and Jorge Noreña. "Bispectrum supersample covariance". In: *Phys. Rev. D* 97.4, 043532 (Feb. 2018), p. 043532. DOI: 10.1103/PhysRevD.97.043532. arXiv: 1709.02473 [astro-ph.CO].

[100] Kwan Chuen Chan et al. "Measurement of Void Bias Using Separate Universe Simulations". In: *Astrophys. J.* 889.2, 89 (Feb. 2020), p. 89. DOI: 10.3847/1538-4357/ab64ec. arXiv: 1909.03736 [astro-ph.CO].

[101] Nicolas Chartier et al. *CARPool: fast, accurate computation of large-scale structure statistics by pairing costly and cheap cosmological simulations*. 2020. arXiv: 2009.08970 [astro-ph.CO].

[102] Serguei Chatrchyan et al. "Observation of a New Boson at a Mass of 125 GeV with the CMS Experiment at the LHC". In: *Phys. Lett.* B716 (2012), pp. 30–61. DOI: 10.1016/j.physletb.2012.08.021. arXiv: 1207.7235 [hep-ex].

[103] Joe Zhiyu Chen, Amol Upadhye, and Yvonne Y. Y. Wong. *One line to run them all: SuperEasy massive neutrino linear response in N-body simulations*. 2020. arXiv: 2011.12504 [astro-ph.CO].

[104] Joe Zhiyu Chen, Amol Upadhye, and Yvonne Y. Y. Wong. *The cosmic neutrino background as a collection of fluids in large-scale structure simulations*. 2020. arXiv: 2011.12503 [astro-ph.CO].

[105] Ming-Hui Chen, Qi-Man Shao, and Joseph G. Ibrahim. "Estimating Ratios of Normalizing Constants". In: *Monte Carlo Methods in Bayesian Computation*. New York, NY: Springer New York, 2000, pp. 124–190. ISBN: 978-1-4612-1276-8. DOI: 10.1007/978-1-4612-1276-8_5. URL: https://doi.org/10.1007/978-1-4612-1276-8%5C_5.

[106] Ming-Hui Chen, Qi-Man Shao, and Joseph G. Ibrahim. *Monte Carlo Methods in Bayesian Computation*. Springer New York, 2000. DOI: 10.1007/978-1-4612-1276-8. URL: https://doi.org/10.1007%5C%2F978-1-4612-1276-8.

[107] Shi-Fan Chen et al. "Redshift-space distortions in Lagrangian perturbation theory". In: *Journal of Cosmology and Astroparticle Physics* 2021.03 (Mar. 2021), p. 100. ISSN: 1475-7516. DOI: 10.1088/1475-7516/2021/03/100. URL: http://dx.doi.org/10.1088/1475-7516/2021/03/100.

[108] Xingang Chen. "Primordial Non-Gaussianities from Inflation Models". In: *Advances in Astronomy* 2010 (2010), pp. 1–43. ISSN: 1687-7977. DOI: 10.1155/2010/638979. URL: http://dx.doi.org/10.1155/2010/638979.

[109] Sihao Cheng and Brice Ménard. "Weak lensing scattering transform: dark energy and neutrino mass sensitivity". In: *Monthly Notices of the Royal Astronomical Society* 507.1 (Oct. 2021), pp. 1012–1020. DOI: 10.1093/mnras/stab2102. arXiv: 2103.09247 [astro-ph.CO].

[110] Sihao Cheng et al. "A new approach to observational cosmology using the scattering transform". In: *Monthly Notices of the Royal Astronomical Society* 499.4 (Dec. 2020), pp. 5902–5914. DOI: 10.1093/mnras/staa3165. arXiv: 2006.08561 [astro-ph.CO].

[111] Shouvik Roy Choudhury and Steen Hannestad. *Updated results on neutrino mass and mass hierarchy from cosmology with Planck 2018 likelihoods*. 2019. arXiv: 1907.12598 [astro-ph.CO].

[112] Paweł Ciecielg et al. "Gaussianity of cosmic velocity fields and linearity of the velocity-gravity relation". In: *Mon. Not. R. Astron. Soc.* 339.3 (Mar. 2003), pp. 641–651. DOI: 10.1046/j.1365-8711.2003.06202.x. arXiv: astro-ph/0010364 [astro-ph].

[113] DESI Collaboration et al. *The DESI Experiment Part I: Science,Targeting, and Survey Design*. 2016. arXiv: 1611.00036 [astro-ph.IM].

[114] LSST Science Collaboration et al. *LSST Science Book, Version 2.0*. 2009. arXiv: 0912.0201 [astro-ph.IM].

[115] Planck Collaboration et al. *Planck 2018 results. IX. Constraints on primordial non-Gaussianity*. 2019. arXiv: 1905.05697 [astro-ph.CO].

[116] Stéphane Colombi, Michał J. Chodorowski, and Romain Teyssier. "Cosmic velocity-gravity relation in redshift space". In: *Mon. Not. R. Astron. Soc.* 375.1 (Feb. 2007), pp. 348–370. DOI: 10.1111/j.1365-2966.2006.11330.x. arXiv: 0805.1693 [astro-ph].

[117] S. Contarini et al. "Euclid: Cosmological forecasts from the void size function". In: *arXiv e-prints*, arXiv:2205.11525 (May 2022), arXiv:2205.11525. arXiv: 2205.11525 [astro-ph.CO].

[118] Sofia Contarini et al. "Cosmological exploitation of the size function of cosmic voids identified in the distribution of biased tracers". In: *Monthly Notices of the Royal Astronomical Society* 488.3 (July 2019), pp. 3526–3540. ISSN: 1365-2966. DOI: 10.1093/mnras/stz1989. URL: http://dx.doi.org/10.1093/mnras/stz1989.

[119] Matteo Costanzi et al. "Cosmology with massive neutrinos III: the halo mass function and an application to galaxy clusters". In: *Journal of Cosmology and Astroparticle Physics* 2013.12 (Dec. 2013), pp. 012–012. ISSN: 1475-7516. DOI: 10.1088/1475-7516/2013/12/012. URL: http://dx.doi.org/10.1088/1475-7516/2013/12/012.

[120] W. R. Coulton et al. "Constraining neutrino mass with the tomographic weak lensing bispectrum". In: *Journal of Cosmology and Astroparticle Physics* 5, 043 (May 2019), p. 043. DOI: 10.1088/1475-7516/2019/05/043. arXiv: 1810.02374.

[121] William R Coulton et al. "Quijote PNG: The information content of the halo power spectrum and bispectrum". In: *arXiv e-prints*, arXiv:2206.15450 (June 2022), arXiv:2206.15450. arXiv: 2206.15450 [astro-ph.CO].

[122] Hélène M. Courtois et al. "Three-dimensional Velocity and Density Reconstructions of the Local Universe with Cosmicflows-1". In: *Astrophys. J.* 744.1, 43 (Jan. 2012), p. 43. DOI: 10.1088/0004-637X/744/1/43. arXiv: 1109.3856 [astro-ph.CO].

[123] Robert D. Cousins. "The Jeffreys–Lindley paradox and discovery criteria in high energy physics". In: *Synthese* 194.2 (July 2014), pp. 395–432. ISSN: 1573-0964. DOI: 10.1007/s11229-014-0525-z. URL: http://dx.doi.org/10.1007/s11229-014-0525-z.

[124] Glen Cowan et al. "Asymptotic formulae for likelihood-based tests of new physics". In: *The European Physical Journal C* 71.2 (Feb. 2011). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-011-1554-0. URL: http://dx.doi.org/10.1140/epjc/s10052-011-1554-0.

[125] Paolo Creminelli et al. "Limits on non-Gaussianities from WMAP data". In: *Journal of Cosmology and Astroparticle Physics* 2006.05 (May 2006), pp. 004–004. ISSN: 1475-7516. DOI: 10.1088/1475-7516/2006/05/004. URL: http://dx.doi.org/10.1088/1475-7516/2006/05/004.

[126] Dylan Cromer et al. *Towards 1% accurate galaxy cluster masses: Including baryons in weak-lensing mass inference*. 2021. arXiv: 2104.06925 [astro-ph.CO].

[127] Guido D'Amico et al. "The cosmological analysis of the SDSS/BOSS data from the Effective Field Theory of Large-Scale Structure". In: *Journal of Cosmology and Astroparticle Physics* 2020.05 (May 2020), pp. 005–005. ISSN: 1475-7516. DOI: 10.1088/1475-7516/2020/05/005. URL: http://dx.doi.org/10.1088/1475-7516/2020/05/005.

[128] Biwei Dai, Yu Feng, and Uroš Seljak. "A gradient based method for modeling baryons and matter in halos of fast simulations". In: *Journal of Cosmology and Astroparticle Physics* 2018.11 (Nov. 2018), pp. 009–009. ISSN: 1475-7516. DOI: 10.1088/1475-7516/2018/11/009. URL: http://dx.doi.org/10.1088/1475-7516/2018/11/009.

[129] Ji-Ping Dai, Licia Verde, and Jun-Qing Xia. "What can we learn by combining the skew spectrum and the power spectrum?" In: *Journal of Cosmology and Astroparticle Physics* 2020.8, 007 (Aug. 2020), p. 007. DOI: 10.1088/1475-7516/2020/08/007. arXiv: 2002.09904 [astro-ph.CO].

[130] Jeppe Dakin et al. "νCONCEPT: cosmological neutrino simulations from the non-linear Boltzmann hierarchy". In: *Journal of Cosmology and Astroparticle Physics* 2019.02 (Feb. 2019), pp. 052–052. DOI: 10.1088/1475-7516/2019/02/052. URL: https://doi.org/10.1088%2F1475-7516%2F2019%2F02%2F052.

[131] Dark Energy Survey Collaboration et al. "The Dark Energy Survey: more than dark energy - an overview". In: *Mon. Not. R. Astron. Soc.* 460.2 (Aug. 2016), pp. 1270–1299. DOI: 10.1093/mnras/stw641. arXiv: 1601.00329 [astro-ph.CO].

[132] Christopher T. Davies et al. "Constraining cosmology with weak lensing voids". In: *arXiv e-prints*, arXiv:2010.11954 (Oct. 2020), arXiv:2010.11954. arXiv: 2010.11954 [astro-ph.CO].

[133] R. B. Davies. "Hypothesis Testing When a Nuisance Parameter is Present Only Under the Alternative". In: *Biometrika* 64.2 (1977), pp. 247–254. ISSN: 00063444. DOI: 10.2307/2335690. URL: http://www.jstor.org/stable/2335690.

[134] Robert B. Davies. "Hypothesis testing when a nuisance parameter is present only under the alternative". In: *Biometrika* 74 (1987), pp. 33–43. DOI: 10.1093/biomet/74.1.33.

[135] M. Davis et al. "The evolution of large-scale structure in a universe dominated by cold dark matter". In: *The Astrophysical Journal* 292 (May 1985), pp. 371–394. DOI: 10.1086/163168.

[136] Kyle S. Dawson et al. "THE SDSS-IV EXTENDED BARYON OSCILLATION SPECTROSCOPIC SURVEY: OVERVIEW AND EARLY DATA". In: *The Astronomical Journal* 151.2 (Feb. 2016), p. 44. ISSN: 1538-3881. DOI: 10.3847/0004-6256/151/2/44. URL: http://dx.doi.org/10.3847/0004-6256/151/2/44.

[137] Roland de Putter et al. "Thinking outside the box: effects of modes larger than the survey on matter power spectrum covariance". In: *Journal of Cosmology and Astroparticle Physics* 2012.4, 019 (Apr. 2012), p. 019. DOI: 10.1088/1475-7516/2012/04/019. arXiv: 1111.6596 [astro-ph.CO].

[138] Stijn N. B. Debackere, Joop Schaye, and Henk Hoekstra. "How baryons can significantly bias cluster count cosmology". In: *Mon. Not. R. Astron. Soc.* 505.1 (July 2021), pp. 593–609. DOI: 10.1093/mnras/stab1326. arXiv: 2101.07800 [astro-ph.CO].

[139] A. Dekel et al. "POTENT Reconstruction from Mark III Velocities". In: *Astrophys. J.* 522.1 (Sept. 1999), pp. 1–38. DOI: 10.1086/307636. arXiv: astro-ph/9812197 [astro-ph].

