UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DEPARTAMENTO DE FÍSICA

INFORMED TOTAL-ERROR-MINIMIZING (ITEM) PRIORS: INTERPRETABLE COSMOLOGICAL PARAMETER CONSTRAINTS DESPITE COMPLEX NUISANCE EFFECTS

TESIS PARA OPTAR AL GRADO DE MAGÍSTER EN CIENCIAS, MENCIÓN FÍSICA

BERNARDITA RIED GUACHALLA

PROFESOR GUÍA:
DOMENICO SAPONE

MIEMBROS DE LA COMISIÓN:
DANIEL GRUEN
GONZALO PALMA GUILODRÁN
NELSON PADILLA

Este trabajo ha sido parcialmente financiado por:
FONDECYT Regular - N. 1200171
DAAD (Forschungsstipendien) Kurzstipendien - N. 91818589
ANID (Subdirección de Capital Humano) Magíster Nacional - N. 22210491
Deutsche Forschungsgemeinschaft (Germany's Excellence Strategy) EXC-2094 – N. 390783311

SANTIAGO DE CHILE
2022

RESUMEN DE LA MEMORIA PARA OPTAR
AL TÍTULO DE MAGÍSTER EN CIENCIAS
MENCIÓN EN FÍSICA
POR: BERNARDITA RIED GUACHALLA
FECHA: 2022
PROF. GUÍA: DOMENICO SAPONE

PRIORS INFORMATIVOS QUE MINIMIZAN EL ERROR TOTAL: RESTRICCIONES INTERPRETABLES DE LOS PARÁMETROS COSMOLÓGICOS A PESAR DE LOS COMPLEJOS EFECTOS NUISANCE

La cosmología moderna pretende comprender la física del universo, es decir, los procesos cinemáticos y dinámicos, el contenido, la estructura, el origen y la evolución cósmica. Para comprenderla plenamente, los surveys cosmológicos deben analizarse con los conocimientos estadísticos más avanzados posibles, debido a que los próximos surveys abarcarán todo el cielo. Los conjuntos de datos venideros nos obligarán a considerar modelos físicos de escalas más pequeñas y estadísticas de mayor orden que las habitualmente estudiadas hasta la fecha. Por ello, llamamos a esta época la era de la precisión cosmológica.

Particularmente, cuando los modelos aumentan su complejidad añadiendo parámetros, sin restringir su información previa, surge un desafío en un análisis Bayesiano: el efecto de volumen de prior (o *prior volume* en inglés). Este efecto puede llevar las interpretaciones cosmológicas a supuestos radicalmente erróneos y a tensiones inexistentes. Nuestra propuesta incluye un marco sistemático que deriva distribuciones de priors informadas y minimizadoras del error total de los parámetros nuisance a partir de un conjunto de posibles efectos nuisance.

En este trabajo, presentamos una revisión bibliográfica exhaustiva tanto de la cosmología como de la estadística para sentar las bases de la investigación. Esta incluye los fundamentos de la cosmología y la formación de la estructura cósmica en los capítulos 2 y 3, respectivamente. Las revisiones y discusiones exhaustivas en torno al uso de las probabilidades y la inferencia estadística se incluyen en los capítulos 4 y 5, respectivamente. Incorporamos una revisión del problema del tanque alemán en el capítulo 6 y comentamos su importancia en relación con el análisis cosmológico. El principal resultado científico se muestra en el capítulo 7, donde proponemos un método para corregir el efecto de volumen de prior mediante un análisis de probabilidad simulada. Lo probamos con los modelos de shot-noise del density split statistics [1, 2] con datos del Dark Energy Survey (DES) Y1. Este novedoso método extiende el enfoque del análisis de verosimilitud, no limitado a la cosmología observacional, sino a todas las áreas en las que se tienen en cuenta los análisis Bayesianos.

RESUMEN DE LA MEMORIA PARA OPTAR
AL TÍTULO DE MAGÍSTER EN CIENCIAS
MENCIÓN EN FÍSICA
POR: BERNARDITA RIED GUACHALLA
FECHA: 2022
PROF. GUÍA: DOMENICO SAPONE

INFORMED TOTAL-ERROR-MINIMIZING (ITEM) PRIORS: INTERPRETABLE COSMOLOGICAL PARAMETER CONSTRAINTS DESPITE COMPLEX NUISANCE EFFECTS

Modern physical cosmology aims to understand the physics of the universe, i.e., the kinematical and dynamical processes, the contents, the structure, and the cosmic origin and evolution. To fully understand it, the cosmological probes must be analyzed with the latest statistical knowledge possible, specifically due to the forthcoming entire sky surveys. The upcoming datasets will require us to consider physical models of smaller scales and higher-order statistics than the usually studied in cosmology. Because of that, we call this epoch the era of precision cosmology.

A particular challenge arises when models increase their complexity by adding parameters without constraining their preliminary information for an accurate Bayesian analysis: the prior volume effect. This effect can lead cosmological interpretations to radical wrong assumptions and tensions that might not even exist. Our proposal includes a systematic framework that derives informed and total error minimizing prior distributions to nuisance parameters from a set of possible nuisance effects.

In this work, we present a thorough literature review of both cosmology and statistics to set the investigation's basis. The former includes the basics of physical cosmology and the cosmic structure formation in Chapters 2 and 3, respectively. Comprehensive reviews and discussions around the use of probabilities and statistical inference are included in Chapters 4 and 5, respectively. We incorporate an examination of the German tank problem in Chapter 6 and remark on its importance relative to cosmological analysis. The main scientific result is shown in Chapter 7, where we propose a method to correct the prior volume effect through a simulated likelihood analysis. We test it with the shot-noise models from the density split statistics framework [1, 2] with data from the Dark Energy Survey (DES) Y1. This novel method extends the approach of the likelihood analysis, not limited to observational cosmology, but to all the areas in which Bayesian analyses are taken into account.

Essentially, all models are wrong,
but some are useful.

Box & Draper 1987

Acknowledgments

To educate is to govern was the motto of the president of Chile, Pedro Aguirre Cerda, in 1938. This short but powerful phrase had been in my head when finishing this thesis. I must then acknowledge my profound gratitude to the Chilean and German governments for funding my studies and research without worrying about my roots but for my great efforts. The public institutions that hosted me through the last years, the University of Chile and the University of Munich, allowed me to challenge, develop and think without boundaries. They allowed me to commit myself to the importance of education as a human right.

I want to thank everyone who has taught me something about the world. Especially my supervisor, professor Domenico Sapone, who has believed in me since the first time we worked together and encouraged me to explore the world from an academic perspective. I must also acknowledge professor Gonzalo Palma for showing me the beauty of physics in the second year of my undergraduate studies, making me realize that this field is what I have always been looking for in life.

I can not summarize in words the infinite gratitude I have towards professor Daniel Gruen for both organizing the PAVES program at Stanford University and hosting my research stay that ended up happening in Germany due to the pandemic of COVID19. I also thank Oliver Friedrich for solving my multiple queries and guiding me in hard times of our history. To both of them, I own many of the learnings, ideas, and thoughts presented in this work.

My resting acknowledgments are split into both hemispheres: In the Southern, I must thank all the *Cosmology and Theoretical Astrophysics Group*, as well as the *Centro de Astrofísica y Tecnologías Afines* from the University of Chile for encouraging me to do research and outreach the latest discoveries to the public. I must also mention all the contributors to the social project *Preu en Cuarentena* which helped thousands of high-school students in mathematics in the middle of the chaos of a social revolution and a world pandemic.

In the Northern hemisphere, I must thank all the acquaintances and colleagues I met from September 2021 to March 2022 from North America and Europe during my intense US Ph.D. trip, and stay in Germany, respectively. Especially all the friends I made from the ACAI Group at USM, from which I highlight my office mate and best friend, Elisa.

A page is not enough to note all who made my soul smile through their kindness and loving support. Thanks Emir, Vale, Tirso, Parth, Serge, Erick, Fran, Benja, Felipe, Bruno, Christian, Daniel, Nahir, Martin, Alfa, Joakin, Dylan, Daniil, Jed, Luca, Pepe, Yun-Hsin, Jeremy, Maca, Emi, Nico, Bashin and all the others I have not mentioned, but I am deeply in debt. I finish thanking Chimi for her loyal company throughout the globe.

Table of content

List of Tables

List of Figures

Chapter 1

Introduction

Cosmology is the study of the physical laws that govern the universe. This area of active research aims to describe, through a mathematical language, large-scale observed phenomena such as the homogeneous distribution of the cosmic structures. The standard model of cosmology has successfully predicted findings such as the Cosmic Microwave Background [3, 4] and addressed observations that support an expanding universe [5–7], and more recently, at an accelerated rate [8, 9]. Variations of the theory of a hot Big Bang origin of the cosmos started to be debated among cosmologists [10], being nowadays constantly validated, rejected, or extended through the use of combined probes of cosmology [11].

The intuition relies on tracing a world-line: if the universe is currently expanding, in the past, all the components should have been clustered in a *hot and dense cosmic soup*. The first stages of the young universe would have had primordial interactions that would manifest in today's observations. *Observational cosmology* aims to test these kinds of hypotheses by matching them with the information obtained through telescopes and experiments. Other areas, such as *particle physics*, are fundamental too because they explore different relevant physical scales for early epochs of the universe [12].

Cosmological analyses have experienced enormous advances in the last century due to the exponential growth of data but still tend to be over-simplistic on large scales [10]. In recent decades, telescopes have provided large amounts of data that have decreased the random uncertainties associated with cosmological measurements. Collaborative groups have cooperated to generate surveys that have mapped the sky and traced changes on the lines of sight through time [13].

Statistics allow us to extract information from observations and numerically validate the formulated physical concepts, decreasing the random uncertainties. Nevertheless, with smaller error bars, systematic effects have recently arisen and changed the research focus towards the era of *precision cosmology* due to the rise of statistical disagreements, being *the Hubble tension* one of the prominent examples [14]. This incongruity associated to the Hubble constant when observing the early and late universe could be explained by two non-mutually exclusive solutions: variations of the ΛCDM model and a more thorough statistical approach. Observational cosmology aims to solve this through statistics and machine learning algorithms trained with unprecedented computational power. The combination of physical knowledge and artificial intelligence can help us fill the gaps when the information is hard to disentangle.

The Bayesian theory of probabilities has been widely used among cosmologists to validate or reject cosmological theories. One key aspect of this interpretation of probabilities is the condensation of *prior information* into a *prior distribution* of probabilities. More often than not, this goal has not been addressed with the accuracy the precision cosmology era demands. Some of the classic errors include the assignation of non-informative prior distributions to *nuisance parameters*. Our models include these to account for relevant astrophysical and systematic effects. Some collaborations nowadays contain dozens of these parameters in their analysis, surpassing the cosmological ones in large quantity (See [15] for an example, and [16] for a list of classic nuisance parameters from large galaxy surveys).

One of the biggest problems that come to light with the vast number of parameters is the *prior volume effect*. This corresponds to the biases due to the marginalizations done to nuisance parameters that have non-informative priors and, therefore, provide an excess of volume in the space of parameters. This can lead to mistaken cosmological interpretations for cases in which the probability distribution is highly non-Gaussian in some parameter space dimension.

This thesis presents a standardized methodology to assign prior distributions to nuisance parameters when prior information is not available or is hard to condense in a simple mathematical description. The *Informed Total-Error-Minimizing priors* (or ITEM priors) reduce the biases concerning some realistic data realizations and enforce a frequentist interpretation of the posterior constraints associated with the parameters of interest. Our method splits the vector of the model parameters into target and nuisance and generates posteriors using several different candidate priors for the latter. We next filter the candidates using requirements on the statistical bias and coverage for the marginalized target posteriors. The resulting prior will be the one that minimizes the uncertainty of the total error.

Given the importance of this problem in cosmology, we test this method in a re-analysis of the shot-noise of tracer galaxies from the Density Split Statistics (DSS) [1, 2] measured in Dark Energy Survey Year 1 data [17] and present them at the results section of this thesis.

As this work uses tools from the areas of physical cosmology and statistics, it has been initially divided into three parts. The first one introduces fundamental notions needed to model the cosmic large-scale structure: Chapter 2 presents an overview of cosmology that includes the basics of general relativity and the ΛCDM model of cosmology, while Chapter 3 describes the large-scale structure formation and introduces the density split statistics framework.

The second part describes the basics of the probabilistic and statistical knowledge needed to do cosmological data analysis: Chapter 4 is an overview of probabilities, whereas Chapter 5 covers many of the topics on statistical inference needed to accomplish the goal of this work.

The final part is the presentation of the main result: the ITEM prior formalism in the future publication Ried et al. 2022 (in prep.) in Chapter 7 leads to new constraints on the nuisance parameters and reduces the prior volume effect substantially. We finally discuss the achievements of this work and its contribution to the broad scientific community in Chapter 8.

Chapter 2

Basics of physical cosmology

There is substantial observational evidence to support that the universe is expanding [5–7], and in an accelerated way [8, 9]. Cosmologists, the scientists that study the physics of the universe, have observed that the space between galaxies increases as the cosmic evolution [12]. The further we observe, the higher this effect manifests because the expansion is the same everywhere. From this phenomenon, some logical questions to make are: *What will happen with galaxies in the future? Are they going to continue separating until the gravitational bounds are too faint? Is there going to be a turning point in which galaxies collapse?*

In this chapter, fundamental aspects of cosmology are presented to set the scene to answer those questions, including the cosmological principle in Sec. 2.1, the Einstein's theory of gravitation in Sec. 2.2, the physical evidence for an expanding universe in Sec. 2.3 and the standard model of cosmology, the ΛCDM model at Sec. 2.4. We conclude by describing the different cosmic distances that can be measured in our universe in Sec. 2.5. This introductory chapter is based on the books [12, 18] and lecture notes [19].

2.1. Cosmological principle

When observing physical distances larger than 100 Mpc, the mass distribution of the universe has proved to be essentially *homogeneous* and *isotropic* [20] assuming the *Copernican principle*: humans are not privileged observers and the physical laws are the same [21]. In other words, the average density of matter fluctuations looks the same in any direction (no preferred center). It does not change if there is a change in the position (no edges)[1].

The distributions of galaxies and clusters in space stand as evidence, but how did scientists derive this principle without observations? At the beginning of the twentieth century, there was no data to support the cosmological principle. Still, Einstein's field equations were consistent, allowing physicists like Friedmann and Lemaître to make inferences about general relativity, establishing the basis on which modern cosmology relies.

[1] It seems unlikely that our galaxy is located at the center of a spherically symmetric universe, giving us the impression of homogeneity and isotropic. Billions of galaxies outside the Milky Way seem equally good observers, and it would be an enormous coincidence. [12]

2.2. Gravitation

The general theory of relativity describes gravitation through geometry. It was published by Albert Einstein in 1916 [22] and generalizes the special relativity theory of a flat space-time geometry [23].

2.2.1. Special relativity

Galileo's inertial transformations can be generalized: in special relativity, the fundamental assumption is that the space-time interval $ds^2 = dt^2 - dx^2 - dy^2 - dz^2 = \eta_{\alpha\beta} dx^\alpha dx^{\beta 2}$ is invariant under the Lorentz transformations

$$y^\alpha = \Lambda^\alpha_\beta x^\beta + a^\alpha, \tag{2.1}$$

with $y = (t', x', y', z')$, $x = (t, x, y, z)$ and $dx = (dt, dx, dy, dz)$ 4-vectors and Λ^α_β, a^α constants and where $\eta_{\alpha\gamma}$ is the Minkowski metric:

$$\eta_{\alpha\gamma} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \tag{2.2}$$

The invariance of the space-time interval ds^2 implies that the matrix Λ^α_β must follow the property:

$$\Lambda^\alpha_\beta \Lambda^\gamma_\delta \eta_{\alpha\gamma} = \eta_{\beta\delta}\,. \tag{2.3}$$

2.2.2. General relativity

Eq. 2.3 can be generalize by assuming that the gravitational fields can be described by a general and symmetric metric:

$$g_{\alpha\beta} = \eta_{\mu\nu} \frac{\partial y^\mu}{\partial x^\alpha} \frac{\partial y^\nu}{\partial x^\beta} \tag{2.4}$$

when doing a general coordinate transformation from x^α to y^α. In particular, if $\frac{\partial y^\mu}{\partial x^\alpha} = \text{const.}$, we would recover the Minkowski metric (Eq. 5.2). This means that, locally, the gravitational field can be described by a space-time without curvature.

It is also possible to derive the equation of motion (or geodesic equation) in a system with a general metric $g_{\alpha\beta}$:

$$\frac{d^2 x^\lambda}{ds^2} + \frac{dx^\alpha}{ds} \frac{dx^\beta}{ds} \Gamma^\lambda_{\alpha\beta} = 0, \tag{2.5}$$

where $\Gamma^\lambda_{\alpha\beta} = \frac{\partial^2 x^\mu}{\partial x^\alpha \partial x^\beta} \frac{\partial x^\mu}{\partial x^\lambda}$ is the Christoffel symbol.

The derivation of the metric can be introduce ($g_{\alpha\beta,\mu} = \frac{\partial g_{\alpha\beta}}{\partial x^\mu}$) to obtain the geometry and the dynamical properties of the theory of general relativity [12]:

[2] We considered $c = 1$.

$$\Gamma^{\mu}_{\alpha\beta} = \frac{1}{2} g^{\mu\nu}(g_{\alpha\nu,\beta} + g_{\beta\nu,\alpha} - g_{\alpha\beta,\nu}), \tag{2.6}$$

where $g^{\mu\nu}$ is the inverse of the metric, also called the contravariant metric.

2.2.3. The FLRW metric

During the 1920s-1940s, four physicists (Alexander Friedmann [24], Georges Lemaître [5, 25], Howard P. Robertson [26–28] and Arthur Geoffrey Walker [29]) worked independently and derived a metric that describes a homogeneous and isotropic universe. The name for it is the *Friedmann–Lemaître–Robertson–Walker metric*, or simply, the FLRW metric. It can be expressed with spherical coordinates as:

$$ds^2 = -dt^2 + a^2(t)\left[\frac{dr^2}{1-kr^2} + r^2(d\theta^2 + \sin^2(\theta)d\phi^2)\right], \tag{2.7}$$

where $a(t)$ is the time-dependent scale factor and the constant k describes the curvature and can take any value.[3] The evolution of $a(t)$ depends on the energy components of the universe, as will be showed in Sec. 2.3.3.

It can be shown that this metric solves Einstein's field equations of general relativity [19]. It also allows a universe to expand or contract.

2.2.4. Einstein equations

It is possible to assume that particles in space-time are a perfect fluid, which implies that, with respect the center of mass and in a rest frame, it is homogeneous and isotropic. The conservation laws are:

$$\begin{aligned} \dot{\rho} &= 0, \\ \nabla p &= 0, \end{aligned} \tag{2.8}$$

where ρ is the energy density and $p = (p_x, p_y, p_z)$ is the pressure. We can define the *stress-energy tensor* $T_{\mu\nu}$ as:

$$T_{\mu\nu} = \mathrm{diag}(\rho, p_x, p_y, p_z), \tag{2.9}$$

where diag means a diagonal matrix. This tensor describes the properties of matter in the universe [19]. If isotropy is required, then $T_{\mu\nu} = \mathrm{diag}(\rho, p, p, p)$.

Using the notation proposed by [12], Einstein's equations relate the metric itself (geometry) at the LHS with the presence of matter (energy) at the RHS[5] as:

$$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = \kappa^2 T_{\mu\nu}, \tag{2.10}$$

[3] There are three possible scenarios: A *flat* universe ($k = 0$) correspond to one that follows a Euclidean geometry, i.e. the trajectories of particles in their geodesics remain parallel[4]. In a *closed* universe ($k > 0$), the trajectories of free particles will converge, while in an *open* universe ($k < 0$), the trajectories will diverge. [10]

[5] LHS and RHS are acronyms for Left-Hand Side and Right-Hand Side.

where $R_{\mu\nu}$ is the *Ricci tensor*:

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

which describes the properties of curvature of space-time. R is the *curvature scalar*:

$$R = g^{\alpha\beta}R_{\alpha\beta}. \tag{2.12}$$

Finally, κ^2, the *Einstein gravitational constant* is defined as:

$$\kappa^2 = \frac{8\pi G}{c^4}. \tag{2.13}$$

2.3. An expanding universe

In 1917, Albert Einstein proposed a static universe using General Relativity by an additional term, the *cosmological constant* Λ [30]:

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R - \Lambda g_{\mu\nu} = \kappa^2 T_{\mu\nu}, \tag{2.14}$$

which is independent of the observer's motion. This model is unstable, and therefore, it was discarded shortly after.

In 1918, de Sitter [31] demonstrated that a universe dominated by a cosmological constant would be expanding. But it was not until ten years later that there was observational data to prove it. Therefore, the cosmological constant was first rejected as a possible way of modeling the universe. However, in the last decades, it has become a candidate again with the observation of an accelerated expansion of the universe.

2.3.1. First evidence: The Hubble-Lemaître Law

At the end of the 1920s, scientists Georges Lemaître [5], and Edwin Hubble [6] discovered independently that galaxies farther away from us would move faster than nearby galaxies. They derived a linear relationship between the recession velocity v and the distance from us x:

$$v = H_0 x. \tag{2.15}$$

This is the famous *Hubble–Lemaître Law*. Its importance relies on the fact that it was the first observational evidence to support an expanding universe.

Fig. 2.1 is the original plot made by Edwin Hubble in 1929, which relates the distance and the recessive velocity of galaxies measured from the Earth [6].

In Fig. 2.2 we included the data used by Edwin Hubble in 1929 to derive the resulting esti-

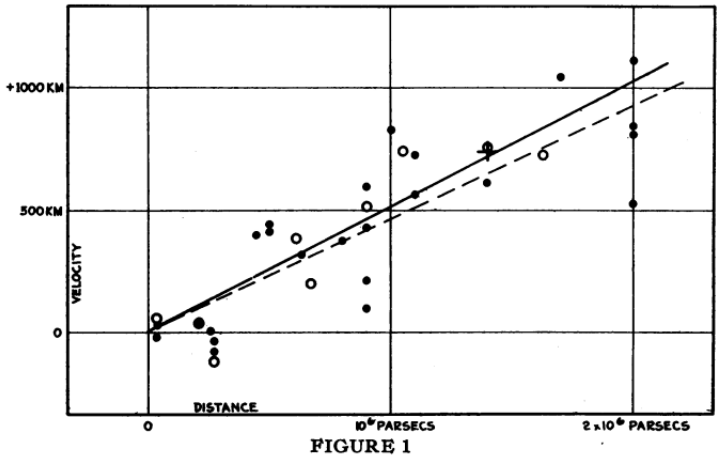


FIGURE 1

Velocity-Distance Relation among Extra-Galactic Nebulae.

Radial velocities, corrected for solar motion, are plotted against distances estimated from involved stars and mean luminosities of nebulae in a cluster. The black discs and full line represent the solution for solar motion using the nebulae individually; the circles and broken line represent the solution combining the nebulae into groups; the cross represents the mean velocity corresponding to the mean distance of 22 nebulae whose distances could not be estimated individually.

Figure 2.1: Original plot made by Edwin Hubble in 1929. [6] It relates the distance and the recessive velocity of galaxies measured.

mates on the Hubble constant[6]. [6]

Nowadays, the estimated value for H_0 differs by a factor of 9 and 7 from the ones found by Lemaître and Hubble, respectively. The inconsistencies are solved with a better calibration of the distances [7]. This discovery changed our conception of the universe.

These observations could be explained by the redshift effect derived from Friedmann's solutions to Einstein's equations of general relativity. [24]

[6] It is worth mentioning that what Lemaître and Hubble measured were the apparent and absolute magnitudes (m and M), not the distances directly. The apparent magnitude measures the brightness of an astronomical object observed from Earth. In contrast, the absolute magnitude is equal to the apparent magnitude that the object would have if it were viewed from a distance of 10 pc. They then derived the proper distance $x(t)$ between us and the galaxies:

$$x(t) = 10^{\left(\frac{m-M+5}{5}\right)}. \tag{2.16}$$

However, the results for H_0 were higher because Hubble used a wrong zero-point calibration of the standard candles. All distances were thus too small (by a factor of 7), and then, H_0 was too large by the same factor.

[7] Hubble used a least-square linear regression on 24 measurements of distances and recessional velocities of galaxies. In Section 5.2.2, we reproduce those results and introduce the least-squares method.