[140] J.-B. Delisle, N. Hara, and D. Ségransan. "Efficient modeling of correlated noise". In: *Astron. Astrophys.* 635 (Mar. 2020), A83. ISSN: 1432-0746. DOI: 10.1051/0004-6361/201936905. URL: http://dx.doi.org/10.1051/0004-6361/201936905.

[141] Joseph DeRose et al. "The Aemulus Project. I. Numerical Simulations for Precision Cosmology". In: *The Astrophysical Journal* 875.1 (Apr. 2019), p. 69. ISSN: 1538-4357. DOI: 10.3847/1538-4357/ab1085. URL: http://dx.doi.org/10.3847/1538-4357/ab1085.

[142] DES Collaboration et al. "Dark Energy Survey Year 3 Results: Constraints on extensions to ΛCDM with weak lensing and galaxy clustering". In: *arXiv e-prints*, arXiv:2207.05766 (July 2022), arXiv:2207.05766. arXiv: 2207.05766 [astro-ph.CO].

[143] DESI Collaboration et al. "The DESI Experiment Part I: Science,Targeting, and Survey Design". In: *arXiv e-prints*, arXiv:1611.00036 (Oct. 2016), arXiv:1611.00036. arXiv: 1611.00036 [astro-ph.IM].

[144] Anne-Sylvie Deutsch et al. "Reconstruction of the remote dipole and quadrupole fields from the kinetic Sunyaev Zel'dovich and polarized Sunyaev Zel'dovich effects". In: *Phys. Rev. D* 98.12 (2018), p. 123501. DOI: 10.1103/PhysRevD.98.123501. arXiv: 1707.08129 [astro-ph.CO].

[145] A. G. Doroshkevich et al. "COSMOLOGICAL IMPACT OF THE NEUTRINO REST MASS". In: *Annals of the New York Academy of Sciences* 375.1 (1981), pp. 32–42. DOI: 10.1111/j.1749-6632.1981.tb33688.x. eprint: https://nyaspubs.onlinelibrary.wiley.com/doi/pdf/10.1111/j.1749-6632.1981.tb33688.x. URL: https://nyaspubs.onlinelibrary.wiley.com/doi/abs/10.1111/j.1749-6632.1981.tb33688.x.

[146] Andrej Dvornik et al. "KiDS-1000: Combined halo-model cosmology constraints from galaxy abundance, galaxy clustering and galaxy-galaxy lensing". In: *arXiv e-prints*, arXiv:2210.03110 (Oct. 2022), arXiv:2210.03110. arXiv: 2210.03110 [astro-ph.CO].

[147] Michael Eickenberg et al. "Wavelet Moments for Cosmological Parameter Estimation". In: *arXiv e-prints*, arXiv:2204.07646 (Apr. 2022), arXiv:2204.07646. arXiv: 2204.07646 [astro-ph.CO].

[148] Daniel J. Eisenstein and Wayne Hu. "Power Spectra for Cold Dark Matter and Its Variants". In: *The Astrophysical Journal* 511.1 (Jan. 1999), pp. 5–15. ISSN: 1538-4357. DOI: 10.1086/306640. URL: http://dx.doi.org/10.1086/306640.

[149] Daniel J. Eisenstein et al. "Improving Cosmological Distance Measurements by Reconstruction of the Baryon Acoustic Peak". In: *The Astrophysical Journal* 664.2 (Aug. 2007), pp. 675–679. ISSN: 1538-4357. DOI: 10.1086/518712. URL: http://dx.doi.org/10.1086/518712.

[150] J. D. Emberson et al. "Cosmological neutrino simulations at extreme scale". In: *Research in Astronomy and Astrophysics* 17.8 (Aug. 2017), p. 085. ISSN: 1674-4527. DOI: 10.1088/1674-4527/17/8/85. URL: http://dx.doi.org/10.1088/1674-4527/17/8/85.

[151] Kimberly Emig, Cecilia Lunardini, and Rogier Windhorst. "Do high energy astrophysical neutrinos trace star formation?" In: *Journal of Cosmology and Astroparticle Physics* 2015.12 (Dec. 2015), pp. 029–029. ISSN: 1475-7516. DOI: 10.1088/1475-7516/2015/12/029. URL: http://dx.doi.org/10.1088/1475-7516/2015/12/029.

[152] Andreas Faltenbacher and Simon D. M. White. "ASSEMBLY BIAS AND THE DYNAMICAL STRUCTURE OF DARK MATTER HALOS". In: *The Astrophysical Journal* 708.1 (Dec. 2009), pp. 469–473. ISSN: 1538-4357. DOI: 10.1088/0004-637x/708/1/469. URL: http://dx.doi.org/10.1088/0004-637X/708/1/469.

[153] Yu Feng, Uroš Seljak, and Matias Zaldarriaga. "Exploring the posterior surface of the large scale structure reconstruction". In: *Journal of Cosmology and Astroparticle Physics* 2018.07 (July 2018), pp. 043–043. ISSN: 1475-7516. DOI: 10.1088/1475-7516/2018/07/043. URL: http://dx.doi.org/10.1088/1475-7516/2018/07/043.

[154] Yu Feng et al. "FASTPM: a new scheme for fast simulations of dark matter and haloes". In: *Monthly Notices of the Royal Astronomical Society* 463.3 (Dec. 2016), pp. 2273–2286. DOI: 10.1093/mnras/stw2123. arXiv: 1603.00476 [astro-ph.CO].

[155] J. R. Fergusson et al. "Combining power spectrum and bispectrum measurements to detect oscillatory features". In: *Phys. Rev.* D91.2 (2015), p. 023502. DOI: 10.1103/PhysRevD.91.023502. arXiv: 1410.5114 [astro-ph.CO].

[156] J. R. Fergusson et al. "Polyspectra searches for sharp oscillatory features in cosmic microwave sky data". In: *Phys. Rev.* D91.12 (2015), p. 123506. DOI: 10.1103/PhysRevD.91.123506. arXiv: 1412.6152 [astro-ph.CO].

[157] Simone Ferraro and Kendrick M. Smith. "Using large scale structure to measure f_{NL}, g_{NL} and τ_{NL}". In: *Phys. Rev. D* 91.4 (2015), p. 043506. DOI: 10.1103/PhysRevD.91.043506. arXiv: 1408.3126 [astro-ph.CO].

[158] Christian Fidler et al. "Relativistic initial conditions for N-body simulations". In: *Journal of Cosmology and Astroparticle Physics* 2017.06 (June 2017), pp. 043–043. ISSN: 1475-7516. DOI: 10.1088/1475-7516/2017/06/043. URL: http://dx.doi.org/10.1088/1475-7516/2017/06/043.

[159] R. A. Fisher. "Theory of Statistical Estimation". In: *Mathematical Proceedings of the Cambridge Philosophical Society* 22.5 (1925), pp. 700–725. DOI: 10.1017/S0305004100009580.

[160] A. Font-Ribera et al. "DESI and other Dark Energy experiments in the era of neutrino mass measurements". In: *Journal of Cosmology and Astroparticle Physics* 5, 023 (May 2014), p. 023. DOI: 10.1088/1475-7516/2014/05/023. arXiv: 1308.4164.

[161] Andreu Font-Ribera et al. "DESI and other Dark Energy experiments in the era of neutrino mass measurements". In: *Journal of Cosmology and Astroparticle Physics* 2014.05 (May 2014), pp. 023–023. DOI: 10.1088/1475-7516/2014/05/023. URL: https://doi.org/10.1088%2F1475-7516%2F2014%2F05%2F023.

[162] Simon Foreman et al. "Baryonic effects on the matter bispectrum". In: *Monthly Notices of the Royal Astronomical Society* 498.2 (Oct. 2020), pp. 2887–2911. DOI: 10.1093/mnras/staa2523. arXiv: 1910.03597 [astro-ph.CO].

[163] Grant Foster. "The Cleanest Fourier Spectrum". In: *AJ* 109 (Apr. 1995), p. 1889. DOI: 10.1086/117416.

[164] A. Fowlie. "Bayesian and frequentist approaches to resonance searches". In: *Journal of Instrumentation* 14.10 (Oct. 2019), P10031–P10031. ISSN: 1748-0221. DOI: 10.1088/1748-0221/14/10/p10031. URL: http://dx.doi.org/10.1088/1748-0221/14/10/P10031.

[165] Wolfram Freudling et al. "Large-Scale Power Spectrum and Cosmological Parameters from SFI Peculiar Velocities". In: *Astrophys. J.* 523.1 (Sept. 1999), pp. 1–15. DOI: 10.1086/307707. arXiv: astro-ph/9904118 [astro-ph].

[166] Y. Fukuda et al. "Evidence for Oscillation of Atmospheric Neutrinos". In: *Physical Review Letters* 81.8 (Aug. 1998), pp. 1562–1567. ISSN: 1079-7114. DOI: 10.1103/physrevlett.81.1562. URL: http://dx.doi.org/10.1103/PhysRevLett.81.1562.

[167] L. Gao and S. D. M. White. "Assembly bias in the clustering of dark matter haloes". In: *Monthly Notices of the Royal Astronomical Society: Letters* 377.1 (May 2007), pp. L5–L9. ISSN: 1745-3933. DOI: 10.1111/j.1745-3933.2007.00292.x. URL: http://dx.doi.org/10.1111/j.1745-3933.2007.00292.x.

[168] Elena Giusarma et al. *Learning neutrino effects in Cosmology with Convolutional Neural Networks*. 2019. arXiv: 1910.04255 [astro-ph.CO].

[169] Nickolay Y. Gnedin, Andrey V. Kravtsov, and Douglas H. Rudd. "Implementing the DC Mode in Cosmological Simulations with Supercomoving Variables". In: *Astrophys. J. Suppl.* 194.2, 46 (June 2011), p. 46. DOI: 10.1088/0067-0049/194/2/46. arXiv: 1104.1428 [astro-ph.CO].

[170] L. E. H. Godfrey et al. "Science at Very High Angular Resolution with the Square Kilometre Array". In: *Publications of the Astronomical Society of Australia* 29.1 (2012), pp. 42–53. ISSN: 1448-6083. DOI: 10.1071/as11050. URL: http://dx.doi.org/10.1071/AS11050.

[171] Alvaro Gonzalez. "Measurement of Areas on a Sphere Using Fibonacci and Latitude–Longitude Lattices". In: *Mathematical Geosciences* 42.1 (Nov. 2009), pp. 49–64. ISSN: 1874-8953. DOI: 10.1007/s11004-009-9257-x. URL: http://dx.doi.org/10.1007/s11004-009-9257-x.

[172] K. M. Gorski et al. "HEALPix: A Framework for High-Resolution Discretization and Fast Analysis of Data Distributed on the Sphere". In: *The Astrophysical Journal* 622.2 (Apr. 2005), pp. 759–771. ISSN: 1538-4357. DOI: 10.1086/427976. URL: http://dx.doi.org/10.1086/427976.

[173] S. Gouyou Beauchamps et al. "Impact of survey geometry and super-sample covariance on future photometric galaxy surveys". In: *Astron. Astrophys.* 659, A128 (Mar. 2022), A128. DOI: 10.1051/0004-6361/202142052. arXiv: 2109.02308 [astro-ph.CO].

[174] Eilam Gross and Ofer Vitells. "Trial factors for the look elsewhere effect in high energy physics". In: *The European Physical Journal C* 70.1-2 (Oct. 2010), pp. 525–530. ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-010-1470-8. URL: http://dx.doi.org/10.1140/epjc/s10052-010-1470-8.

[175] Davide Gualdi et al. "Matter trispectrum: theoretical modelling and comparison to N-body simulations". In: *arXiv e-prints*, arXiv:2009.02290 (Sept. 2020), arXiv:2009.02290. arXiv: 2009.02290 [astro-ph.CO].

[176] T. de Haan et al. "COSMOLOGICAL CONSTRAINTS FROM GALAXY CLUSTERS IN THE 2500 SQUARE-DEGREE SPT-SZ SURVEY". In: *The Astrophysical Journal* 832.1 (Nov. 2016), p. 95. ISSN: 1538-4357. DOI: 10.3847/0004-637x/832/1/95. URL: http://dx.doi.org/10.3847/0004-637X/832/1/95.