TABLE 1

NEBULAE WHOSE DISTANCES HAVE BEEN ESTIMATED FROM STARS INVOLVED OR FROM MEAN LUMINOSITIES IN A CLUSTER

OBJECT	m_s	r	v	m_t	M_t
S. Mag.	..	0.032	+ 170	1.5	−16.0
L. Mag.	..	0.034	+ 290	0.5	17.2
N. G. C. 6822	..	0.214	− 130	9.0	12.7
598	..	0.263	− 70	7.0	15.1
221	..	0.275	− 185	8.8	13.4
224	..	0.275	− 220	5.0	17.2
5457	17.0	0.45	+ 200	9.9	13.3
4736	17.3	0.5	+ 290	8.4	15.1
5194	17.3	0.5	+ 270	7.4	16.1
4449	17.8	0.63	+ 200	9.5	14.5
4214	18.3	0.8	+ 300	11.3	13.2
3031	18.5	0.9	− 30	8.3	16.4
3627	18.5	0.9	+ 650	9.1	15.7
4826	18.5	0.9	+ 150	9.0	15.7
5236	18.5	0.9	+ 500	10.4	14.4
1068	18.7	1.0	+ 920	9.1	15.9
5055	19.0	1.1	+ 450	9.6	15.6
7331	19.0	1.1	+ 500	10.4	14.8
4258	19.5	1.4	+ 500	8.7	17.0
4151	20.0	1.7	+ 960	12.0	14.2
4382	..	2.0	+ 500	10.0	16.5
4472	..	2.0	+ 850	8.8	17.7
4486	..	2.0	+ 800	9.7	16.8
4649	..	2.0	+1090	9.5	17.0
Mean					−15.5

m_s = photographic magnitude of brightest stars involved.
r = distance in units of 10^6 parsecs. The first two are Shapley's values.
v = measured velocities in km./sec. N. G. C. 6822, 221, 224 and 5457 are recent determinations by Humason.
m_t = Holetschek's visual magnitude as corrected by Hopmann. The first three objects were not measured by Holetschek, and the values of m_t represent estimates by the author based upon such data as are available.
M_t = total visual absolute magnitude computed from m_t and r.

Figure 2.2: The galaxies used by Hubble in his original work [6]. He included the distances (in the table labeled as r), the velocities v and the magnitudes m_t and M_t.

2.3.2. Redshift, scale factor and Hubble rate

The expansion of space proportionally elongates the light wavelength ($\lambda_{emit} < \lambda_{obs}$), and inverse proportionally decreases the momentum and energy of the photon. Therefore, we will be observing a different type of light from the one emitted. We call this effect *redshift* because, in the electromagnetic spectra, in the optical range, the color red has a longer wavelength.[8] We can measure this phenomenon with the semi-positive quantity z as follows:

$$1 + z = \frac{\lambda_{obs}}{\lambda_{emit}}. \tag{2.17}$$

Therefore, if $z = 0$, there would be no stretch at all and $\lambda_{emit} = \lambda_{obs}$.[9]

The relative expansion of the universe is parameterized by a dimensionless time-dependent cosmic *scale factor* $a(t)$. This relates the proper physical distance $x(t)$ (which can change

[8] The opposite of this effect is called *blueshift* because blue has a shorter wavelength.

[9] If $z = 1$, it implies that $2\lambda_{emit} = \lambda_{obs}$, which is equivalent to saying that the distances at the time of the observation are twice as they were original.

over time, unlike the distance measured along a null geodesic $ds = 0$, also called comoving distance (see Section 2.5 for a deeper discussion) between two points in different times. The formula for this is:

$$x(t) = a(t)x(t_0), \tag{2.18}$$

where $x(t)$ is the proper distance at the time t, $x(t_0)$ is the comoving distance (t_0 is set to be the present age of the universe), and $a(t)$ is the scale factor.[10] This function allows for homogeneous expansion or contraction of space over time. By definition, $x(t_0) = x_0$ and $a(t_0) = 1$.

To quantify the change in the scale factor, we use the *Hubble rate* $H(t)$ which is the rate of expansion of space. [10] For small and radial propagation distances $d\theta = d\phi = 0$, the Hubble rate can be approximated by dividing the recession velocities of galaxies by their distance from us. This expansion rate is defined as:

$$H(t) = \frac{1}{a}\frac{da}{dt}. \tag{2.19}$$

The Hubble rate varies with time, being the Hubble constant H_0 the current value:

$$H(t = 0) = H_0. \tag{2.20}$$

We can measure the Hubble rate in units of velocity per distance:

$$H(t) \propto \left[\frac{\text{km}}{\text{s} \cdot \text{Mpc}}\right], \tag{2.21}$$

where $\text{Mpc} = 10^6$ pc. If we calculate the derivative of $x(t) = a(t)x_0$, we obtain the velocity:

$$v(t) = \frac{d(x(t))}{dt} = \frac{a(t)x_0}{dt} = \frac{da}{dt}x_0, \tag{2.22}$$

where $\dfrac{d(x_0)}{dt} = 0$ because it is constant. Then, if we substitute the above definition of the Hubble rate gives:

$$v = \frac{da}{dt}x_0 = H_0 x(t) = H_0 x, \tag{2.23}$$

which is the famous Hubble–Lemaître Law. [10] This law applies to any system that expands (or contracts) in a homogeneous and isotropic way. However, this relation is no longer valid for larger distances and velocities.

Furthermore, the final Hubble rate depends on other factors, such as the universe's energy content. In Section 2.4, we will explore the case of a universe filled with matter (baryonic and dark matter), radiation, and dark energy.

[10] The scale factor is determined by the equations of general relativity, which are related to the distribution of density of energy and other details. See Section 2.4 for an example.

2.3.3. Friedmann Equations

It can be shown that the components of the Ricci tensor are:

$$\begin{aligned} R_{00} &= -3\frac{\ddot{a}}{a}, \\ R_{ii} &= -\frac{g_{ii}}{a^2}(a\ddot{a} + 2\dot{a}^2 + 2k) \end{aligned} \tag{2.24}$$

and the R curvature scalar:

$$R = -\frac{6}{a^2}(\dot{a}^2 + a\ddot{a} + k). \tag{2.25}$$

These can be introduced into the Einstein field equations (Eq. 2.10) to obtain the *Friedmann equations*. The temporal and spatial parts (the 00 and ii components) combined will return two independent relations:

$$\begin{aligned} H^2 &= \frac{8\pi}{3}\rho - \frac{k}{a^2}, \\ \frac{\ddot{a}}{a} &= -\frac{4\pi}{3}(\rho + 3p). \end{aligned} \tag{2.26}$$

From which we can derive the *conservation equation*:

$$\dot{\rho} + 3H(\rho + p) = 0. \tag{2.27}$$

These expressions describe an expanding universe in and homogeneous and isotropic regime, filled with a perfect fluid with a given mass density ρ and pressure p.

It is possible to define a the *critical density* ρ_c:

$$\rho_c = \frac{3H^2}{8\pi G}, \tag{2.28}$$

such that the components of the fluid are described by a *density parameter* $\Omega(a)$:

$$\Omega = \frac{\rho}{\rho_c}. \tag{2.29}$$

Introducing Eq. 2.29 into the first Friedmann equation (Eq. 2.26), will imply that

$$1 = \Omega - \frac{k}{a^2H^2}. \tag{2.30}$$

It is common to define a curvature component that accounts for the overall geometry of the universe:

$$\Omega_k(a) = -\frac{k}{a^2H^2} \tag{2.31}$$

and its corresponding critical density

$$\rho_k = -\frac{3k}{8\pi a^2}, \tag{2.32}$$

such that

$$1 = \Omega(a) + \Omega_k(a). \tag{2.33}$$

Eq. 2.33 will vary depending on the considered models and components interacting in the universe at a specific epoch.

2.4. ΛCDM model of the Universe

Also known as the *standard model of cosmology*, the ΛCDM model is one of the simplest ones that fit the experimental observations. It allows us to predict how the energy density should evolve with time and the temperature decrease.

2.4.1. Cosmic inventory

The ΛCDM model divides the energy-momentum sources into three components (also called cosmic inventories):

- **Matter**:

  One of the characteristic of the massive particles, or *matter*, is that

  $$|p| \ll \rho. \tag{2.34}$$

  A fluid with zero pressure ($p = 0$) approximates matter as non-interacting particles with velocities that are non-relativistic.

  From Eq. 2.27, we can derive a density relation[11]:

  $$\frac{\dot{\rho_m}}{\rho_m} = -3\frac{\dot{a}}{a} \Rightarrow \frac{\rho_m}{\rho_c} = \Omega_m(1+z)^3. \tag{2.35}$$

  The ΛCDM model considers two types of matter-components:

  – **Baryonic matter**:

    In this context, *baryons* refers to all the nuclei and electrons of the universe[12]. The main characteristic of the baryonic matter is that it interacts with radiation. Therefore, its observations are related to the light coming from galaxies (e.g., we can infer the amount of matter in a star by studying its luminosity). Baryonic matter is not only distributed in bright objects, like stars, but also in the interstellar medium. This can be observed when studying the spectra of distant quasars and measuring

[11] It is also possible to derive that relation by considering that the energy of a non-relativistic particle, such as matter, will be equal to its rest mass energy [10].

[12] This is somehow wrong because electrons are *leptons*. Still, as nucleons are substantially more massive, this is not a poor assumption.

the absorption lines from intervening hydrogen in the line of sight.

When comparing the matter quantity measured from light sources, there is a substantial difference with the one estimated using the gravitational field of a system. That disagreement is solved with dark matter.

– **Cold dark matter**:

Dark matter is a concept that was first introduced by Zwicky in 1933 [32] and later corroborated by Vera Rubin in 1983 [33, 34] when studying the rotation curve of spiral galaxies. It solved the evident miss-match between the observed baryonic matter and the one estimated from the gravitational interaction.

As dark matter is still a mystery, there are many candidates. Some of them are particles that are not included in the standard model of particles [35].

In the ΛCDM model, the CDM stands for Cold Dark Matter. The *cold* comes from requiring the dark matter particles to be able to clump efficiently in the early universe, having a negligible pressure or, equivalently, with a non-relativistic velocity [18].

- **Radiation**:

Relativistic particles, such as photons, neutrinos, and other quantities from which the energy density is dominated by its kinetic energy, will be denoted *radiation*. The main characteristic is that their pressure is related to the density as:

$$p = \frac{1}{3}\rho. \tag{2.36}$$

The corresponding density relation would be given by:

$$\frac{\dot{\rho_r}}{\rho_r} = -4\frac{\dot{a}}{a} \Rightarrow \frac{\rho_r}{\rho_c} = \Omega_r(1+z)^4. \tag{2.37}$$

- **Dark energy**:

In 1998, the High-redshift Supernova Search Team (HSST) [8] and the Supernova Cosmology Project [9] independently observed supernovae of type Ia (SNIa)[13] and inferred that the universe is lately expanding in an accelerated way. This phenomenon can be described as a negative pressure:

[13] Supernovae are explosions that happens when the *Chandrasekhar limit* is reached, and depending on the nature of the system, there would be a different emission of chemical elements. If there is a lack of a spectral line of hydrogen, we will classify the supernova as a Type I. In addition, if the spectra include an absorption line of ionized silicon, it would be sub-classified as a Type Ia. These supernovae occurs in a binary system of a white dwarf absorbing gas from a companion star. [19]

$$p = -\rho. \tag{2.38}$$

The corresponding density relation would be given by:

$$\frac{\dot{\rho_\Lambda}}{\rho_\Lambda} = 0 \Rightarrow \frac{\rho_\Lambda}{\rho_c} = \Omega_\Lambda. \tag{2.39}$$

Since the energy density does not dilute with the scale factor, energy has to be transformed from another source as the universe expands. [18]

As in the case of dark matter, dark energy is still one of the major challenges in current cosmological research.

The energy density parameter ρ includes the energy density contribution from all the species from the considered model. It includes matter, radiation, and dark energy, and as we are not assuming a flat universe, there would be an additional term for the curvature. Unless stated differently, we will consider a flat ΛCDM universe ($\Omega_k = 0$) throughout this work.

Finally, the Friedmann equation of a ΛCDM model would be equivalent to:

$$H(a) \equiv \frac{\dot{a}}{a} = H_0\sqrt{\Omega_m(1+z)^3 + \Omega_r(1+z)^4 + \Omega_\Lambda} \tag{2.40}$$

from which

$$\Omega_m + \Omega_r + \Omega_\Lambda = 1. \tag{2.41}$$

It is possible to determine the epochs in which an element dominated the scale factor and, therefore, the cosmic expansion.