[177] ChangHoon Hahn and Francisco Villaescusa-Navarro. *Constraining M_ν with the Bispectrum II: The Total Information Content of the Galaxy Bispectrum*. 2020. arXiv: 2012.02200 [astro-ph.CO].

[178] ChangHoon Hahn et al. "Constraining M_ν with the bispectrum. Part I. Breaking parameter degeneracies". In: *Journal of Cosmology and Astroparticle Physics* 2020.03 (Mar. 2020), pp. 040–040. ISSN: 1475-7516. DOI: 10.1088/1475-7516/2020/03/040. URL: http://dx.doi.org/10.1088/1475-7516/2020/03/040.

[179] Zoltán Haiman, Joseph J. Mohr, and Gilbert P. Holder. "Constraints on Cosmological Parameters from Future Galaxy Cluster Surveys". In: *The Astrophysical Journal* 553.2 (June 2001), pp. 545–561. DOI: 10.1086/320939. arXiv: astro-ph/0002336 [astro-ph].

[180] Anik Halder and Alexandre Barreira. "Response approach to the integrated shear 3-point correlation function: the impact of baryonic effects on small scales". In: *arXiv e-prints*, arXiv:2201.05607 (Jan. 2022), arXiv:2201.05607. arXiv: 2201.05607 [astro-ph.CO].

[181] N. Hamaus et al. "Euclid: Forecasts from redshift-space distortions and the Alcock-Paczynski test with cosmic voids". In: *Astron. Astrophys.* 658, A20 (Feb. 2022), A20. DOI: 10.1051/0004-6361/202142073. arXiv: 2108.10347 [astro-ph.CO].

[182] Nico Hamaus et al. "Precision cosmology with voids in the final BOSS data". In: *Journal of Cosmology and Astroparticle Physics* 2020.12, 023 (Dec. 2020), p. 023. DOI: 10.1088/1475-7516/2020/12/023. arXiv: 2007.07895 [astro-ph.CO].

[183] A. J. S. Hamilton. "Towards optimal measurement of power spectra - I. Minimum variance pair weighting and the Fisher matrix". In: *Mon. Not. R. Astron. Soc.* 289.2 (Aug. 1997), pp. 285–294. DOI: 10.1093/mnras/289.2.285. arXiv: astro-ph/9701008 [astro-ph].

[184] Andrew J. S. Hamilton, Christopher D. Rimes, and Román Scoccimarro. "On measuring the covariance matrix of the non-linear power spectrum from simulations". In: *Monthly Notices of the Royal Astronomical Society* 371.3 (Sept. 2006), pp. 1188–1204. DOI: 10.1111/j.1365-2966.2006.10709.x. arXiv: astro-ph/0511416 [astro-ph].

[185] Nick Hand et al. "nbodykit: An Open-source, Massively Parallel Toolkit for Large-scale Structure". In: *The Astronomical Journal* 156.4 (Sept. 2018), p. 160. ISSN: 1538-3881. DOI: 10.3847/1538-3881/aadae0. URL: http://dx.doi.org/10.3847/1538-3881/aadae0.

[186] J H Hannay and J F Nye. "Fibonacci numerical integration on a sphere". In: *Journal of Physics A: Mathematical and General* 37.48 (Nov. 2004), pp. 11591–11601. DOI: 10.1088/0305-4470/37/48/005. URL: https://doi.org/10.1088%2F0305-4470%2F37%2F48%2F005.

[187] S. Hannestad, H. Tu, and Y. Y. Wong. "Measuring neutrino masses and dark energy with weak lensing tomography". In: *Journal of Cosmology and Astroparticle Physics* 6, 025 (June 2006), p. 025. DOI: 10.1088/1475-7516/2006/06/025. eprint: astro-ph/0603019.

[188] Joachim Harnois-Déraps et al. "Cosmic Shear Cosmology Beyond 2-Point Statistics: A Combined Peak Count and Correlation Function Analysis of DES-Y1". In: *arXiv e-prints*, arXiv:2012.02777 (Dec. 2020), arXiv:2012.02777. arXiv: 2012.02777 [astro-ph.CO].

[189] M. Hazumi et al. "LiteBIRD: a small satellite for the study of B-mode polarization and inflation from cosmic background radiation detection". In: *Space Telescopes and Instrumentation 2012: Optical, Infrared, and Millimeter Wave*. Ed. by Mark C. Clampin et al. Vol. 8442. International Society for Optics and Photonics. SPIE, 2012, pp. 451–459. DOI: 10.1117/12.926743. URL: https://doi.org/10.1117/12.926743.

[190] D. J. Heath. "The growth of density perturbations in zero pressure Friedmann-Lemaitre universes." In: *Monthly Notices of the Royal Astronomical Society* 179 (May 1977), pp. 351–358. DOI: 10.1093/mnras/179.3.351.

[191] A. F. Heavens, T. D. Kitching, and L. Verde. "On model selection forecasting, dark energy and modified gravity". In: *Monthly Notices of the Royal Astronomical Society* 380 (Sept. 2007), pp. 1029–1035. DOI: 10.1111/j.1365-2966.2007.12134.x. arXiv: astro-ph/0703191 [astro-ph].

[192] Alan Heavens. "Statistical techniques in cosmology". In: *arXiv e-prints*, arXiv:0906.0664 (June 2009), arXiv:0906.0664. arXiv: 0906.0664 [astro-ph.CO].

[193] J. Colin Hill and Enrico Pajer. "Cosmology from the thermal Sunyaev-Zel'dovich power spectrum: Primordial non-Gaussianity and massive neutrinos". In: *Phys. Rev. D* 88 (6 Sept. 2013), p. 063526. DOI: 10.1103/PhysRevD.88.063526. URL: https://link.aps.org/doi/10.1103/PhysRevD.88.063526.

[194] Yehuda Hoffman, Hélène M. Courtois, and R. Brent Tully. "Cosmic bulk flow and the local motion from Cosmicflows-2". In: *Mon. Not. R. Astron. Soc.* 449.4 (June 2015), pp. 4494–4505. DOI: 10.1093/mnras/stv615. arXiv: 1503.05422 [astro-ph.CO].

[195] Yehuda Hoffman and Erez Ribak. "Constrained Realizations of Gaussian Fields: A Simple Algorithm". In: *Astrophys. J. Lett.* 380 (Oct. 1991), p. L5. DOI: 10.1086/186160.

[196] B. Horowitz, U. Seljak, and G. Aslanyan. "Efficient optimal reconstruction of linear fields and band-powers from cosmological data". In: *Journal of Cosmology and Astroparticle Physics* 2019.10 (Oct. 2019), pp. 035–035. ISSN: 1475-7516. DOI: 10.1088/1475-7516/2019/10/035. URL: http://dx.doi.org/10.1088/1475-7516/2019/10/035.

[197] Benjamin Horowitz et al. "TARDIS. I. A Constrained Reconstruction Approach to Modeling the z $\sim$ 2.5 Cosmic Web Probed by Lyα Forest Tomography". In: *The Astrophysical Journal* 887.1 (Dec. 2019), p. 61. ISSN: 1538-4357. DOI: 10.3847/1538-4357/ab4d4c. URL: http://dx.doi.org/10.3847/1538-4357/ab4d4c.

[198] Selim C. Hotinli et al. "Transverse Velocities with the Moving Lens Effect". In: *Phys. Rev. Lett.* 123.6 (2019), p. 061301. DOI: 10.1103/PhysRevLett.123.061301. arXiv: 1812.03167 [astro-ph.CO].

[199] Wayne Hu, Daniel J. Eisenstein, and Max Tegmark. "Weighing Neutrinos with Galaxy Surveys". In: *Physical Review Letters* 80.24 (June 1998), pp. 5255–5258. ISSN: 1079-7114. DOI: 10.1103/physrevlett.80.5255. URL: http://dx.doi.org/10.1103/PhysRevLett.80.5255.

[200] Wayne Hu and Andrey V. Kravtsov. "Sample Variance Considerations for Cluster Surveys". In: *Astrophys. J.* 584.2 (Feb. 2003), pp. 702–715. DOI: 10.1086/345846. arXiv: astro-ph/0203169 [astro-ph].

[201] Paul Hunt and Subir Sarkar. "Search for features in the spectrum of primordial perturbations using Planck and other datasets". In: *Journal of Cosmology and Astroparticle Physics* 2015.12 (Dec. 2015), pp. 052–052. ISSN: 1475-7516. DOI: 10.1088/1475-7516/2015/12/052. URL: http://dx.doi.org/10.1088/1475-7516/2015/12/052.

[202] Kiyotomo Ichiki and Masahiro Takada. "Impact of massive neutrinos on the abundance of massive clusters". In: *Phys. Rev. D* 85, 063521 (Mar. 2012), p. 063521. DOI: 10.1103/PhysRevD.85.063521. arXiv: 1108.4688 [astro-ph.CO].

[203] Derek Inman and Ue-Li Pen. "Cosmic neutrinos: A dispersive and nonlinear fluid". In: *Physical Review D* 95.6 (Mar. 2017). ISSN: 2470-0029. DOI: 10.1103/physrevd.95.063535. URL: http://dx.doi.org/10.1103/PhysRevD.95.063535.

[204] Mikhail M. Ivanov, Marko Simonović, and Matias Zaldarriaga. "Cosmological parameters and neutrino masses from the final Planck and full-shape BOSS data". In: *Physical Review D* 101.8 (Apr. 2020). ISSN: 2470-0029. DOI: 10.1103/physrevd.101.083504. URL: http://dx.doi.org/10.1103/PhysRevD.101.083504.

[205] Drew Jamieson and Marilena Loverde. "Separate universe void bias". In: *Phys. Rev. D* 100.12, 123528 (Dec. 2019), p. 123528. DOI: 10.1103/PhysRevD.100.123528. arXiv: 1909.05313 [astro-ph.CO].

[206] J. Jasche and G. Lavaux. "Physical Bayesian modelling of the non-linear matter distribution: New insights into the nearby universe". In: *Astron. Astrophys.* 625, A64 (May 2019), A64. DOI: 10.1051/0004-6361/201833710. arXiv: 1806.11117 [astro-ph.CO].

[207] Jens Jasche and Benjamin D. Wandelt. "Bayesian physical reconstruction of initial conditions from large-scale structure surveys". In: *Mon. Not. R. Astron. Soc.* 432.2 (June 2013), pp. 894–913. DOI: 10.1093/mnras/stt449. arXiv: 1203.3639 [astro-ph.CO].

[208] E. Jennings, Y. Li, and W. Hu. "The abundance of voids and the excursion set formalism". In: *Monthly Notices of the Royal Astronomical Society* 434.3 (July 2013), pp. 2167–2181. ISSN: 1365-2966. DOI: 10.1093/mnras/stt1169. URL: http://dx.doi.org/10.1093/mnras/stt1169.

[209] He Jia and Uroš Seljak. *Normalizing Constant Estimation with Gaussianized Bridge Sampling*. 2019. arXiv: 1912.06073 [stat.ML].

[210] Y. P. Jing. "Correcting for the Alias Effect When Measuring the Power Spectrum Using a Fast Fourier Transform". In: *Astrophys. J.* 620.2 (Feb. 2005), pp. 559–563. DOI: 10.1086/427087. arXiv: astro-ph/0409240 [astro-ph].

[211] T. Kacprzak et al. "Cosmology constraints from shear peak statistics in Dark Energy Survey Science Verification data". In: *ArXiv e-prints* (Mar. 2016). arXiv: 1603.05040.

[212] Manoj Kaplinghat, Lloyd Knox, and Yong-Seon Song. "Determining Neutrino Mass from the Cosmic Microwave Background Alone". In: *Phys. Rev. Lett.* 91 (24 Dec. 2003), p. 241301. DOI: 10.1103/PhysRevLett.91.241301. URL: https://link.aps.org/doi/10.1103/PhysRevLett.91.241301.