Originally, radiation dominated due to the low spatial density, which forced particles to have a larger pressure than matter. This transition (the matter-radiation equality) happened at $z \sim 3400$. Recent observations have shown that the scale factor does not follow the matter-dominated predictions but the one described by a dark energy-dominated era (dark energy-matter equality happened at $z = 0.4$). See Figs. 1.2 and 1.3 from [10] and Table 3.1 from [18].

One of the main extensions that can be made to the ΛCDM model is *cosmic inflation*. It corresponds to a phase that occurred approximately 10^{-33} seconds after the theoretical Big Bang, and it lasted $\sim 10^{-36}$ seconds. It is a theory of exponential growth of space in the early universe, and it would explain the homogeneity observed in the CMB. [36]. However, despite that it solves the flaws of the ΛCDM model, inflation is not predicted by the standard model of particle physics and is still largely debated.

2.4.2. Observational evidence

Some of the observational evidence that supports the ΛCDM model of the universe is:

- *Cosmic microwave background*: One of the key results from the CMB observation is the small anisotropies present on it, i.e. temperature fluctuations also support homogeneity. It is still a major discovery that the early universe was very smooth.

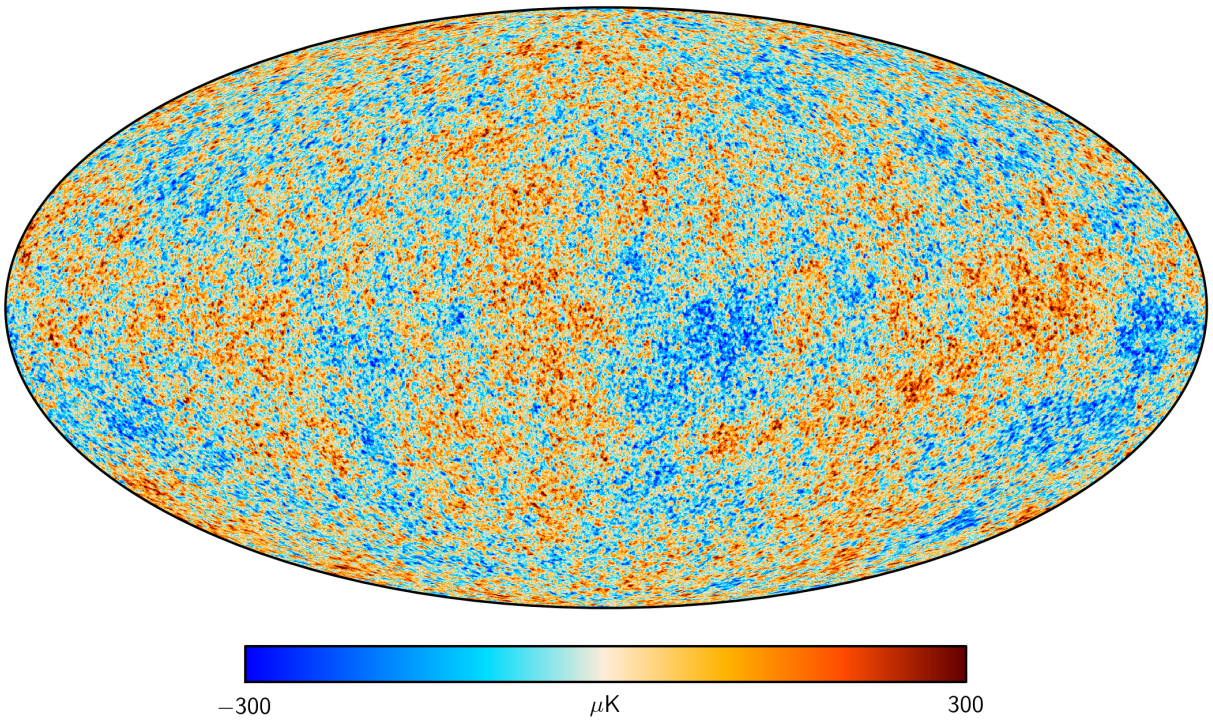


Figure 2.3: The Cosmic Microwave Background data from the Planck satellite mission [37].

- *Big bang nucleosynthesis*: Shortly after the discovery of the CMB, Peebles studied the consequences it had over the formation of nuclei in the first few minutes of the hot dense universe, determining the abundance of light elements created in that epoch [38].
- *Hubble-Lemaître Law*: Observations made independently by Hubble and Lemaître showed that galaxies are moving away from Earth at speeds proportional to their distance. In Sec. 2.3.1 we address in more detail on this observational evidence of cosmic structure expansion.
- *Large-scale structure*: Since the fist large surveys of galaxies of the last decades of the twentieth century (like the Sloan Digital Sky Survey [39] and the Two-degree-Field Galaxy Redshift Survey [40]) we have substantial evidence to state that galaxies are not distributed randomly in the sky (see Fig. 2.4). These features are obtained from the ΛCDM model when perturbations around the smooth background are included [10]. In Chapter 3 we will study the theory of cosmic structure formation.

2.5. Cosmic distances

Distances in cosmology are estimated by inferring a celestial object's absolute magnitude (intrinsic brightness). A common approach is to use a *standard candle*. They are defined as the type of astronomical object that has a known absolute magnitude, i.e., the same intrinsic

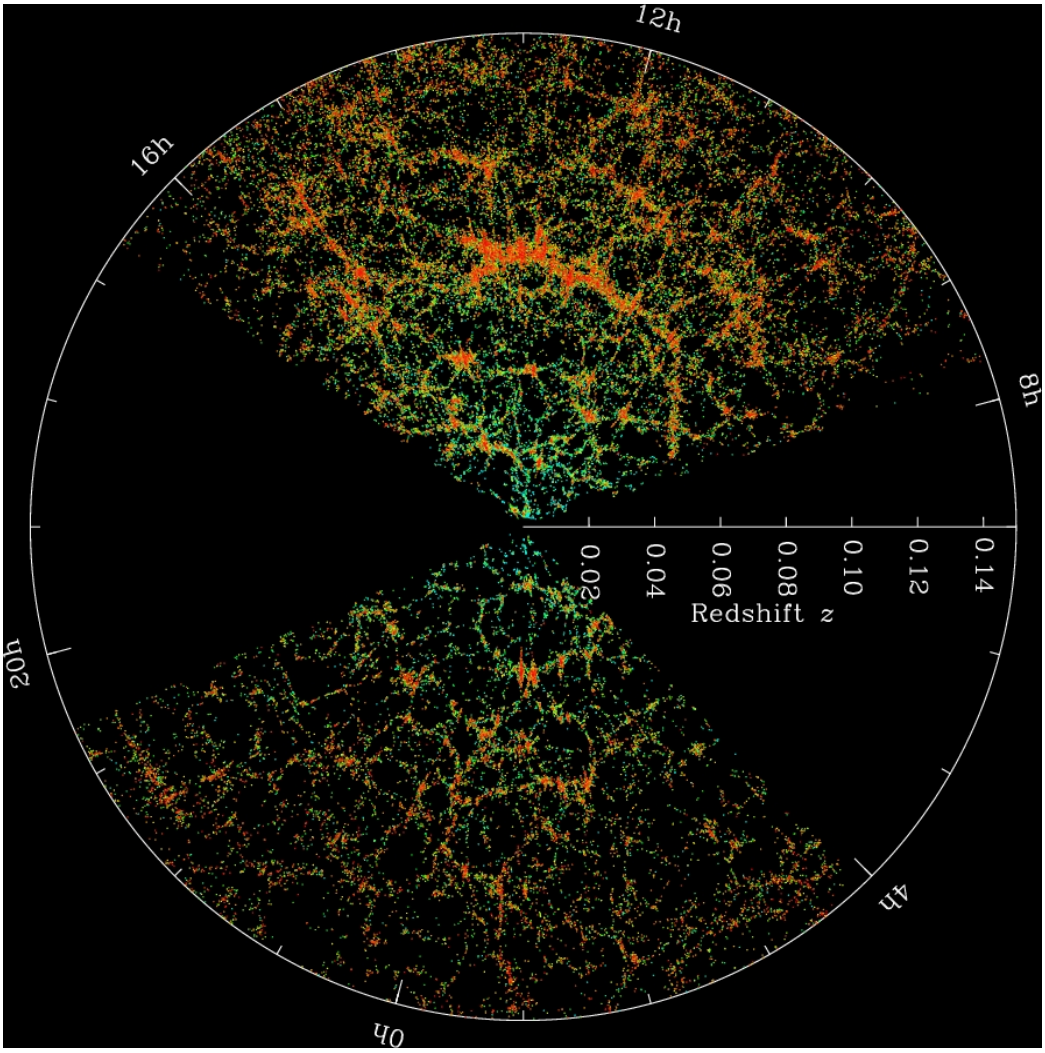


Figure 2.4: The SDSS's map of the Universe. Each dot is a galaxy; the color is the g-r color of that galaxy. As redshift increases, the filament structure, also known as the cosmic web, shows no random distribution of galaxies. Image Credit: M. Blanton and SDSS https://www.sdss.org/science/orangepie/

brightness. Two examples are *Cepheid variables*[14] and *Type Ia supernovae*[15].

From the Friedmann equation (Eq. 2.26) with different contributions to the energy density, and using the same notation as [43], we can derive an expression for the cosmological expansion:

$$H(z) = H_0\sqrt{\Omega_m(1+z)^3 + \Omega_\Lambda}, \tag{2.42}$$

where the density parameters are as they would be measured today. This expression gives us direct information about the composition of the universe. We can use it to measure distances,

[14] Red giant near the end of their life stars that have a clear relation between the pulsation period and their luminosity [19, 41]. They can be find in low redshift ($\sim$ 10Mpc)

[15] They are distributed in a large redshift range, but most distant known is in redshift $z = 1.914$ [42]

which is not trivial since points in space are not static. The *Hubble distance* d_H is given by:

$$d_H = \frac{c}{H_0}. \tag{2.43}$$

Another fundamental way of measuring distances in the universe is the *comoving distance* d_C (remains fixed as the universe expands), which is measured between two nearby objects that are moving with the Hubble flow. Therefore, it depends on the redshift as:

$$d_C(z) = d_H \int_0^z \frac{dz'}{E(z')}, \tag{2.44}$$

with $E(z)$ being equivalent to:

$$E(z) = \frac{H(z)}{H_0} = \sqrt{\Omega_m(1+z)^3 + \Omega_\Lambda}. \tag{2.45}$$

When two events are at the same redshift but are separated in the sky by some angle, we use the *transverse comoving distance* d_M, and it is related to the line-of-sight comoving distance d_C as:

$$d_M(z) = \begin{cases} \dfrac{d_H}{\sqrt{\Omega_k}} \sinh\left(\dfrac{\sqrt{\Omega_k} d_C(z)}{d_H}\right) & \Omega_k > 0 \\ d_C(z) & \Omega_k = 0 \\ \dfrac{d_H}{\sqrt{|\Omega_k|}} \sin\left(\dfrac{\sqrt{|\Omega_k|} d_C(z)}{d_H}\right) & \Omega_k < 0 \end{cases}, \tag{2.46}$$

where Ω_k gives us the curvature of space being derived from the other energy density quantities:

$$\Omega_k = 1 - \Omega_m - \Omega_\Lambda. \tag{2.47}$$

If we re-scale $d_M(z)$ by the scale factor, we obtain the *angular diameter distance*:

$$d_A(z) = a d_M(z), \tag{2.48}$$

which for the case $\Omega_k = 0$, this would be the proper distance between a source and the observer at the time of emission.

The final distance we will derive is the *luminosity distance* d_L, that is related to the transverse comoving distance by

$$d_L(z) = (1+z) d_M(z). \tag{2.49}$$

This last one is defined in terms of the relationship between the apparent magnitude m and absolute magnitude $\mathcal{M}$ of an astronomical object.

$$m = 5 \log_{10}(H_0 d_L) + \mathcal{M}. \tag{2.50}$$

For a deeper explanation of distance measures in cosmology, see [43].

Chapter 3

Cosmic structure formation

The cosmological principle states that the average matter density universe is homogeneous and isotropic on large scales. However, when observing the spatial distribution of galaxies, as illustrated in Fig. 2.4, galaxies are not randomly distributed: they form clusters, voids, and filaments. Specific points in space with high matter density (such as galaxies) and large zones that are almost empty (voids) in the late universe imply that it is highly inhomogeneous.