[213] Issha Kayo, Masahiro Takada, and Bhuvnesh Jain. "Information content of weak lensing power spectrum and bispectrum: including the non-Gaussian error covariance matrix". In: *Monthly Notices of the Royal Astronomical Society* 429 (Feb. 2013), pp. 344–371. DOI: 10.1093/mnras/sts340. arXiv: 1207.6322 [astro-ph.CO].

[214] Alex Kim et al. "Testing Gravity Using Type Ia Supernovae Discovered by Next-Generation Wide-Field Imaging Surveys". In: *Bulletin of the AAS* 51.3, 140 (May 2019), p. 140. arXiv: 1903.07652 [astro-ph.CO].

[215] Alex G. Kim and Eric V. Linder. "Complementarity of peculiar velocity surveys and redshift space distortions for testing gravity". In: *Phys. Rev. D* 101.2, 023516 (Jan. 2020), p. 023516. DOI: 10.1103/PhysRevD.101.023516. arXiv: 1911.09121 [astro-ph.CO].

[216] Mischa Knabenhans et al. "Euclid preparation: II. The EuclidEmulator – a tool to compute the cosmology dependence of the nonlinear matter power spectrum". In: *Monthly Notices of the Royal Astronomical Society* 484.4 (Jan. 2019), pp. 5509–5529. ISSN: 1365-2966. DOI: 10.1093/mnras/stz197. URL: http://dx.doi.org/10.1093/mnras/stz197.

[217] David G Koch et al. "Kepler mission design, realized photometric performance, and early science". In: *The Astrophysical Journal Letters* 713.2 (2010), p. L79.

[218] Doogesh Kodi Ramanah et al. "Super-resolution emulator of cosmological simulations using deep physical models". In: *Monthly Notices of the Royal Astronomical Society* 495.4 (May 2020), pp. 4227–4236. ISSN: 0035-8711. DOI: 10.1093/mnras/staa1428. eprint: https://academic.oup.com/mnras/article-pdf/495/4/4227/33372045/staa1428.pdf. URL: https://doi.org/10.1093/mnras/staa1428.

[219] E. Komatsu et al. "FIVE-YEAR WILKINSON MICROWAVE ANISOTROPY PROBE OBSERVATIONS: COSMOLOGICAL INTERPRETATION". In: *The Astrophysical Journal Supplement Series* 180.2 (Feb. 2009), pp. 330–376. ISSN: 1538-4365. DOI: 10.1088/0067-0049/180/2/330. URL: http://dx.doi.org/10.1088/0067-0049/180/2/330.

[220] Ehsan Kourkchi et al. "Cosmicflows-4: The Calibration of Optical and Infrared Tully-Fisher Relations". In: *Astrophys. J.* 896.1, 3 (June 2020), p. 3. DOI: 10.3847/1538-4357/ab901c. arXiv: 2004.14499 [astro-ph.GA].

[221] Christina D. Kreisch, Francis-Yan Cyr-Racine, and Olivier Doré. "Neutrino puzzle: Anomalies, interactions, and cosmological tensions". In: *Phys. Rev. D* 101.12, 123505 (June 2020), p. 123505. DOI: 10.1103/PhysRevD.101.123505. arXiv: 1902.00534 [astro-ph.CO].

[222] Christina D. Kreisch et al. "Massive neutrinos leave fingerprints on cosmic voids". In: *Monthly Notices of the Royal Astronomical Society* 488.3 (Sept. 2019), pp. 4413–4426. DOI: 10.1093/mnras/stz1944. arXiv: 1808.07464 [astro-ph.CO].

[223] Christina D. Kreisch et al. "The GIGANTES Data Set: Precision Cosmology from Voids in the Machine-learning Era". In: *Astrophys. J.* 935.2, 100 (Aug. 2022), p. 100. DOI: 10.3847/1538-4357/ac7d4b. arXiv: 2107.02304 [astro-ph.CO].

[224] Joseph Kuruvilla and Nabila Aghanim. *Information content in mean pairwise velocity and mean relative velocity between pairs in a triplet*. 2021. arXiv: 2102.06709 [astro-ph.CO].

[225] Juliana Kwan et al. "Cosmic Emulation: Fast Predictions for the Galaxy Power Spectrum". In: *Astrophys. J.* 810.1, 35 (Sept. 2015), p. 35. DOI: 10.1088/0004-637X/810/1/35. arXiv: 1311.6444 [astro-ph.CO].

[226] Fabien Lacasa and Julien Grain. "Fast and easy super-sample covariance of large-scale structure observables". In: *Astron. Astrophys.* 624, A61 (Apr. 2019), A61. DOI: 10.1051/0004-6361/201834343. arXiv: 1809.05437 [astro-ph.CO].

[227] Fabien Lacasa et al. "Efficient Computation of Super-Sample Covariance for Stage IV Galaxy Surveys". In: *arXiv e-prints*, arXiv:2209.14421 (Sept. 2022), arXiv:2209.14421. arXiv: 2209.14421 [astro-ph.CO].

[228] I. Lacerna et al. "CENTRAL GALAXIES IN DIFFERENT ENVIRONMENTS: DO THEY HAVE SIMILAR PROPERTIES?" In: *The Astrophysical Journal* 788.1 (May 2014), p. 29. ISSN: 1538-4357. DOI: 10.1088/0004-637x/788/1/29. URL: http://dx.doi.org/10.1088/0004-637X/788/1/29.

[229] Ivan Lacerna and Nelson Padilla. "The nature of assembly bias - II. Halo spin". In: *Mon. Not. R. Astron. Soc.* 426.1 (Oct. 2012), pp. L26–L30. DOI: 10.1111/j.1745-3933.2012.01316.x. arXiv: 1207.4476 [astro-ph.CO].

[230] Pierre Simon Laplace. "Memoir on the Probability of the Causes of Events". In: *Statistical Science* 1.3 (1986), pp. 364–378. DOI: 10.1214/ss/1177013621. URL: https://doi.org/10.1214/ss/1177013621.

[231] R. Laureijs et al. *Euclid Definition Study Report*. 2011. arXiv: 1110.3193 [astro-ph.CO].

[232] Guilhem Lavaux. "Bayesian 3D velocity field reconstruction with VIRBIUS". In: *Mon. Not. R. Astron. Soc.* 457.1 (Mar. 2016), pp. 172–197. DOI: 10.1093/mnras/stv2915. arXiv: 1512.04534 [astro-ph.CO].

[233] Earl Lawrence et al. "The Coyote Universe. III. Simulation Suite and Precision Emulator for the Nonlinear Matter Power Spectrum". In: *The Astrophysical Journal* 713.2 (Apr. 2010), pp. 1322–1331. DOI: 10.1088/0004-637X/713/2/1322. arXiv: 0912.4490 [astro-ph.CO].

[234] Florent Leclercq, Jens Jasche, and Benjamin Wandelt. "Bayesian analysis of the dynamic cosmic web in the SDSS galaxy survey". In: *Journal of Cosmology and Astroparticle Physics* 2015.6 (June 2015), pp. 015–015. DOI: 10.1088/1475-7516/2015/06/015. arXiv: 1502.02690 [astro-ph.CO].

[235] Florent Leclercq et al. "The phase-space structure of nearby dark matter as constrained by the SDSS". In: *Journal of Cosmology and Astroparticle Physics* 2017.6, 049 (June 2017), p. 049. DOI: 10.1088/1475-7516/2017/06/049. arXiv: 1601.00093 [astro-ph.CO].

[236] J LESGOURGUES and S PASTOR. "Massive neutrinos and cosmology". In: *Physics Reports* 429.6 (July 2006), pp. 307–379. ISSN: 0370-1573. DOI: 10.1016/j.physrep.2006.04.001. URL: http://dx.doi.org/10.1016/j.physrep.2006.04.001.

[237] Julien Lesgourgues et al. "Recent times: neutrinos and structure formation". In: *Neutrino Cosmology*. Cambridge University Press, 2013, pp. 273–347. DOI: 10.1017/CBO9781139012874.007.

[238] Antony Lewis and Sarah Bridle. "Cosmological parameters from CMB and other data: A Monte Carlo approach". In: *Physical Review D* 66.10 (Nov. 2002). ISSN: 1089-4918. DOI: 10.1103/physrevd.66.103511. URL: http://dx.doi.org/10.1103/PhysRevD.66.103511.

[239] Antony Lewis, Anthony Challinor, and Anthony Lasenby. "Efficient Computation of Cosmic Microwave Background Anisotropies in Closed Friedmann-Robertson-Walker Models". In: *The Astrophysical Journal* 538.2 (Aug. 2000), pp. 473–476. DOI: 10.1086/309179. URL: https://doi.org/10.1086%5C%2F309179.

[240] Yin Li, Wayne Hu, and Masahiro Takada. "Super-sample covariance in simulations". In: *Phys. Rev. D* 89.8, 083519 (Apr. 2014), p. 083519. DOI: 10.1103/PhysRevD.89.083519. arXiv: 1401.0385 [astro-ph.CO].

[241] Yin Li, Marcel Schmittfull, and Uroš Seljak. "Galaxy power-spectrum responses and redshift-space super-sample effect". In: *Journal of Cosmology and Astroparticle Physics* 2018.2, 022 (Feb. 2018), p. 022. DOI: 10.1088/1475-7516/2018/02/022. arXiv: 1711.00018 [astro-ph.CO].

[242] Z. Li et al. "Constraining neutrino mass with tomographic weak lensing peak counts". In: *Phys. Rev. D* 99.6, 063527 (Mar. 2019), p. 063527. DOI: 10.1103/PhysRevD.99.063527. arXiv: 1810.01781.

[243] D. V. Lindley. "A Statistical Paradox". In: *Biometrika* 44.1/2 (1957), pp. 187–192. ISSN: 00063444. URL: http://www.jstor.org/stable/2333251.

[244] Adrian Liu et al. "Eliminating the optical depth nuisance from the CMB with 21 cm cosmology". In: *Phys. Rev. D* 93 (4 Feb. 2016), p. 043013. DOI: 10.1103/PhysRevD.93.043013. URL: https://link.aps.org/doi/10.1103/PhysRevD.93.043013.

[245] J. Liu et al. "Cosmology constraints from the weak lensing peak counts and the power spectrum in CFHTLenS data". In: *Phys. Rev. D* 91.6, 063507 (Mar. 2015), p. 063507. DOI: 10.1103/PhysRevD.91.063507. arXiv: 1412.0757.

[246] Jia Liu and Mathew S. Madhavacheril. "Constraining neutrino mass with the tomographic weak lensing one-point probability distribution function and power spectrum". In: *Phys. Rev. D* 99 (8 Apr. 2019), p. 083508. DOI: 10.1103/PhysRevD.99.083508. URL: https://link.aps.org/doi/10.1103/PhysRevD.99.083508.

[247] Jia Liu et al. "MassiveNuS: Cosmological massive neutrino simulations". In: *Journal of Cosmology and Astroparticle Physics* 2018.3 (2018). ISSN: 14757516. DOI: 10.1088/1475-7516/2018/03/049. URL: http://columbialensing.org..

[248] X. Liu et al. "Cosmological constraints from weak lensing peak statistics with Canada-France-Hawaii Telescope Stripe 82 Survey". In: *Monthly Notices of the Royal Astronomical Society* 450 (July 2015), pp. 2888–2902. DOI: 10.1093/mnras/stv784. arXiv: 1412.3683.

[249] N. R. Lomb. "Least-Squares Frequency Analysis of Unequally Spaced Data". In: *Astrophys. and Space Science* 39.2 (Feb. 1976), pp. 447–462. DOI: 10.1007/BF00648343.

[250] M. LoVerde and M. Zaldarriaga. "Neutrino clustering around spherical dark matter halos". In: *Phys. Rev. D* 89.6, 063502 (Mar. 2014), p. 063502. DOI: 10.1103/PhysRevD.89.063502. arXiv: 1310.6459.

[251] David J. C. MacKay. *Information Theory, Inference, and Learning Algorithms*. Copyright Cambridge University Press, 2003.

[252] Mathew S. Madhavacheril, Nicholas Battaglia, and Hironao Miyatake. "Fundamental physics from future weak-lensing calibrated Sunyaev-Zel'dovich galaxy cluster counts". In: *Phys. Rev. D* 96 (10 Nov. 2017), p. 103525. DOI: 10.1103/PhysRevD.96.103525. URL: https://link.aps.org/doi/10.1103/PhysRevD.96.103525.