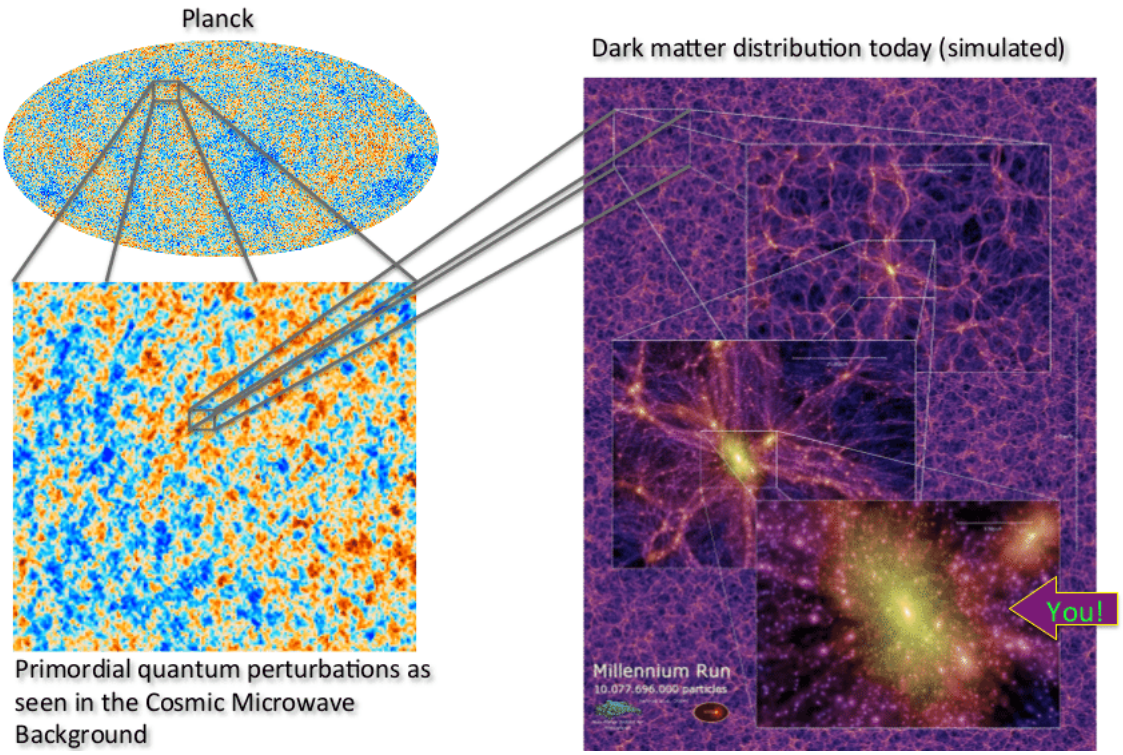


Figure 3.1: On the left part of the figure, we have Planck's measurements [44] of the anisotropies from the CMB, while in the right, we show the Millenium Run simulation [45] today's dark matter distribution following the seeds from the CMB. Figure from [46].

This was not always the case: the measured CMB temperature fluctuations are small, suggesting that the distribution of matter was highly homogeneous in the early universe. One

implication is the symmetry of the underdense and overdense regions (abundance and amplitudes) in very early times. The initial overdensities generated a gravitational potential that made matter collapse into the current structures, while initial underdensities grew and stayed mainly underdense. In Fig. 3.1 we show how primordial perturbations seed the *large-scale structure*[16] (LSS) in the universe through a gravitational instability mechanism.

Everything points that initial and small perturbations of the homogeneous density were amplified, generating the large structures that we find in the universe nowadays. If matter distribution started from a smooth density field that had slight inhomogeneities, *what mechanism explains this phenomenon?* Or in other words, *how did matter density fluctuations evolve through time and the universe's expansion?* [20]

We answer those questions in this chapter: In Sec. 3.1 we present the preliminary mathematical and physical knowledge needed to model the evolution of the matter field. In Sec. 3.2 a linear perturbation solution of the equations of motion of the cosmic fields is summarized, in addition to a description of the point statistics and the cosmic probability density function of matter density. In Sec. 3.3 we relate the matter density field with the observations we can make of the cosmic structure. We summarize some of the current and future sky surveys that allow us to infer the density field at Sec. 3.4. We end outlying a higher-order technique to trace the matter density field called density split statistics at Sec. 3.5 and focus on the specific models that describe the galaxy-matter connection. This chapter is based on the books [36, 38], the review [47] and the thesis work [48].

3.1. Preliminaries

There are three main quantities we will focus when studying the cosmic structure formation of the universe: mass density $\rho(\boldsymbol{x}, t)$, velocity $\boldsymbol{v}(\boldsymbol{x}, t)$, pressure $p(\boldsymbol{x}, t)$ and a self-gravitational potential $\Phi(\boldsymbol{x}, t)$, where $\boldsymbol{x}$ and t corresponds to a given location and time respectively.

3.1.1. Equations of motion

Matter, in cosmological scales, can be assumed as a self-gravitating classic fluid without shear or viscosity. Fluids like this are described by three fundamental equations: the *continuity equation*, the *Euler equation* and the *Poisson equation*. These three relate to the fluid's mass density, streaming velocity, and pressure.

- **Continuity equation**: Describes the evolution of the matter density field due to the fluxes in a system, in which the fluid moves but it is neither destroyed nor created:

$$\frac{\partial \rho}{\partial t} + \nabla(\rho \boldsymbol{v}) = 0. \tag{3.1}$$

- **Euler equation**: Describes the evolution of the velocity field due to the forces, in which the fluid is accelerated by both, pressure and gravity:

[16] Refer to all structures that are bigger than galaxies.

$$\frac{\partial \boldsymbol{v}}{\partial t} + \boldsymbol{v}\nabla\boldsymbol{v} + \frac{1}{\rho}\nabla p = -\nabla\Phi. \tag{3.2}$$

- **Poisson equation**: Relates the matter density field to its self-gravitational potential:

$$\nabla^2\Phi = 4\pi\rho. \tag{3.3}$$

3.1.2. Density fluctuations

We can consider perturbations on the matter density contrast $\delta(\boldsymbol{x}, t)$ as:

$$\delta(\boldsymbol{x}, t) = \frac{\rho(\boldsymbol{x}, t) - \overline{\rho}}{\overline{\rho}}, \tag{3.4}$$

where $\overline{\rho}$ is the mean density of all space, and because of this definition, $-1 \leq \delta < \infty$. It also mathematically implies that we consider density perturbations on top of a homogeneous background density:

$$\rho(\boldsymbol{x}, t) = \overline{\rho}(1 + \delta(\boldsymbol{x}, t)). \tag{3.5}$$

These perturbations will cause deviations of the other quantities, such as the velocity field and the gravitational potential (deviations from the Hubble flow due to the density fluctuations). A way of decoupling the perturbations from the Hubble flow is to work with *co-moving coordinates*. Following the approach from Chapter 6 of [36], the equations of motion of the perturbation can be written given by:

- **Continuity equation**:

$$\delta' + \nabla([1 + \delta]\boldsymbol{u}) = 0. \tag{3.6}$$

- **Euler equation**:

$$\boldsymbol{u}' + (\boldsymbol{u}\nabla)\boldsymbol{u} + \mathcal{H}\boldsymbol{u} = -\nabla\phi. \tag{3.7}$$

- **Poisson equation**:

$$\nabla^2\phi = 4\pi a^2\overline{\rho}\delta. \tag{3.8}$$

where $\boldsymbol{u}$ is the co-moving velocity perturbations, ϕ the scalar metric perturbation, $'$ denotes derivation in the conformal time τ and $\mathcal{H}$ is the conformal expansion rate: $\mathcal{H} = \dfrac{a'}{a}$.

These equations are equivalent to introducing small deviations to a homogeneous background universe. These perturbations can be added to the FLRW metric for a universe, e.g., filled with matter, affecting the distribution fields directly [36].

3.1.3. Initial Conditions

When photons from the CMB were emitted, the universe was not entirely homogeneous: there were small differences (of the order 10^{-5}) in density and temperature that can be observed in the left panel of 3.1. The inflationary theory (briefly explained in Sec. 2.4) predicts that the primordial density fluctuations are homogeneous, isotropic and Gaussian[17] [47]:

$$P(\delta)d\delta = \frac{1}{\sqrt{2\pi}\sigma}e^{-\delta^2/(2\sigma^2)}d\delta. \tag{3.9}$$

Small perturbations, present in the early universe, are such that $\delta \ll 1$. This can be understand as that in space, $\rho(\boldsymbol{x}, t) \sim \overline{\rho}$, and therefore, the difference is relatively small due to the symmetry between under-dense ($\rho(\boldsymbol{x}_{\text{under-dense}}, t) < \overline{\rho}$) and over-dense regions ($\rho(\boldsymbol{x}_{\text{over-dense}}, t) > \overline{\rho}$). Due to the effect of gravity, these fluctuations grew with the cosmic history, and remain Gaussian until the onset of non-linearity, which implied $\delta \geq 1$ in the later universe. The solutions of perturbation theory must account for this feature in their initial conditions.

3.2. Physics of the cosmic structure

In the ΛCDM model of the universe, Eqs. 3.6, 3.7, and 3.8 can be solved through different methods, including exact and numerical solutions. In this work, we solve the linear equations of motion through perturbation theory to illustrate how fields evolve through the cosmic history in the simplest case. We refer to [36] for a review on non-linear perturbation theory.

3.2.1. Linear perturbation theory

The terms $\delta\boldsymbol{u}$ and $(\boldsymbol{u}\nabla)\boldsymbol{u}$ from Eqs. 3.6 and 3.7 respectively, are second order. For small perturbations, these terms can be neglected, and therefore, offer simpler equations of motion and solutions. In the linear regime, the equations of motion can be written as:

$$\delta' + \nabla\boldsymbol{u} = 0, \tag{3.10}$$

$$\boldsymbol{u}' + \mathcal{H}\boldsymbol{u} = -\nabla\phi. \tag{3.11}$$

After taking the divergence of Eq. 3.11 and combining 3.8 and 3.10, it is possible to derive a third equation:

$$\delta'' + \mathcal{H}\delta' - 4\pi a^2\overline{\rho}\delta = 0, \tag{3.12}$$

which has the general solution:

$$\delta_L(\boldsymbol{x}, \tau) = \frac{D(\tau)}{D(\tau_0)}\delta(\boldsymbol{x}, \tau_0), \tag{3.13}$$

where δ_L stands for linear, $D(\tau)$ is the *linear growth factor* and τ_0 refers to the conformal time today. It is common to set $D(\tau_0) = 1$ and to use the following expression for D from [49]:

[17] See Section 4.2.2 for an explanation on continuous probability distributions.

$$D(z) = \frac{5}{2}\Omega_m E(z) \int_z^\infty \frac{1+z'}{E^3(z')} dz'. \tag{3.14}$$

3.2.2. Correlation functions

In statistics, the n-point correlation function correlates n variables in some coordinate (e.g. spatial or temporal distances). Let $\boldsymbol{x}_1, ..., \boldsymbol{x}_N$ be locations in the universe. The n-point correlation function of the density contrast δ will be:

$$\xi^{(n)}(\boldsymbol{x}_1, ..., \boldsymbol{x}_N, \tau) = \langle \delta(\boldsymbol{x}_1, \tau)\delta(\boldsymbol{x}_2, \tau)...\delta(\boldsymbol{x}_N, \tau)\rangle_c, \tag{3.15}$$

where the subscript c denotes the *connected parts* of the joint ensemble average of the density in an arbitrary number of locations. (See Eq. 125 from [47] for a derivation.)

In particular, the two-point correlation function can be interpreted as the average of the density at two different locations, or as the measure of clustering of a distribution in space:

$$\xi(r) = \langle \delta(\boldsymbol{x})\delta(\boldsymbol{x} + \boldsymbol{r})\rangle. \tag{3.16}$$

where we assume a fixed τ, that $\xi(r)$ depends only on the norm of $\boldsymbol{r}$ due to statistical homogeneity and isotropy [47].