[253] Juan Maldacena. "Non-gaussian features of primordial fluctuations in single field inflationary models". In: *Journal of High Energy Physics* 2003.05 (May 2003), pp. 013–013. ISSN: 1029-8479. DOI: 10.1088/1126-6708/2003/05/013. URL: http://dx.doi.org/10.1088/1126-6708/2003/05/013.

[254] Gabriela A. Marques et al. "Constraining neutrino mass with weak lensing Minkowski Functionals". In: *Journal of Cosmology and Astroparticle Physics* 2019.6, 019 (June 2019), p. 019. DOI: 10.1088/1475-7516/2019/06/019. arXiv: 1812.08206 [astro-ph.CO].

[255] N. Martinet et al. "KiDS-450: cosmological constraints from weak-lensing peak statistics - II: Inference from shear peaks using N-body simulations". In: *Monthly Notices of the Royal Astronomical Society* 474 (Feb. 2018), pp. 712–730. DOI: 10.1093/mnras/stx2793. arXiv: 1709.07678.

[256] Shogo Masaki, Takahiro Nishimichi, and Masahiro Takada. "Anisotropic separate universe simulations". In: *Mon. Not. R. Astron. Soc.* 496.1 (July 2020), pp. 483–496. DOI: 10.1093/mnras/staa1579. arXiv: 2003.10052 [astro-ph.CO].

[257] Elena Massara et al. "Using the Marked Power Spectrum to Detect the Signature of Neutrinos in Large-Scale Structure". In: *Phys. Rev. Lett.* 126.1, 011301 (Jan. 2021), p. 011301. DOI: 10.1103/PhysRevLett.126.011301. arXiv: 2001.11024 [astro-ph.CO].

[258] Elena Massara et al. "Voids in massive neutrino cosmologies". In: *Journal of Cosmology and Astroparticle Physics* 2015.11, 018 (Nov. 2015), p. 018. DOI: 10.1088/1475-7516/2015/11/018. arXiv: 1506.03088 [astro-ph.CO].

[259] Brendan McKay et al. "Solving the Bible Code Puzzle". In: *Statist. Sci.* 14.2 (May 1999), pp. 150–173. DOI: 10.1214/ss/1009212243. URL: https://doi.org/10.1214/ss/1009212243.

[260] P. Daniel Meerburg et al. *Primordial Non-Gaussianity*. 2019. arXiv: 1903.04409 [astro-ph.CO].

[261] Cody Messick et al. "Analysis framework for the prompt discovery of compact binary mergers in gravitational-wave data". In: *Physical Review D* 95.4 (Feb. 2017). ISSN: 2470-0029. DOI: 10.1103/physrevd.95.042001. URL: http://dx.doi.org/10.1103/PhysRevD.95.042001.

[262] Marius Millea and Uroš Seljak. "Marginal unbiased score expansion and application to CMB lensing". In: *Phys. Rev. D* 105.10, 103531 (May 2022), p. 103531. DOI: 10.1103/PhysRevD.105.103531. arXiv: 2112.09354 [astro-ph.CO].

[263] Rupert G. Miller. *Simultaneous Statistical Inference*. Springer New York, 1981. DOI: 10.1007/978-1-4613-8122-8. URL: https://doi.org/10.1007%5C%2F978-1-4613-8122-8.

[264] C. Modi, F. Lanusse, and U. Seljak. "FlowPM: Distributed TensorFlow implementation of the FastPM cosmological N-body solver". In: *Astronomy and Computing* 37, 100505 (Oct. 2021), p. 100505. DOI: 10.1016/j.ascom.2021.100505. arXiv: 2010.11847 [astro-ph.CO].

[265] Chirag Modi. "Reconstruction of Cosmological Fields in Forward Model Framework". PhD thesis. UC, Berkeley (main), 2020.

[266] Chirag Modi, Yu Feng, and Uroš Seljak. "Cosmological reconstruction from galaxy light: neural network based light-matter connection". In: *Journal of Cosmology and Astroparticle Physics* 2018.10 (Oct. 2018), pp. 028–028. ISSN: 1475-7516. DOI: 10.1088/1475-7516/2018/10/028. URL: http://dx.doi.org/10.1088/1475-7516/2018/10/028.

[267] Chirag Modi et al. "Intensity mapping with neutral hydrogen and the Hidden Valley simulations". In: *Journal of Cosmology and Astroparticle Physics* 2019.9, 024 (Sept. 2019), p. 024. DOI: 10.1088/1475-7516/2019/09/024. arXiv: 1904.11923 [astro-ph.CO].

[268] Irshad Mohammed and Uroš Seljak. "Analytic model for the matter power spectrum, its covariance matrix and baryonic effects". In: *Mon. Not. R. Astron. Soc.* 445.4 (Dec. 2014), pp. 3382–3400. DOI: 10.1093/mnras/stu1972. arXiv: 1407.0060 [astro-ph.CO].

[269] Eva-Maria Mueller et al. "Constraints on massive neutrinos from the pairwise kinematic Sunyaev-Zel'dovich effect". In: *Physical Review D* 92.6 (Sept. 2015). ISSN: 1550-2368. DOI: 10.1103/physrevd.92.063501. URL: http://dx.doi.org/10.1103/PhysRevD.92.063501.

[270] Moritz Münchmeyer et al. "Constraining local non-Gaussianities with kinetic Sunyaev-Zel'dovich tomography". In: *Phys. Rev. D* 100.8, 083508 (Oct. 2019), p. 083508. DOI: 10.1103/PhysRevD.100.083508. arXiv: 1810.13424 [astro-ph.CO].

[271] Seshadri Nadathur. "Testing cosmology with a catalogue of voids in the BOSS galaxy surveys". In: *Mon. Not. R. Astron. Soc.* 461.1 (Sept. 2016), pp. 358–370. DOI: 10.1093/mnras/stw1340. arXiv: 1602.04752 [astro-ph.CO].

[272] Savvas Nesseris and Juan García-Bellido. "Is the Jeffreys' scale a reliable tool for Bayesian model comparison in cosmology?" In: *Journal of Cosmology and Astroparticle Physics* 2013.08 (Aug. 2013), pp. 036–036. ISSN: 1475-7516. DOI: 10.1088/1475-7516/2013/08/036. URL: http://dx.doi.org/10.1088/1475-7516/2013/08/036.

[273] Nhat-Minh Nguyen et al. "Impacts of the physical data model on the forward inference of initial conditions from biased tracers". In: *Journal of Cosmology and Astroparticle Physics* 2021.3, 058 (Mar. 2021), p. 058. DOI: 10.1088/1475-7516/2021/03/058. arXiv: 2011.06587 [astro-ph.CO].

[274] Nhat-Minh Nguyen et al. "Taking measurements of the kinematic Sunyaev-Zel'dovich effect forward: including uncertainties from velocity reconstruction with forward modeling". In: *JCAP* 12 (2020), p. 011. DOI: 10.1088/1475-7516/2020/12/011. arXiv: 2007.13721 [astro-ph.CO].

[275] Adi Nusser, Enzo Branchini, and Marc Davis. "Gaia: A Window to Large-scale Motions". In: *Astrophys. J.* 755.1, 58 (Aug. 2012), p. 58. DOI: 10.1088/0004-637X/755/1/58. arXiv: 1202.4138 [astro-ph.CO].

[276] Nikhil Padmanabhan et al. "A 2 per cent distance toz= 0.35 by reconstructing baryon acoustic oscillations – I. Methods and application to the Sloan Digital Sky Survey". In: *Monthly Notices of the Royal Astronomical Society* 427.3 (Nov. 2012), pp. 2132–2145. ISSN: 1365-2966. DOI: 10.1111/j.1365-2966.2012.21888.x. URL: http://dx.doi.org/10.1111/j.1365-2966.2012.21888.x.

[277] Enrique Paillas et al. "Baryon effects on void statistics in the EAGLE simulation". In: *Monthly Notices of the Royal Astronomical Society* 470.4 (Oct. 2017), pp. 4434–4452. DOI: 10.1093/mnras/stx1514. arXiv: 1609.00101 [astro-ph.CO].

[278] Nathalie Palanque-Delabrouille et al. "Constraint on neutrino masses from SDSS-III/BOSS Lyα forest and other cosmological probes". In: *Journal of Cosmology and Astroparticle Physics* 2015.02 (Feb. 2015), pp. 045–045. ISSN: 1475-7516. DOI: 10.1088/1475-7516/2015/02/045. URL: http://dx.doi.org/10.1088/1475-7516/2015/02/045.

[279] Nathalie Palanque-Delabrouille et al. "Neutrino masses and cosmology with Lyman-alpha forest power spectrum". In: *Journal of Cosmology and Astroparticle Physics* 2015.11 (Nov. 2015), pp. 011–011. ISSN: 1475-7516. DOI: 10.1088/1475-7516/2015/11/011. URL: http://dx.doi.org/10.1088/1475-7516/2015/11/011.

[280] A. Palmese and A. G. Kim. "Probing gravity and growth of structure with gravitational waves and galaxies' peculiar velocity". In: *Phys. Rev. D* 103.10, 103507 (May 2021), p. 103507. DOI: 10.1103/PhysRevD.103.103507. arXiv: 2005.04325 [astro-ph.CO].

[281] S. Pandey et al. "Dark Energy Survey year 3 results: Constraints on cosmological parameters and galaxy-bias models from galaxy clustering and galaxy-galaxy lensing using the redMaGiC sample". In: *Phys. Rev. D* 106.4, 043520 (Aug. 2022), p. 043520. DOI: 10.1103/PhysRevD.106.043520. arXiv: 2105.13545 [astro-ph.CO].

[282] Aseem Paranjape, Oliver Hahn, and Ravi K. Sheth. "The dependence of galaxy clustering on tidal environment in the Sloan Digital Sky Survey". In: *Mon. Not. R. Astron. Soc.* 476.4 (June 2018), pp. 5442–5452. DOI: 10.1093/mnras/sty633. arXiv: 1801.04568 [astro-ph.GA].

[283] Christian Partmann et al. *Fast simulations of cosmic large-scale structure with massive neutrinos*. 2020. arXiv: 2003.07387 [astro-ph.CO].

[284] Christian Pedersen et al. "Massive neutrinos and degeneracies in Lyman-alpha forest simulations". In: *Journal of Cosmology and Astroparticle Physics* 2020.04 (Apr. 2020), pp. 025–025. ISSN: 1475-7516. DOI: 10.1088/1475-7516/2020/04/025. URL: http://dx.doi.org/10.1088/1475-7516/2020/04/025.

[285] P. J. E. Peebles. *Principles of Physical Cosmology*. Princeton University Press, 1993.

[286] P. J. E. Peebles. *The large-scale structure of the universe*. Princeton University Press, 1980.

[287] Oliver H. E. Philcox, David N. Spergel, and Francisco Villaescusa-Navarro. "Effective halo model: Creating a physical and accurate model of the matter power spectrum and cluster counts". In: *Phys. Rev. D* 101.12, 123520 (June 2020), p. 123520. DOI: 10.1103/PhysRevD.101.123520. arXiv: 2004.09515 [astro-ph.CO].

[288] Alice Pisani et al. "Counting voids to probe dark energy". In: *Phys. Rev. D* 92 (8 Oct. 2015), p. 083531. DOI: 10.1103/PhysRevD.92.083531. URL: https://link.aps.org/doi/10.1103/PhysRevD.92.083531.

[289] Planck Collaboration et al. "Planck 2013 results. I. Overview of products and scientific results". In: *Astronomy & Astrophysics* 571, A1 (Nov. 2014), A1. DOI: 10.1051/0004-6361/201321529. arXiv: 1303.5062 [astro-ph.CO].

[290] Planck Collaboration et al. "Planck 2013 results. XV. CMB power spectra and likelihood". In: *Astronomy & Astrophysics* 571, A15 (Nov. 2014), A15. DOI: 10.1051/0004-6361/201321573. arXiv: 1303.5075 [astro-ph.CO].