3.2.3. Power spectrum

The density contrast $\delta(\boldsymbol{x})$ can be written on its Fourier components:

$$\delta(\boldsymbol{x}) = \int \delta(\boldsymbol{k}) \exp\{i\boldsymbol{k} \cdot \boldsymbol{x}\} d^3\boldsymbol{k}. \tag{3.17}$$

The *matter power spectrum* $P(k)$ is the Fourier transform of the two-point correlation function:

$$\langle \delta(\boldsymbol{k})\delta(\boldsymbol{k}')\rangle = \delta_D(\boldsymbol{k} + \boldsymbol{k}')P(k), \tag{3.18}$$

where δ_D is the Dirac delta function and:

$$P(k) = \frac{1}{(2\pi)^3} \int \xi(r) \exp\{i\boldsymbol{k} \cdot \boldsymbol{r}\} d^3\boldsymbol{r}. \tag{3.19}$$

The power spectrum can be used to estimate the abundance of the cosmic contents (radiation, baryonic matter, and dark matter) through the analysis of its shape in the early universe. It can also be used to determine the growth of structure because it is possible to see its evolution through different observables (CMB and galaxy distribution). In Fig. 3.2 we illustrate today's power spectrum, where k is the wavenumber and it related to the wavelength as $\lambda = 2\pi/k$. In large scales, the linear growth factor from Eq. 3.14 acts on the power spectrum as:

$$P(k, \tau) = D^2(\tau)P(k, 0), \tag{3.20}$$

where $\tau = 0$ is set to be a time shortly after the end of inflation. $P(k, 0)$ is also known as the *primordial matter power spectrum*.

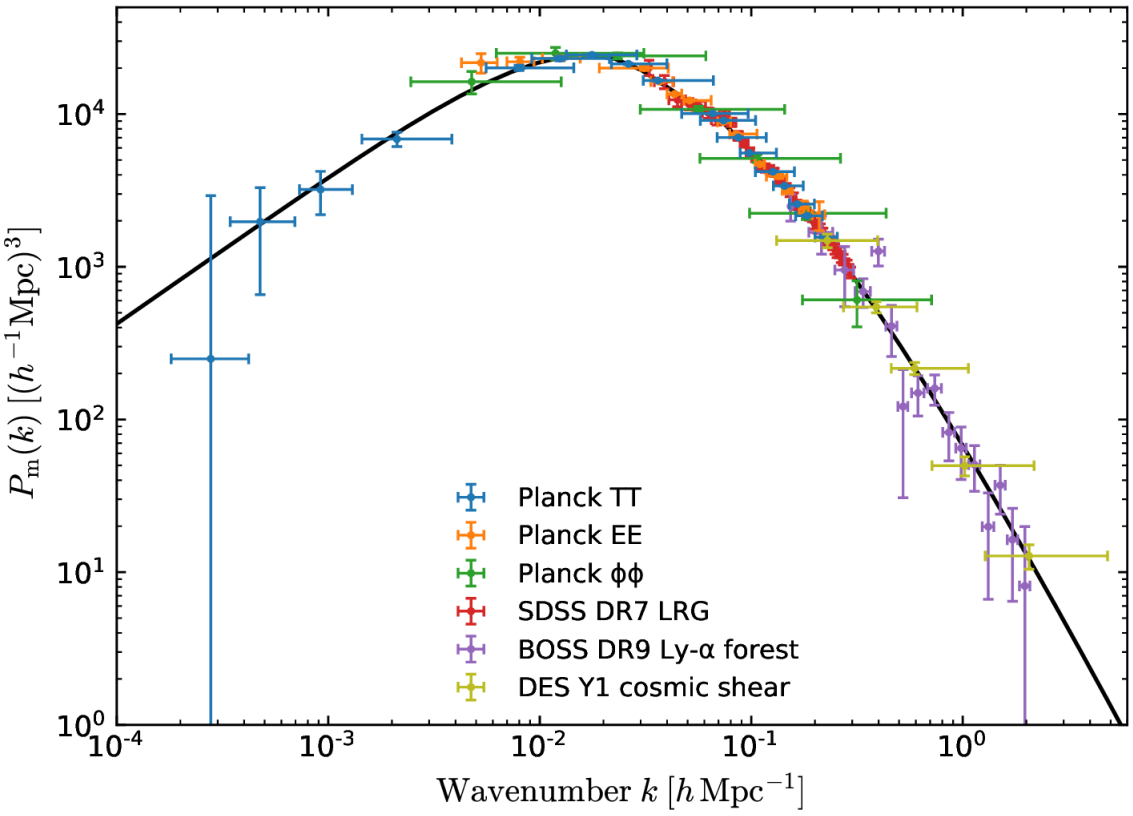


Figure 3.2: The linear matter power spectrum at $z = 0$ constructed from different cosmological probes. Image from [50].

This expression should be extended to the non-linear regime because the linear approximation underestimates the growth of structures on the scales in which collapses occur.

The power spectrum, as the correlation function, can be generalized into an n-point descriptor:

$$\langle \delta(\boldsymbol{k}_1)...\delta(\boldsymbol{k}_N)\rangle_c = \delta_D(\boldsymbol{k}_1 + ... + \boldsymbol{k}_N)P_N(\boldsymbol{k}_1 + ... + \boldsymbol{k}_N). \tag{3.21}$$

For $n = 3$, $P_3(\boldsymbol{k}_1 + \boldsymbol{k}_2 + \boldsymbol{k}_3)$ is called the *bispectrum* and usually denoted $B(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3)$.

3.2.4. The cosmic density PDF

It is common to define the probability density function (PDF) of matter density fluctuations in the universe using the correlation functions. We will first illustrate the case for the two-point correlation function.

The probability dP that two particles at volume elements dV_1 and dV_2 are separated by distance $x_{12} = |\boldsymbol{x}_1 - \boldsymbol{x}_2|$, is given by:

$$dP = n^2[1 + \xi(x_{12})]dV_1dV_2, \tag{3.22}$$

where n is the mean density. [47]

If $\xi(x_{12}) = 0$, i.e., there is no clustering, the probability is simply the mean density squared, independently of the distance. For the case in which $\xi(x_{12}) > 0$, there would be an excessive clustering, while for $\xi(x_{12}) < 0$, objects would be anti-correlated and probabilities of finding mass concentrated will be lower.

With a similar reasoning, the probability of having three particles at volume elements dV_1, dV_2 and dV_3 will be:

$$dP = n^3[1 + \xi(x_{12}) + \xi(x_{23}) + \xi(x_{13}) + \xi^{(3)}(x_{12}, x_{23}, x_{13})]dV_1 dV_2 dV_3, \tag{3.23}$$

where $\xi^{(3)}$ is the three-point correlation function.

One of the main characteristics of Gaussian random fields is that they can be described entirely with their first two moments. The theory of inflation predicts that in the early universe, the density contrast field $\delta(\boldsymbol{x}, \tau)$ followed a Gaussian distribution. This is equivalent to stating that all the cosmological information was initially condensed in its two-point correlation function and, therefore, in its power spectrum. With the formation of the cosmic structure due to fluctuations, the n-point correlation functions become nonzero, developing *higher-order statistics* of the density field.

It is common to *smooth* the density distribution with some filter W_R of size R and derive the moments δ_R:

$$\delta_R(\boldsymbol{x}, \tau) = \int W_R(|\boldsymbol{x} - \boldsymbol{x}'|)\delta(\boldsymbol{x}')d^3\boldsymbol{x}', \tag{3.24}$$

where W_R can be a Gaussian window:

$$W_R^{\text{Gaussian}}(r) = \frac{1}{(4\pi/3)R^3} \exp\Big\{-r^2/(2R^2)\Big\}, \tag{3.25}$$

or a top-hat window:

$$W_R^{\text{Top-hat}}(r) = \frac{1}{(2\pi)^{3/2}R^3} H(R - r), \tag{3.26}$$

where $H(x)$ is the Heaviside step function[18]. The fourier transform of both functions would be:

$$W_R^{\text{Gaussian}}(k) = \exp\Big\{-k^2R^2/2\Big\}, \tag{3.27}$$

$$W_R^{\text{Top-hat}}(k) = \frac{3j_1(kR)}{kR}, \tag{3.28}$$

where $j_1(x)$ is the spherical Bessel function of order one.

With all these relations, it is possible to derive a fundamental quantity: σ_8. The autocorrelation function $\xi(0)$ will be equivalent to:

[18] $H(x) = 1$ for $x > 0$ and $H(x) = 0$ for $x < 0$.

$$\begin{aligned}\xi(0) &= \lim_{r\longrightarrow 0}\int_0^\infty P(k)\sin(kr)kdk \\ &= \int_0^\infty \Delta^2(k)|W_R(k)|^2 d\log k,\end{aligned} \tag{3.29}$$

where $\Delta^2(k) = \dfrac{k^3P(k)}{2\pi^2}$. Finally, using the top-hat filter, the above equation would be equivalent to:

$$\sigma^2(R) = \int_0^\infty \Delta^2(k)\left(\frac{3j_1(kR)}{kR}\right)^2 d\log k. \tag{3.30}$$

This quantity is the *rms amplitude (squared) of mass fluctuations* smoothed over the scale R. Historically, the size $R = 8h^{-1}$Mpc have been used in cosmology because it is the typical size of the galaxy clusters:

$$\sigma_8 = \sigma(R = 8h^{-1}\text{Mpc}, z = 0). \tag{3.31}$$

We finally define the *cumulant generating function* (CGF) as:

$$\psi_R(y,\tau) = \sum_{n=0}^{\infty}\frac{\langle\delta_R(\tau)^n\rangle_c}{n!}y^n, \tag{3.32}$$

where we have averaged over the spatial coordinates $\boldsymbol{x}$, and due to homogeneity, it will not have a spatial dependence. The CGF will allow us to derive the PDF $P(\delta_R|\tau)$:

$$P(\delta_R|\tau) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\exp\{-iy\delta_R(\tau) + \psi_R(iy)\}dy. \tag{3.33}$$

A final quantity that can be defined to account for higher-order features of the PDF is the S_n constant [51]:

$$S_n = \frac{\langle\delta_R(\tau)^n\rangle_c}{\{\langle\delta_R(\tau)^2\rangle_c\}^{n-1}}. \tag{3.34}$$

where, in particular, the S_3 and S_4 are proportional to the skewness and the kurtosis, respectively.

3.3. Observations of the cosmic structure

Galaxies and clusters are used to infer the *density fields* of the whole universe and, therefore, study the large-scale structure. It is not possible to measure the total density field $\delta(\boldsymbol{x},\tau)$ neither the power spectrum $P(\boldsymbol{k},\tau)$ directly. Usually, we can observe the light emitted from the interactions of baryonic matter (specifically the galaxy density field) or the effects of lensing (cosmic shear field) that trace the matter density.

Cosmologists calculate the variance of density fluctuations as a function of scale through the *two-point statistics* to deduce the statistical properties of the density fields. Other higher-order features are also studied, e.g., the skewness, that accounts for the unbalanced evolution of the underdense and overdense regions in the sky.

However, there are statistical and astrophysical challenges to overcome, such as noise in the observations, the redshift estimation of galaxies, and other systematic biases that we will address in Sec. 3.5.

3.3.1. Galaxy density field

One of the main supports of the theory of the cosmic structure is the observation of the *galaxy density field* in the sky. Depending on the type of survey, the data might contain the spectroscopic or photometric redshift, respectively.

The *spectroscopic redshift* z_s is measured by identifying a shift of the emission and absorption lines spectra for atoms well studied from Earth. On the other hand, *photometric redshift* z_p is commonly estimated from the magnitudes of the observed objects in different chromatic filters. In the latter, an observed distribution of galaxies will have a specific redshift range: $n_g(z_p)$. We can think of this as the distribution of galaxies in some area or volume, restricted to z_p.

It is also common to define the galaxy density contrast, similar as in Sec 3.1.2:

$$\delta_g(\boldsymbol{x}, z) = \frac{n_g(\boldsymbol{x}, z) - \overline{n}_g}{\overline{n}_g}, \tag{3.35}$$

where $\overline{n}_g$ is the mean galaxy density field.

3.3.2. Galaxy bias

As mentioned in Sec. 3.3, there is a systematic bias when trying to trace δ from δ_g. Galaxy clustering can be used as a cosmological probe when the mass-to-light ratio is modeled. Theory predicts that the halos of dark matter were initially originated from peaks of higher density as illustrated in Fig. 3.3.

There are other complexities that arise from mixing a discrete field with a continuous one, or from the selection process: a galaxy might or not appear in a survey due to its optical properties. From an astrophysical perspective, galaxy formation is a complicated process, and locally, the density of galaxies is non-linear: massive galaxies will preferentially form in a dense environment, amplifying the universe's matter density contrast. Therefore, galaxies are stochastic tracers and are biased when trying to trace the matter field.

The systematic mismatch is usually approached with a linear approximation:

$$\delta_g \approx b \cdot \delta, \tag{3.36}$$

where the factor b is called the linear *galaxy bias* and is only valid on large scales. There is an additional component associated with the random effects arising from smaller scales to the larger ones to consider: *the galaxy stochasticity*. In [1, 2], this intrinsic shot-noise of the distribution of galaxies is considered in two possible models that we do explore in the primary research from this thesis, shown in Chapter 7.

Given some number density of galaxy $n(z)$ centered at a redshift distribution z, and assuming

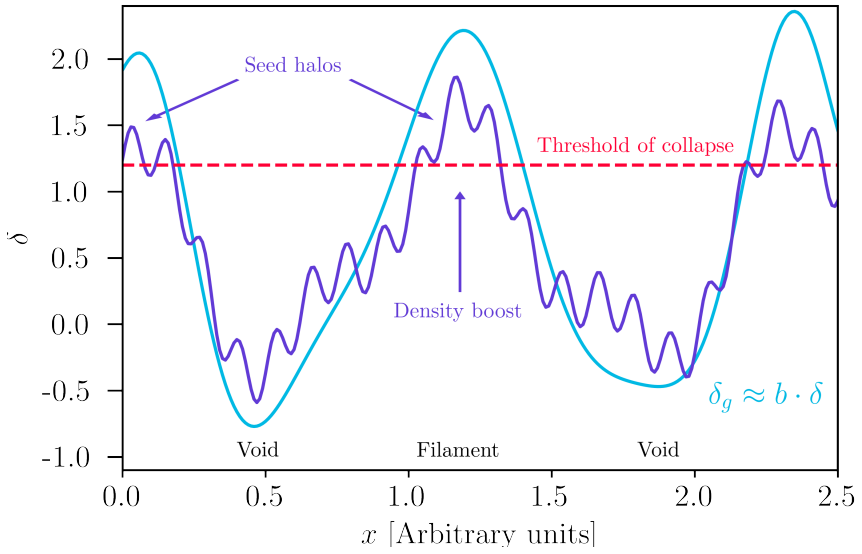


Figure 3.3: Illustration of the galaxy bias. A long wave fluctuation present a density increase, called *density boost*, which surpass the threshold of collapse. Galaxies are believed to form in the small peaks from the zones of high density, being their smoothed density contrast δ_g augmented with respect to δ. This is parameterized with the bias parameter b. Figure adapted from Fig. 1 from [16] and Fig. 9.1 from [48].

isotropy, we can define the line-of-sight projection of the 3D density contrast δ_{2D} as:

$$\delta_{2D} = \int q(w)\delta(w)dw, \tag{3.37}$$

where w is the co-moving distance corresponding to a redshift z and the projection kernel $q(w)$ is given by:

$$q(w) = n(z[w])\frac{dz[w]}{dw}. \tag{3.38}$$

We can also define the average density field δ_m as the average δ_{2D} field over some filter. In Eq. II.3 from [2], an average density field is defined using a top-hat filter with aperture radius θ_T. In Sec. 3.5 the field δ_m is extensively used due to the observational limitations.

3.3.3. Higher-order statistics

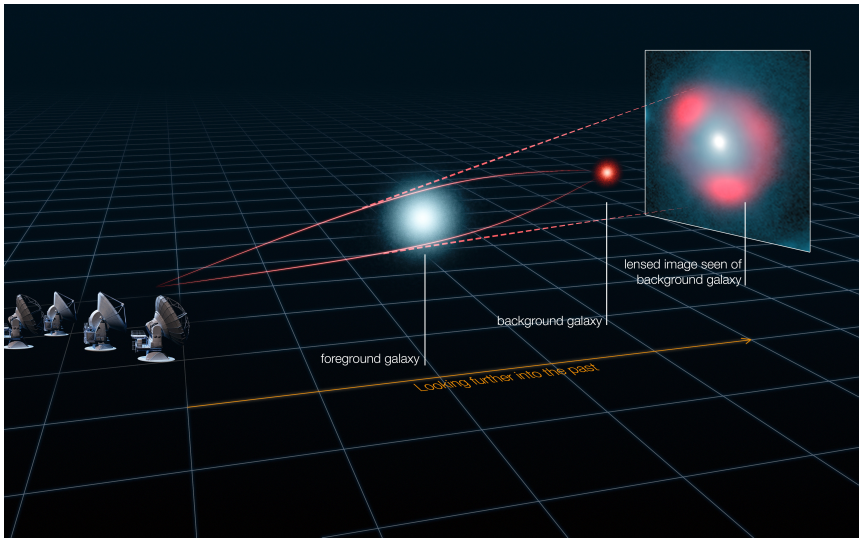
As mentioned earlier, two-point correlation function measurements only allow us to recover the first two moments of the PDF. Higher-order moments do not necessarily return more complete estimates on the cosmological parameters. Still, they enable us to break degeneracies with two-point statistics and explore the properties of galaxies and matter density. Alternative methods exist to explore the universe's large-scale structure on non-linear scales. These non-linearities can be studied when computing the three-point correlation function or when observing, e.g., peak statistics and the cluster mass function [47].

3.3.4. Gravitational lensing

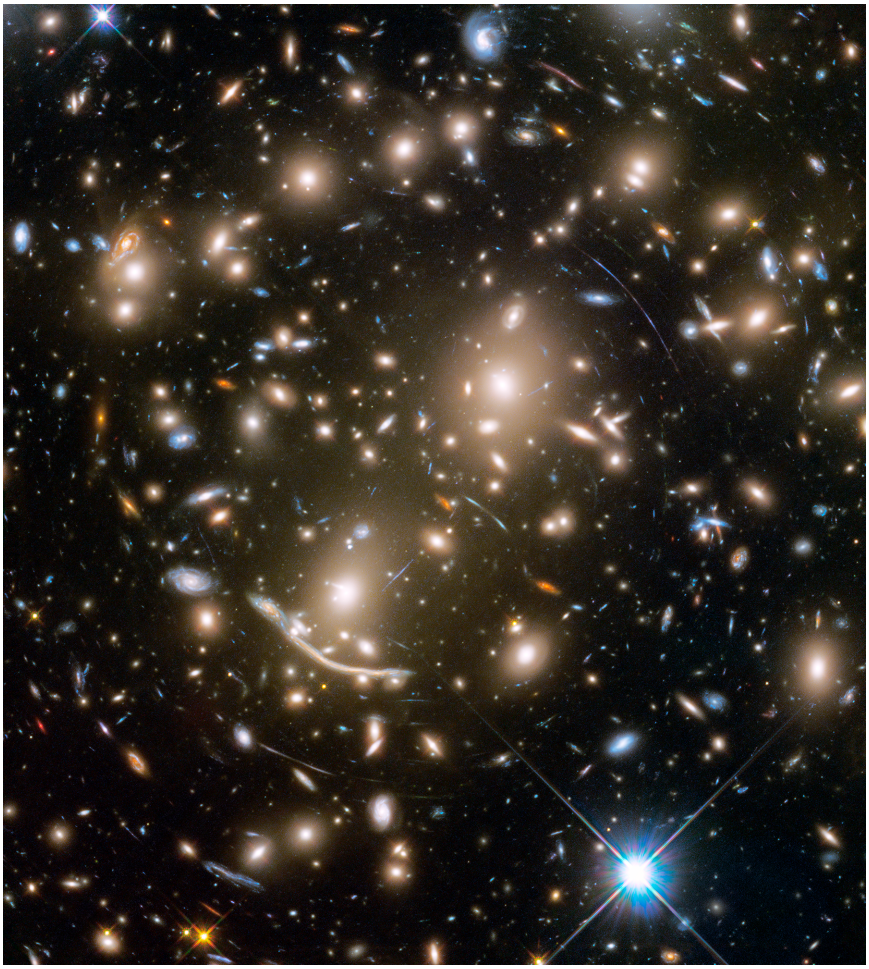
Due to the gravitational tidal forces that the large-scale structure generates, light traveling from distant sources deflects, virtually distorting the images from the background. That effect can be visualized in the upper panel from Fig. 3.4. The light gets *deflected* (i.e. the source changes its position), the observation gets *magnified* and *sheared* (i.e. the shapes gets distorted). These three effects can be appreciated in the lower panel from Fig. 3.4. Gravitational lensing can be either faint or intense, receiving the names of *weak lensing* and *strong lensing* respectively. In the later, images of individual light sources end at distinct positions or stretched into large arcs.

We can analyze patterns from the deformed galaxies and estimate the intervening total matter density field (both baryonic and dark matter). The gravitationally distorted galaxy field is called the *cosmic shear field*, which include all the galaxies observed because the shear effect is present in the whole line of sight due to the presence of intermediate matter. The measurement of this weak lensing effect averages the shear on a set of galaxies, considering assumptions like the mean shape of galaxies (elliptical) and their orientation (random). If we measure deviations from the random orientations, we will be under the presence of a lensing effect. See [52] for a theory review with applications of weak lensing.

As done with the galaxy field, cosmologists have applied the two-point correlation function to the cosmic shear signals [53–55] and then combine the analysis with the information with the galaxy density field through a cross-correlation function [15, 56].



(a) Illustration of how gravitational lensing affects the background galaxy sources. Image credit: ALMA (ESO/NRAO/NAOJ), L. Calçada (ESO), Y. Hezaveh et al.



(b) Abell 370 cluster of galaxies generating gravitational lensing due to the hundreds of galaxies that compose it. Credits: NASA, ESA, R. Bouwens and G. Illingworth (University of California, Santa Cruz)

Figure 3.4: Explanation and example of the gravitational lensing effect.

3.4. Sky surveys

A *sky survey* is a map of a section of the sky that establishes all its measurable contents [13]. With the evolution of astronomical science and instrumental experiments, these observations have provided fundamental information about the universe's contents, galaxy formation, and evolution. The data is in the format of images and spectra of galaxies, with some redshift information (photometric or spectroscopic), providing a three-dimensional map of the universe. In addition, observations contain information encoded in the characteristics of galaxies, such as their positions, colors, and shapes.

In 2006, the *Dark Energy Task Force report* [57] defined the Stage-III and Stage-IV classification of galaxy surveys as the ongoing (started in the 2010s) and future (will begin in the 2020s) dark energy experiments, respectively.

For Stage-III galaxy surveys, the data has been taken in the last decade and it is still generating production of articles and discussions among the community. Several experiments were displayed all over the world, including the *Dark Energy Survey* (DES) [58, 59], the *Kilo-Degree Survey* (KiDS) [60], the *Hyper Suprime-Cam* HSC [61] and the *Extended Baryon Oscillation Spectroscopic Survey* (eBOSS) [62].

On the other hand, Stage-IV surveys are generating large expectations due to the unprecedented amount of data that will be collected through the telescopes, in addition to the concept of time-domain characteristic they will have. The images and spectra will cover nearly the entire extra-galactic sky. Some of these are already observing the large-scale structure, such as the *Dark Energy Spectroscopic Instrument* (DESI) [63]. There are others that will start to operate relatively soon, such as the *Legacy Survey of Space and Time* (LSST) at the Vera Rubin Observatory [64, 65], *Euclid* [66] and the *Nancy Grace Roman Space Telescope* [67].

3.4.1. The Dark Energy Survey

In particular, in Section 3.5 we present a framework that is based on the data products generated from the Dark Energy Survey Year 1 Results (DES Y1) [68] and in Chapter 7 we work on them directly.

Originally, DES mapped a total of 5000 deg^2 of the sky ($\sim 1/8$ of the whole sky, see Fig. 3.5) using optical imaging observations of a total of 525 nights at the Blanco Telescope in the Cerro Tololo Inter-American Observatory in Chile, using a digital camera with 570 Megapixels in 62 CCD's called DECam. This task took 6 years (2013-2019), and it provides the redshift and estimate of shapes of a total of 300 million galaxies up to $z \sim 1.5$ in five filters ($grizY$), being the most extensive connecting map of the universe. It is composed of more than 400 scientists from over 25 institutions grouped in a collaboration called the DES collaboration[19]. The different combined probes, such as galaxy clustering and weak gravitational measurements of three two-point correlation functions, have provided precise constraints on the cosmological parameters. However, it also has exposed that systematic effects must be considered in the Stage-IV sky surveys.