[291] Planck Collaboration et al. "Planck 2013 results. XXIV. Constraints on primordial non-Gaussianity". In: *Astronomy & Astrophysics* 571, A24 (Nov. 2014), A24. DOI: 10.1051/0004-6361/201321554. arXiv: 1303.5084 [astro-ph.CO].

[292] Planck Collaboration et al. "Planck 2018 results. VI. Cosmological parameters". In: *ArXiv e-prints* (July 2018). arXiv: 1807.06209.

[293] E. Platen, R. Van De Weygaert, and B. J. T. Jones. "A cosmic watershed: the WVF void detection technique". In: *Monthly Notices of the Royal Astronomical Society* 380.2 (Sept. 2007), pp. 551–570. ISSN: 1365-2966. DOI: 10.1111/j.1365-2966.2007.12125.x. URL: http://dx.doi.org/10.1111/j.1365-2966.2007.12125.x.

[294] Erwin Platen, Rien van de Weygaert, and Bernard J. T. Jones. "Alignment of voids in the cosmic web". In: *Monthly Notices of the Royal Astronomical Society* 387.1 (June 2008), pp. 128–136. ISSN: 1365-2966. DOI: 10.1111/j.1365-2966.2008.13019.x. URL: http://dx.doi.org/10.1111/j.1365-2966.2008.13019.x.

[295] G Pollina et al. "On the relative bias of void tracers in the Dark Energy Survey". In: *Monthly Notices of the Royal Astronomical Society* 487.2 (May 2019), pp. 2836–2852. ISSN: 1365-2966. DOI: 10.1093/mnras/stz1470. URL: http://dx.doi.org/10.1093/mnras/stz1470.

[296] Natalia Porqueres et al. "Inferring high-redshift large-scale structure dynamics from the Lyman-α forest". In: *Astron. Astrophys.* 630, A151 (Oct. 2019), A151. DOI: 10.1051/0004-6361/201936245. arXiv: 1907.02973 [astro-ph.CO].

[297] James Prideaux-Ghee et al. "Field-Based Physical Inference From Peculiar Velocity Tracers". In: *arXiv e-prints*, arXiv:2204.00023 (Mar. 2022), arXiv:2204.00023. arXiv: 2204.00023 [astro-ph.CO].

[298] Michael Proschan and Myron Waclawiw. "Practical Guidelines for Multiplicity Adjustment in Clinical Trials". In: *Controlled clinical trials* 21 (Jan. 2001), pp. 527–39. DOI: 10.1016/S0197-2456(00)00106-9.

[299] Roland de Putter, Oliver Zahn, and Eric V. Linder. "CMB lensing constraints on neutrinos and dark energy". In: *Physical Review D* 79.6 (Mar. 2009). ISSN: 1550-2368. DOI: 10.1103/physrevd.79.065033. URL: http://dx.doi.org/10.1103/PhysRevD.79.065033.

[300] Thomas Quinn et al. *Time stepping N-body simulations*. 1997. arXiv: astro-ph/9710043 [astro-ph].

[301] Gioacchino Ranucci and Marco Rovere. "Periodogram and likelihood periodicity search in the SNO solar neutrino data". In: *Phys. Rev. D* 75 (1 Jan. 2007), p. 013010. DOI: 10.1103/PhysRevD.75.013010. URL: https://link.aps.org/doi/10.1103/PhysRevD.75.013010.

[302] Annika Reinert and Martin Wolfgang Winkler. "A precision search for WIMPs with charged cosmic rays". In: *Journal of Cosmology and Astroparticle Physics* 2018.01 (Jan. 2018), pp. 055–055. DOI: 10.1088/1475-7516/2018/01/055. URL: https://doi.org/10.1088%5C%2F1475-7516%5C%2F2018%5C%2F01%5C%2F055.

[303] S. O. Rice. "Mathematical Analysis of Random Noise". In: *Bell System Technical Journal* 24.1 (1945), pp. 46–156. DOI: https://doi.org/10.1002/j.1538-7305.1945.tb00453.x. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/j.1538-7305.1945.tb00453.x. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/j.1538-7305.1945.tb00453.x.

[304] A. G. Riess. "Peculiar Velocities from Type Ia Supernovae". In: *Cosmic Flows Workshop*. Ed. by Stephane Courteau and Jeffrey Willick. Vol. 201. Astronomical Society of the Pacific Conference Series. Jan. 2000, p. 80. arXiv: astro-ph/9908237 [astro-ph].

[305] Adam G. Riess et al. "Large Magellanic Cloud Cepheid Standards Provide a 1% Foundation for the Determination of the Hubble Constant and Stronger Evidence for Physics beyond ΛCDM". In: *Astrophys. J.* 876.1, 85 (May 2019), p. 85. DOI: 10.3847/1538-4357/ab1422. arXiv: 1903.07603 [astro-ph.CO].

[306] Christopher D. Rimes and Andrew J. S. Hamilton. "Information content of the non-linear matter power spectrum". In: *Monthly Notices of the Royal Astronomical Society* 360 (June 2005), pp. L82–L86. DOI: 10.1111/j.1745-3933.2005.00051.x. arXiv: astro-ph/0502081 [astro-ph].

[307] Jakob Robnik and Uroš Seljak. "Kepler Data Analysis: Non-Gaussian Noise and Fourier Gaussian Process Analysis of Stellar Variability". In: *The Astronomical Journal* 159.5 (Apr. 2020), p. 224. DOI: 10.3847/1538-3881/ab8460. URL: https://doi.org/10.3847%5C%2F1538-3881%5C%2Fab8460.

[308] Jakob Robnik and Uroš Seljak. "Matched filtering with non-Gaussian noise for planet transit detections". In: *arXiv preprint arXiv:2010.03470* (2020).

[309] M. Roncarelli, F. Villaescusa-Navarro, and M. Baldi. "The kinematic Sunyaev–Zel'dovich effect of the large-scale structure (I): dependence on neutrino mass". In: *Monthly Notices of the Royal Astronomical Society* (Jan. 2017), stx170. ISSN: 1365-2966. DOI: 10.1093/mnras/stx170. URL: http://dx.doi.org/10.1093/mnras/stx170.

[310] Martin Sahlén. "Cluster-void degeneracy breaking: Neutrino properties and dark energy". In: *Physical Review D* 99.6 (Mar. 2019). ISSN: 2470-0029. DOI: 10.1103/physrevd.99.063525. URL: http://dx.doi.org/10.1103/PhysRevD.99.063525.

[311] Shun Saito, Masahiro Takada, and Atsushi Taruya. "Impact of Massive Neutrinos on the Nonlinear Matter Power Spectrum". In: *Physical Review Letters* 100.19 (May 2008). ISSN: 1079-7114. DOI: 10.1103/physrevlett.100.191301. URL: http://dx.doi.org/10.1103/PhysRevLett.100.191301.

[312] P. F. de Salas et al. "Status of neutrino oscillations 2018: 3σ hint for normal mass ordering and improved CP sensitivity". In: *Phys. Lett.* B782 (2018), pp. 633–640. DOI: 10.1016/j.physletb.2018.06.019. arXiv: 1708.01186 [hep-ph].

[313] Till Sawala et al. "Setting the stage: structures from Gaussian random fields". In: *Mon. Not. R. Astron. Soc.* 501.4 (Mar. 2021), pp. 4759–4776. DOI: 10.1093/mnras/staa3568. arXiv: 2003.04321 [astro-ph.CO].

[314] J. D. Scargle. "Studies in astronomical time series analysis. II. Statistical aspects of spectral analysis of unevenly spaced data." In: *Astrophys. J.* 263 (Dec. 1982), pp. 835–853. DOI: 10.1086/160554.

[315] Emmanuel Schaan, Masahiro Takada, and David N. Spergel. "Joint likelihood function of cluster counts and n -point correlation functions: Improving their power through including halo sample variance". In: *Phys. Rev. D* 90, 123523 (Dec. 2014), p. 123523. DOI: 10.1103/PhysRevD.90.123523. arXiv: 1406.3330 [astro-ph.CO].

[316] Andreas S. Schmidt et al. "Cosmological N-body simulations with a large-scale tidal field". In: *Mon. Not. R. Astron. Soc.* 479.1 (Sept. 2018), pp. 162–170. DOI: 10.1093/mnras/sty1430. arXiv: 1803.03274 [astro-ph.CO].

[317] Fabian Schmidt. "An n-th order Lagrangian forward model for large-scale structure". In: *Journal of Cosmology and Astroparticle Physics* 2021.4, 033 (Apr. 2021), p. 033. DOI: 10.1088/1475-7516/2021/04/033. arXiv: 2012.09837 [astro-ph.CO].

[318] Fabian Schmidt. "Sigma-eight at the percent level: the EFT likelihood in real space". In: *Journal of Cosmology and Astroparticle Physics* 2021.4, 032 (Apr. 2021), p. 032. DOI: 10.1088/1475-7516/2021/04/032. arXiv: 2009.14176 [astro-ph.CO].

[319] Marcel Schmittfull and Uroš Seljak. "Parameter constraints from cross-correlation of CMB lensing with galaxy clustering". In: *Physical Review D* 97.12 (June 2018). ISSN: 2470-0029. DOI: 10.1103/physrevd.97.123540. URL: http://dx.doi.org/10.1103/PhysRevD.97.123540.

[320] Marcel Schmittfull et al. "Modeling galaxies in redshift space at the field level". In: *Journal of Cosmology and Astroparticle Physics* 2021.5, 059 (May 2021), p. 059. DOI: 10.1088/1475-7516/2021/05/059. arXiv: 2012.03334 [astro-ph.CO].

[321] Román Scoccimarro. "The Bispectrum: From Theory to Observations". In: *Astrophys. J.* 544.2 (Dec. 2000), pp. 597–615. DOI: 10.1086/317248. arXiv: astro-ph/0004086 [astro-ph].

[322] E. Sefusatti et al. "Accurate estimators of correlation functions in Fourier space". In: *Mon. Not. R. Astron. Soc.* 460.4 (Aug. 2016), pp. 3624–3636. DOI: 10.1093/mnras/stw1229. arXiv: 1512.07295 [astro-ph.CO].

[323] Emiliano Sefusatti et al. "Cosmology and the bispectrum". In: *Phys. Rev. D* 74, 023522 (July 2006), p. 023522. DOI: 10.1103/PhysRevD.74.023522. arXiv: astro-ph/0604505 [astro-ph].

[324] Norio Sekiya, Noriko Y. Yamasaki, and Kazuhisa Mitsuda. "A search for a keV signature of radiatively decaying dark matter with Suzaku XIS observations of the X-ray diffuse background". In: *Publications of the Astronomical Society of Japan* 68.SP1 (Sept. 2015). S31. ISSN: 0004-6264. DOI: 10.1093/pasj/psv081. eprint: https://academic.oup.com/pasj/article-pdf/68/SP1/S31/7971976/psv081.pdf. URL: https://doi.org/10.1093/pasj/psv081.

[325] Uroš Seljak. "Cosmography and Power Spectrum Estimation: A Unified Approach". In: *Astrophys. J.* 503.2 (Aug. 1998), pp. 492–501. DOI: 10.1086/306019. arXiv: astro-ph/9710269 [astro-ph].

[326] Uroš Seljak. "Extracting Primordial Non-Gaussianity without Cosmic Variance". In: *Phys. Rev. Lett.* 102 (2 Jan. 2009), p. 021302. DOI: 10.1103/PhysRevLett.102.021302. URL: https://link.aps.org/doi/10.1103/PhysRevLett.102.021302.

[327] Uros Seljak and Byeonghee Yu. *Posterior inference unchained with* EL_2O. 2019. arXiv: 1901.04454 [stat.ML].

[328] Uroš Seljak et al. "Cosmological parameter analysis including SDSS Lyα forest and galaxy bias: Constraints on the primordial spectrum of fluctuations, neutrino mass, and dark energy". In: *Phys. Rev. D* 71 (10 May 2005), p. 103515. DOI: 10.1103/PhysRevD.71.103515. URL: https://link.aps.org/doi/10.1103/PhysRevD.71.103515.

[329] Uroš Seljak et al. "Towards optimal extraction of cosmological information from nonlinear data". In: *Journal of Cosmology and Astroparticle Physics* 2017.12 (Dec. 2017), pp. 009–009. ISSN: 1475-7516. DOI: 10.1088/1475-7516/2017/12/009. URL: http://dx.doi.org/10.1088/1475-7516/2017/12/009.

[330] Leonardo Senatore and Matias Zaldarriaga. *The Effective Field Theory of Large-Scale Structure in the presence of Massive Neutrinos*. 2017. arXiv: 1707.04698 [astro-ph.CO].

[331] J P Shaffer. "Multiple Hypothesis Testing". In: *Annual Review of Psychology* 46.1 (1995), pp. 561–584. DOI: 10.1146/annurev.ps.46.020195.003021. eprint: https://doi.org/10.1146/annurev.ps.46.020195.003021. URL: https://doi.org/10.1146/annurev.ps.46.020195.003021.

[332] H. Shan et al. "KiDS-450: cosmological constraints from weak lensing peak statistics - I. Inference from analytical prediction of high signal-to-noise ratio convergence peaks". In: *Monthly Notices of the Royal Astronomical Society* 474 (Feb. 2018), pp. 1116–1134. DOI: 10.1093/mnras/stx2837. arXiv: 1709.07651.

[333] Blake D. Sherwin et al. "Two-season Atacama Cosmology Telescope polarimeter lensing power spectrum". In: *Phys. Rev. D* 95 (12 June 2017), p. 123529. DOI: 10.1103/PhysRevD.95.123529. URL: https://link.aps.org/doi/10.1103/PhysRevD.95.123529.

[334] R. K. Sheth and G. Tormen. "An excursion set model of hierarchical clustering: ellipsoidal collapse and the moving barrier". In: *Monthly Notices of the Royal Astronomical Society* 329 (Jan. 2002), pp. 61–75. DOI: 10.1046/j.1365-8711.2002.04950.x. eprint: arXiv:astro-ph/0105113.

[335] Ravi K. Sheth and Giuseppe Tormen. "Large-scale bias and the peak background split". In: *Monthly Notices of the Royal Astronomical Society* 308.1 (Sept. 1999), pp. 119–126. ISSN: 1365-2966. DOI: 10.1046/j.1365-8711.1999.02692.x. URL: http://dx.doi.org/10.1046/j.1365-8711.1999.02692.x.

[336] Jingjing Shi and Ravi K. Sheth. "Dependence of halo bias on mass and environment". In: *Mon. Not. R. Astron. Soc.* 473.2 (Jan. 2018), pp. 2486–2492. DOI: 10.1093/mnras/stx2277. arXiv: 1707.04096 [astro-ph.CO].

[337] Masatoshi Shoji and Eiichiro Komatsu. *Massive Neutrinos in Cosmology: Analytic Solutions and Fluid Approximation*. 2010. arXiv: 1003.0942 [astro-ph.CO].

[338] Zbyněk Šidák. "Rectangular Confidence Regions for the Means of Multivariate Normal Distributions". In: *Journal of the American Statistical Association* 62.318 (1967), pp. 626–633. DOI: 10.1080/01621459.1967.10482935. eprint: https://doi.org/10.1080/01621459.1967.10482935. URL: https://doi.org/10.1080/01621459.1967.10482935.

[339] Fergus Simpson and Sarah Bridle. "Illuminating dark energy with cosmic shear". In: *Phys. Rev. D* 71.8, 083501 (Apr. 2005), p. 083501. DOI: 10.1103/PhysRevD.71.083501. arXiv: astro-ph/0411673 [astro-ph].

[340] Edwin Sirko. "Initial Conditions to Cosmological N-Body Simulations, or, How to Run an Ensemble of Simulations". In: *Astrophys. J.* 634.2 (Nov. 2005), pp. 728–743. DOI: 10.1086/497090. arXiv: astro-ph/0503106 [astro-ph].

[341] Kendrick M. Smith et al. "KSZ tomography and the bispectrum". In: *arXiv e-prints*, arXiv:1810.13423 (Oct. 2018), arXiv:1810.13423. arXiv: 1810.13423 [astro-ph.CO].

[342] R. E. Smith et al. "Stable clustering, the halo model and non-linear cosmological power spectra". In: *Monthly Notices of the Royal Astronomical Society* 341.4 (June 2003), pp. 1311–1332. ISSN: 1365-2966. DOI: 10.1046/j.1365-8711.2003.06503.x. URL: http://dx.doi.org/10.1046/j.1365-8711.2003.06503.x.

[343] I.M. Soboĺ. "Quasi-Monte Carlo methods". In: *Progress in Nuclear Energy* 24.1 (1990). Monte Carlo Methods for Neutrons and Photon Transport Calculations, pp. 55–61. ISSN: 0149-1970. DOI: https://doi.org/10.1016/0149-1970(90)90022-W. URL: http://www.sciencedirect.com/science/article/pii/014919709090022W.

[344] Jenny G. Sorce et al. "Cosmicflows Constrained Local UniversE Simulations". In: *Mon. Not. R. Astron. Soc.* 455.2 (Jan. 2016), pp. 2078–2090. DOI: 10.1093/mnras/stv2407. arXiv: 1510.04900 [astro-ph.CO].

[345] D. Spergel et al. *Wide-Field InfraRed Survey Telescope-Astrophysics Focused Telescope Assets WFIRST-AFTA Final Report*. 2013. arXiv: 1305.5422 [astro-ph.IM].

[346] Volker Springel. "The cosmological simulation code gadget-2". In: *Monthly Notices of the Royal Astronomical Society* 364.4 (Dec. 2005), pp. 1105–1134. ISSN: 1365-2966. DOI: 10.1111/j.1365-2966.2005.09655.x. URL: http://dx.doi.org/10.1111/j.1365-2966.2005.09655.x.

[347] Benjamin E. Stahl et al. "Peculiar-velocity cosmology with Types Ia and II supernovae". In: *Mon. Not. R. Astron. Soc.* 505.2 (Aug. 2021), pp. 2349–2360. DOI: 10.1093/mnras/stab1446. arXiv: 2105.05185 [astro-ph.CO].

[348] John D Storey and Robert Tibshirani. "Statistical significance for genomewide studies". In: *Proceedings of the National Academy of Sciences of the United States of America* 100.16 (Aug. 2003), pp. 9440–9445. DOI: 10.1073/pnas.1530509100. URL: https://pubmed.ncbi.nlm.nih.gov/12883005.

[349] P. M. Sutter et al. "A FIRST APPLICATION OF THE ALCOCK-PACZYNSKI TEST TO STACKED COSMIC VOIDS". In: *The Astrophysical Journal* 761.2 (Dec. 2012), p. 187. ISSN: 1538-4357. DOI: 10.1088/0004-637x/761/2/187. URL: http://dx.doi.org/10.1088/0004-637X/761/2/187.

[350] P. M. Sutter et al. *VIDE: The Void IDentification and Examination toolkit*. 2014. arXiv: 1406.1191 [astro-ph.CO].

[351] Richard Swinbank and R. James Purser. "Fibonacci grids: A novel approach to global modelling". In: *Quarterly Journal of the Royal Meteorological Society* 132.619 (2006), pp. 1769–1793. DOI: 10.1256/qj.05.227. eprint: https://rmets.onlinelibrary.wiley.com/doi/pdf/10.1256/qj.05.227. URL: https://rmets.onlinelibrary.wiley.com/doi/abs/10.1256/qj.05.227.

[352] Masahiro Takada and Wayne Hu. "Power spectrum super-sample covariance". In: *Phys. Rev. D* 87.12, 123504 (June 2013), p. 123504. DOI: 10.1103/PhysRevD.87.123504. arXiv: 1302.6994 [astro-ph.CO].

[353] Masahiro Takada and Bhuvnesh Jain. "Cosmological parameters from lensing power spectrum and bispectrum tomography". In: *Monthly Notices of the Royal Astronomical Society* 348 (Mar. 2004), pp. 897–915. DOI: 10.1111/j.1365-2966.2004.07410.x. arXiv: astro-ph/0310125 [astro-ph].

[354] Ryuichi Takahashi et al. "REVISING THE HALOFIT MODEL FOR THE NONLINEAR MATTER POWER SPECTRUM". In: *The Astrophysical Journal* 761.2 (Dec. 2012), p. 152. ISSN: 1538-4357. DOI: 10.1088/0004-637x/761/2/152. URL: http://dx.doi.org/10.1088/0004-637X/761/2/152.

[355] Ryuichi Takahashi et al. "Simulations of Baryon Acoustic Oscillations. II. Covariance Matrix of the Matter Power Spectrum". In: *Astrophys. J.* 700.1 (July 2009), pp. 479–490. DOI: 10.1088/0004-637X/700/1/479. arXiv: 0902.0371 [astro-ph.CO].

[356] Svetlin Tassev, Matias Zaldarriaga, and Daniel J Eisenstein. "Solving large scale structure in ten easy steps with COLA". In: *Journal of Cosmology and Astroparticle Physics* 2013.06 (June 2013), pp. 036–036. ISSN: 1475-7516. DOI: 10.1088/1475-7516/2013/06/036. URL: http://dx.doi.org/10.1088/1475-7516/2013/06/036.

[357] Max Tegmark. "How to measure CMB power spectra without losing information". In: *Phys. Rev. D* 55.10 (May 1997), pp. 5895–5907. DOI: 10.1103/PhysRevD.55.5895. arXiv: astro-ph/9611174 [astro-ph].

[358] Max Tegmark, Andy N. Taylor, and Alan F. Heavens. "Karhunen-Loève Eigenvalue Problems in Cosmology: How Should We Tackle Large Data Sets?" In: *The Astrophysical Journal* 480 (May 1997), pp. 22–35. DOI: 10.1086/303939. arXiv: astro-ph/9603021 [astro-ph].

[359] Ryo Terasawa et al. "Separate universe approach to evaluate nonlinear matter power spectrum for nonflat Λ CDM model". In: *Phys. Rev. D* 106.8, 083504 (Oct. 2022), p. 083504. DOI: 10.1103/PhysRevD.106.083504. arXiv: 2205.10339 [astro-ph.CO].

[360] The Simons Observatory Collaboration et al. "The Simons Observatory: Science goals and forecasts". In: *ArXiv e-prints* (Aug. 2018). arXiv: 1808.07445.

[361] Jeremy Tinker et al. "Toward a Halo Mass Function for Precision Cosmology: The Limits of Universality". In: *The Astrophysical Journal* 688.2 (Dec. 2008), pp. 709–728. ISSN: 1538-4357. DOI: 10.1086/591439. URL: http://dx.doi.org/10.1086/591439.

[362] C. To et al. "Dark Energy Survey Year 1 Results: Cosmological Constraints from Cluster Abundances, Weak Lensing, and Galaxy Correlations". In: *Phys. Rev. Lett.* 126.14, 141301 (Apr. 2021), p. 141301. DOI: 10.1103/PhysRevLett.126.141301. arXiv: 2010.01138 [astro-ph.CO].

[363] M. A. Troxel et al. "Dark Energy Survey Year 1 results: Cosmological constraints from cosmic shear". In: *Phys. Rev. D* 98 (4 Aug. 2018), p. 043528. DOI: 10.1103/PhysRevD.98.043528. URL: https://link.aps.org/doi/10.1103/PhysRevD.98.043528.

[364] R. Brent Tully et al. "Our Peculiar Motion Away from the Local Void". In: *Astrophys. J.* 676.1 (Mar. 2008), pp. 184–205. DOI: 10.1086/527428. arXiv: 0705.4139 [astro-ph].

[365] R. Brent Tully et al. "The Laniakea supercluster of galaxies". In: *Nature* 513.7516 (Sept. 2014), pp. 71–73. DOI: 10.1038/nature13674. arXiv: 1409.0880 [astro-ph.CO].

[366] Cora Uhlemann et al. "Fisher for complements: extracting cosmology and neutrino mass from the counts-in-cells PDF". In: *Monthly Notices of the Royal Astronomical Society* 495.4 (May 2020), pp. 4006–4027. DOI: 10.1093/mnras/staa1155. arXiv: 1911.11158 [astro-ph.CO].

[367] Cora Uhlemann et al. "It takes two to know one: Computing accurate one-point PDF covariances from effective two-point PDF models". In: *arXiv e-prints*, arXiv:2210.07819 (Oct. 2022), arXiv:2210.07819. arXiv: 2210.07819 [astro-ph.CO].

[368] Amol Upadhye et al. "Redshift-space distortions in massive neutrino and evolving dark energy cosmologies". In: *Physical Review D* 93.6 (Mar. 2016). ISSN: 2470-0029. DOI: 10.1103/physrevd.93.063515. URL: http://dx.doi.org/10.1103/PhysRevD.93.063515.

[369] Eleonora Di Valentino, Alessandro Melchiorri, and Joseph Silk. "Cosmological constraints in extended parameter space from the Planck 2018 Legacy release". In: *Journal of Cosmology and Astroparticle Physics* 2020.01 (Jan. 2020), pp. 013–013. ISSN: 1475-7516. DOI: 10.1088/1475-7516/2020/01/013. URL: http://dx.doi.org/10.1088/1475-7516/2020/01/013.

[370] Georgios Valogiannis and Cora Dvorkin. "Going Beyond the Galaxy Power Spectrum: an Analysis of BOSS Data with Wavelet Scattering Transforms". In: *arXiv e-prints*, arXiv:2204.13717 (Apr. 2022), arXiv:2204.13717. arXiv: 2204.13717 [astro-ph.CO].

[371] Georgios Valogiannis and Cora Dvorkin. "Towards an Optimal Estimation of Cosmological Parameters with the Wavelet Scattering Transform". In: *arXiv e-prints*, arXiv:2108.07821 (Aug. 2021), arXiv:2108.07821. arXiv: 2108.07821 [astro-ph.CO].

[372] Jacob T. VanderPlas. "Understanding the Lomb–Scargle Periodogram". In: *The Astrophysical Journal Supplement Series* 236.1 (May 2018), p. 16. ISSN: 1538-4365. DOI: 10.3847/1538-4365/aab766. URL: http://dx.doi.org/10.3847/1538-4365/aab766.

[373] L. Verde. "Statistical Methods in Cosmology". In: *Lecture Notes in Physics* (2010), pp. 147–177. ISSN: 1616-6361. DOI: 10.1007/978-3-642-10598-2_4. URL: http://dx.doi.org/10.1007/978-3-642-10598-2_4.

[374] Giovanni Verza et al. "The void size function in dynamical dark energy cosmologies". In: *Journal of Cosmology and Astroparticle Physics* 2019.12 (Dec. 2019), pp. 040–040. ISSN: 1475-7516. DOI: 10.1088/1475-7516/2019/12/040. URL: http://dx.doi.org/10.1088/1475-7516/2019/12/040.

[375] Matteo Viel, Martin G Haehnelt, and Volker Springel. "The effect of neutrinos on the matter distribution as probed by the intergalactic medium". In: *Journal of Cosmology and Astroparticle Physics* 2010.06 (June 2010), pp. 015–015. ISSN: 1475-7516. DOI: 10.1088/1475-7516/2010/06/015. URL: http://dx.doi.org/10.1088/1475-7516/2010/06/015.

[376] F. Villaescusa-Navarro et al. "Neutrino halos in clusters of galaxies and their weak lensing signature". In: *Journal of Cosmology and Astroparticle Physics* 6, 027 (June 2011), p. 027. DOI: 10.1088/1475-7516/2011/06/027. arXiv: 1104.4770.

[377] F. Villaescusa-Navarro et al. "Non-linear evolution of the cosmic neutrino background". In: *Journal of Cosmology and Astroparticle Physics* 3, 019 (Mar. 2013), p. 019. DOI: 10.1088/1475-7516/2013/03/019. arXiv: 1212.4855.

[378] Francisco Villaescusa-Navarro. *Pylians: Python libraries for the analysis of numerical simulations*. Astrophysics Source Code Library, record ascl:1811.008. Nov. 2018. ascl: 1811.008.

[379] Francisco Villaescusa-Navarro, Philip Bull, and Matteo Viel. "WEIGHING NEUTRINOS WITH COSMIC NEUTRAL HYDROGEN". In: *The Astrophysical Journal* 814.2 (Nov. 2015), p. 146. ISSN: 1538-4357. DOI: 10.1088/0004-637x/814/2/146. URL: http://dx.doi.org/10.1088/0004-637X/814/2/146.

[380] Francisco Villaescusa-Navarro et al. "Cosmology with massive neutrinos I: towards a realistic modeling of the relation between matter, haloes and galaxies". In: *Journal of Cosmology and Astroparticle Physics* 2014.03 (Mar. 2014), pp. 011–011. ISSN: 1475-7516. DOI: 10.1088/1475-7516/2014/03/011. URL: http://dx.doi.org/10.1088/1475-7516/2014/03/011.

[381] Francisco Villaescusa-Navarro et al. "The CAMELS project: Cosmology and Astrophysics with MachinE Learning Simulations". In: *arXiv e-prints*, arXiv:2010.00619 (Oct. 2020), arXiv:2010.00619. arXiv: 2010.00619 [astro-ph.CO].

[382] Francisco Villaescusa-Navarro et al. "The Imprint of Neutrinos on Clustering in Redshift Space". In: *The Astrophysical Journal* 861.1 (July 2018), p. 53. ISSN: 1538-4357. DOI: 10.3847/1538-4357/aac6bf. URL: http://dx.doi.org/10.3847/1538-4357/aac6bf.

[383] Francisco Villaescusa-Navarro et al. *The Quijote simulations*. 2019. arXiv: 1909.05273 [astro-ph.CO].

[384] C. Wagner et al. "Separate universe simulations." In: *Mon. Not. R. Astron. Soc.* 448 (Mar. 2015), pp. L11–L15. DOI: 10.1093/mnrasl/slu187. arXiv: 1409.6294 [astro-ph.CO].

[385] Duncan J. Watts et al. "A Projected Estimate of the Reionization Optical Depth Using the CLASS Experiment's Sample Variance Limited E-mode Measurement". In: *The Astrophysical Journal* 863.2 (Aug. 2018), p. 121. DOI: 10.3847/1538-4357/aad283. URL: https://doi.org/10.3847/1538-4357/aad283.

[386] Risa H. Wechsler et al. "The Dependence of Halo Clustering on Halo Formation History, Concentration, and Occupation". In: *The Astrophysical Journal* 652.1 (Nov. 2006), pp. 71–84. ISSN: 1538-4357. DOI: 10.1086/507120. URL: http://dx.doi.org/10.1086/507120.

[387] M. White. "The Zel'dovich approximation". In: *Mon. Not. R. Astron. Soc.* 439 (Apr. 2014), pp. 3630–3640. DOI: 10.1093/mnras/stu209. arXiv: 1401.5466.

[388] Benjamin D Wibking et al. "Cosmology with galaxy–galaxy lensing on non-perturbative scales: emulation method and application to BOSS LOWZ". In: *Monthly Notices of the Royal Astronomical Society* 492.2 (Dec. 2019), pp. 2872–2896. ISSN: 1365-2966. DOI: 10.1093/mnras/stz3423. URL: http://dx.doi.org/10.1093/mnras/stz3423.

[389] S. S. Wilks. "The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses". In: *Ann. Math. Statist.* 9.1 (Mar. 1938), pp. 60–62. DOI: 10.1214/aoms/1177732360. URL: https://doi.org/10.1214/aoms/1177732360.

[390] Jeffrey A. Willick and Michael A. Strauss. "Maximum Likelihood Comparison of Tully-Fisher and Redshift Data. II. Results from an Expanded Sample". In: *Astrophys. J.* 507.1 (Nov. 1998), pp. 64–83. DOI: 10.1086/306314. arXiv: astro-ph/9801307 [astro-ph].

[391] Jeffrey A. Willick et al. "Maximum Likelihood Comparisons of Tully-Fisher and Redshift Data: Constraints on Ω and Biasing". In: *Astrophys. J.* 486.2 (Sept. 1997), pp. 629–664. DOI: 10.1086/304551. arXiv: astro-ph/9612240 [astro-ph].

[392] Bill S. Wright, Hans A. Winther, and Kazuya Koyama. "COLA with massive neutrinos". In: *Journal of Cosmology and Astroparticle Physics* 2017.10 (Oct. 2017), pp. 054–054. ISSN: 1475-7516. DOI: 10.1088/1475-7516/2017/10/054. URL: http://dx.doi.org/10.1088/1475-7516/2017/10/054.

[393] S. Yahia-Cherif et al. "Validating the Fisher approach for stage IV spectroscopic surveys". In: *Astronomy & Astrophysics* 649 (May 2021), A52. ISSN: 1432-0746. DOI: 10.1051/0004-6361/201937312. URL: http://dx.doi.org/10.1051/0004-6361/201937312.

[394] Christophe Yèche et al. "Constraints on neutrino masses from Lyman-alpha forest power spectrum with BOSS and XQ-100". In: *Journal of Cosmology and Astroparticle Physics* 2017.06 (June 2017), pp. 047–047. ISSN: 1475-7516. DOI: 10.1088/1475-7516/2017/06/047. URL: http://dx.doi.org/10.1088/1475-7516/2017/06/047.

[395] Byeonghee Yu et al. *Towards Neutrino Mass from Cosmology without Optical Depth Information*. 2018. arXiv: 1809.02120 [astro-ph.CO].

[396] Hao-Ran Yu et al. "Differential neutrino condensation onto cosmic structure". In: *Nature Astronomy* 1.7 (June 2017). ISSN: 2397-3366. DOI: 10.1038/s41550-017-0143. URL: http://dx.doi.org/10.1038/s41550-017-0143.

[397] Saleem Zaroubi, Yehuda Hoffman, and Avishai Dekel. "Wiener Reconstruction of Large-Scale Structure from Peculiar Velocities". In: *Astrophys. J.* 520.2 (Aug. 1999), pp. 413–425. DOI: 10.1086/307473. arXiv: astro-ph/9810279 [astro-ph].

[398] Y. B. Zel'dovich. "Gravitational instability: An approximate theory for large density perturbations." In: *Astron. Astrophys.* 5 (Mar. 1970), pp. 84–89.

[399] M. Zennaro et al. "Initial conditions for accurate N-body simulations of massive neutrino cosmologies". In: *Monthly Notices of the Royal Astronomical Society* 466 (Apr. 2017), pp. 3244–3258. DOI: 10.1093/mnras/stw3340. arXiv: 1605.05283.

[400] Matteo Zennaro et al. "How to add massive neutrinos to your ΛCDM simulation – extending cosmology rescaling algorithms". In: *Monthly Notices of the Royal Astronomical Society* 489.4 (Sept. 2019), pp. 5938–5951. ISSN: 1365-2966. DOI: 10.1093/mnras/stz2612. URL: http://dx.doi.org/10.1093/mnras/stz2612.

[401] Zhongxu Zhai et al. "The Aemulus Project. III. Emulation of the Galaxy Correlation Function". In: *The Astrophysical Journal* 874.1 (Mar. 2019), p. 95. ISSN: 1538-4357. DOI: 10.3847/1538-4357/ab0d7b. URL: http://dx.doi.org/10.3847/1538-4357/ab0d7b.

[402] Jing-Fei Zhang, Bo Wang, and Xin Zhang. "Forecast for weighing neutrinos in cosmology with SKA". In: *Science China Physics, Mechanics & Astronomy* 63.8 (Apr. 2020). ISSN: 1869-1927. DOI: 10.1007/s11433-019-1516-y. URL: http://dx.doi.org/10.1007/s11433-019-1516-y.

[403] Hong-Ming Zhu et al. "Reconstruction with velocities". In: *Mon. Not. Roy. Astron. Soc.* 494.3 (2020), pp. 4244–4254. DOI: 10.1093/mnras/staa1002. arXiv: 1910.02318 [astro-ph.CO].

[404] José Manuel Zorrilla Matilla et al. "Geometry and growth contributions to cosmic shear observables". In: *Phys. Rev. D* 96.2, 023513 (July 2017), p. 023513. DOI: 10.1103/PhysRevD.96.023513. arXiv: 1706.05133 [astro-ph.CO].