[19] https://www.darkenergysurvey.org/

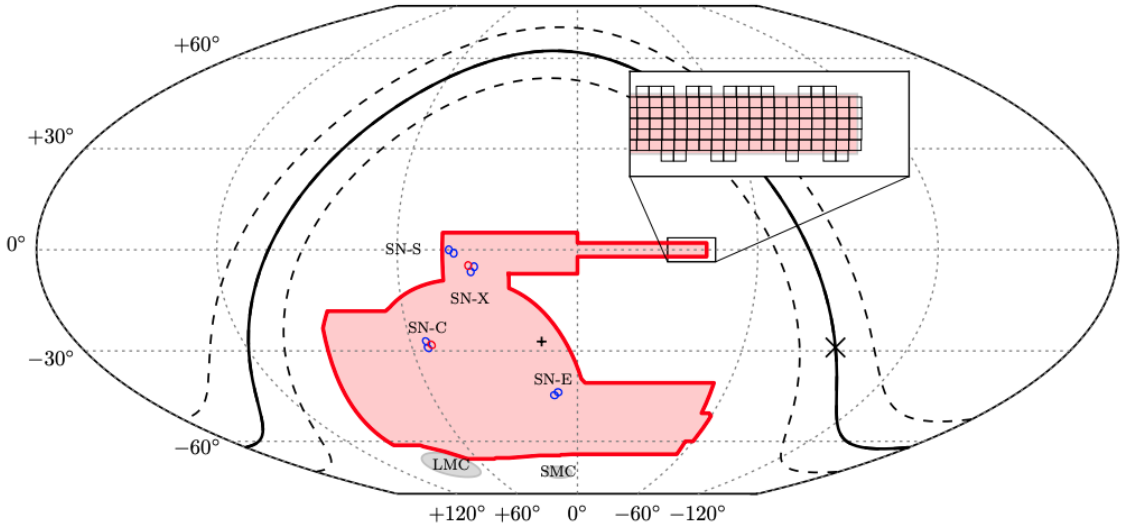


Figure 3.5: Sky-map plot including the DES footprint in equatorial coordinates. Image from [15].

3.5. Density split statistics

The *density split statistics* (DSS) is a method proposed by [1, 2] to infer the matter density field, the skewness of the PDF and the galaxy bias from observations of the gravitational lensing, and the count of galaxies in cells from DES Y1 data. It derives two-point and three-point statistics of the large-scale cosmic structure by splitting the sky (on the line of sight) in subareas of different galaxy density and studies how light from farther galaxies get distorted due to lensing physics.

Throughout Sec. 3.5, we will resume the methodology and main results of the DSS, giving a particular focus to the shot-noise parametrizations used on the galaxy bias. In Sec. 7, we will answer one of the open questions from the DSS: the derivation of informed priors of the shot-noise models to derive more certain cosmological constraints.

3.5.1. Methodology

The methodology of the density split statistics can be summarized in three steps, as presented in [2]:

1. **Split of the observed foreground galaxies in different densities:**

   The first step is to separate a sample of low redshift galaxies with distribution $n_l(z)$, denoted as *foreground*, by the spatial number density in a sub-areas of equal size. This is done by smoothing the position and counting the number of galaxies N_T inside a circular aperture, in particular, a top-hat aperture of radius $\theta_T = 20'$. Different parts of the sky will give a distinct value for N_T. Then, the sky is divided into regions of different galaxy number density in quintiles of growing density. (See the top panel from Fig. 1 of [2] for an illustration.)

2. **Use gravitational lensing to trace the matter density field in each sky quintile of the foreground sample:**

   The second step is to consider a second sample of galaxies, denoted as *background*, with higher redshift and study how the light emitted from them suffers gravitational lensing effects when passing through the foreground sample. For each galaxy density quintile, the tangential shear is measured, and a *density split lensing signal* is obtained to trace the foreground matter density field. The advantage of dividing the total signal is that it allows detecting the skewness present in the density PDF and the overall underdensity of the observed universe. (See the right panel from Fig. 1 of [2] to visualize a lensing signal split by the quintiles of galaxy density.)

3. **Obtain the mean counts-in-cells per density quintile:**

   Cosmological constraints derived from the lensing signals can be improved when connecting the galaxies observed in the foreground sample with the density field inferred from the cosmic shear field. Three different bias models are proposed in [2], from which two account for galaxy stochasticity. We will analyse these models at Sec. 3.5.2 and Sec. 3.5.3.

Therefore, the main question that DSS aims to answer is *what would be the matter density contrast $\delta_{m,T}$ given a number of galaxies N_T found around a line-of-sight?* Or equivalent, *what would be $P(\delta_{m,T}|N_T)$?* This can be answered using the Bayes' theorem (see Sec. 4.3.2 for an explanation):

$$P(\delta_{m,T}|N_T) = \frac{P(N_T|\delta_{m,T})P(\delta_{m,T})}{P(N_T)}, \tag{3.39}$$

where $P(N_T|\delta_{m,T})$ is the probability of finding N_T galaxies given $\delta_{m,T}$, $P(\delta_{m,T})$ is the PDF of matter density contrast and $P(N_T)$ is the probability of having N_T galaxies in the light of side. In the original DSS work, different models were propose to describe $P(\delta_{m,T})$, the expected value of the lensing convergence signal given the density contrast $\langle\kappa|\delta_{m,T}\rangle$ and $P(N_T|\delta_{m,T})$. We will focus on the latter and refer to chapters IV.A and IV.B from [2] for an explanation of the first two.

The conditional probability $P(N_T|\delta_{m,T})$ models the probability of finding N_T galaxies in a circular aperture of radius θ_T given the projected density contrast $\delta_{m,T}$ in that area. In [69] it is shown that for a simple linear bias connection for the matter and density contrast (Eq. 3.36) and pure Poissonian shot-noise, $P(N_T|\delta_{m,T})$ would be equivalent to:

$$P(N_T|\delta_{g,T}) = \frac{[\overline{N}(1+\delta_{g,T})]^{N_T}}{N_T!}e^{-\overline{N}(1+\delta_{g,T})} = \frac{[\overline{N}(1+b\delta_{m,T})]^{N_T}}{N_T!}e^{-\overline{N}(1+b\delta_{m,T})}, \tag{3.40}$$

where $\overline{N} = \langle N_T\rangle$ and b being the only free parameter. This model implies that:

$$\langle N_T|\delta_{g,T}\rangle = \overline{N}(1+\delta_{g,T}), \tag{3.41}$$

and

$$\mathrm{Var}[N_T|\delta_{g,T}] = \langle N_T|\delta_{g,T}\rangle. \tag{3.42}$$

In the DSS, two variations are proposed to extend this relation into higher-orders:

3.5.2. Shot-noise model 1: joint log-normal distribution

The first model presented in the DSS framework uses the ansatz that *the matter and galaxy fields are joint log-normal random variables* with two free parameters: b (the galaxy bias) and r (the galaxy stochasticity). In particular, for a bivariate log-normal distribution, there will be a total of five parameters describing the joint distribution: the means, the standard deviations, and the covariance of both.[20] In particular, in the DSS framework, the variance and the skewness of the density fields are related as:

$$\langle \delta_{g,T}^2 \rangle = b^2 \langle \delta_{m,T}^2 \rangle, \tag{3.43}$$

$$\langle \delta_{g,T}^3 \rangle = b^3 \langle \delta_{m,T}^3 \rangle \tag{3.44}$$

and

$$\langle \delta_{g,T} \delta_{m,T} \rangle = br \langle \delta_{m,T}^2 \rangle. \tag{3.45}$$

See Sec. IV A 1 and Appendix D from [2] for a formal derivation.

The r parameter can be interpreted as a Pearson correlation coefficient:

$$r = \frac{\langle \delta_{g,T} \delta_{m,T} \rangle}{\sqrt{\langle \delta_{g,T}^2 \rangle \langle \delta_{m,T}^2 \rangle}}, \tag{3.46}$$

for which the case of $r = 1$ will not only recover the original Poisson distribution from Eq. 3.40 but will imply that there is a direct correlation between the fields $\delta_{g,T}$ and $\delta_{m,T}$. This leaves no space for higher-order information or galaxy stochasticity. The general distribution of $P(N_T|\delta_{m,T})$ will be given by the numerical evaluation of

$$P(N_T, \delta_{m,T}) = \int P(\delta_{g,T}|\delta_{m,T}) P(N_T, \delta_{g,T}) d\delta_{g,T}, \tag{3.47}$$

where $P(\delta_{g,T}|\delta_{m,T})$ can be derived from the joint log-normal distribution derivation as done in Appendix D from [2].

3.5.3. Shot-noise model 2: super-Poissonianity

The second model is an extension of the Poisson distribution from Eq. 3.40. It allows stochasticity to be more flexible than the previous shot-noise model because of the addition of two stochasticity parameters. In total, the model of super-Poissonianity has three free parameters: b, α_0 and α_1 (two galaxy stochasticities) that arise from allowing the following extensions in Eq. 3.42:

$$\mathrm{Var}[N_T|\delta_{g,T}] = \langle N_T|\delta_{g,T} \rangle \cdot \alpha(\delta_{m,T}), \tag{3.48}$$

where

$$\alpha(\delta_{m,T}) = \alpha_0 + \alpha_1 \cdot \delta_{m,T}. \tag{3.49}$$

[20] One interesting feature is that the skewness can be derived from the listed quantities. See Sec. 4.2.2 for a summary of this distribution.

This will impact the general distribution as:

$$P(N_T|\delta_{m,T}) \propto \frac{1}{\alpha(\delta_{m,T})} \times \frac{[\overline{N}(1+b\delta_{m,T})/\alpha(\delta_{m,T})]^{N_T/\alpha(\delta_{m,T})}}{(N_T/\alpha(\delta_{m,T}))!} e^{-\overline{N}(1+b\delta_{m,T})/\alpha(\delta_{m,T})}, \tag{3.50}$$

where we have set a proportional relation because it must be normalized in case α is not an integer. In particular, if $\alpha_0 = 1$ and $\alpha_1 = 0$, we recover the original Poissonian distribution from Eq. 3.40.

3.5.4. Priors assigned to the shot-noise models

The two models listed in Sec. 3.5.2 and in Sec. 3.5.3 introduce two and three free parameters to the general DSS likelihood analysis, respectively. The Bayesian statistical inference performed in the DSS requires restrictions summarized in their corresponding prior distributions. (See Sec. 4.3.2 for an introduction to Bayesian probabilities and the corresponding prior distributions).

In the shot-noise model 1, the originally assigned priors are:

$$P(b) = \mathcal{U}(0.8, 2.5) \text{ and } P(r) = \mathcal{U}(0, 1), \tag{3.51}$$

where $\mathcal{U}$ denotes an uniform distribution (see Sec. 4.2.2 for a definition). The prior on b was determined from MCMC analytic posterior runs. In contrast, the prior on r is motivated by the fact that matter and galaxy densities must be positively correlated, and therefore $0 \leq r \leq 1$. If the universe shows low stochasticity, the value of r would be close to the unity.

In the shot-noise model 2, the originally assigned priors are:

$$P(b) = \mathcal{U}(0.8, 2.5),\ P(\alpha_0) = \mathcal{U}(0.1, 3.0) \text{ and } P(\alpha_1) = \mathcal{U}(-1.0, 4.0). \tag{3.52}$$

The priors on the galaxy stochasticity parameters α_0 and α_1 are derived from the following arguments: the lower bound in α_0 is not 0.0 because of computational costs, and the other prior bounds are determined after visualizing the $2-\sigma$ confidence region of Buzzard simulations of the DES Y1 like signals. [70]

One of the weaknesses of the priors assigned to the α parameters is that they are *mildly informative* [2]. (See final discussion at Sec. 4.3.2.) The same work states that stronger priors can be motivated and additional articles are enumerated with findings on physical bounds for both α_0 and α_1.

3.5.5. Results from DSS

The considered cosmological parameters to test the DSS were Ω_m and σ_8 (and therefore S_8). The skewness of the matter density field was included in an additional parameter S_3:

$$S_3 = \frac{\langle \delta_m^3 \rangle}{\langle \delta_m^2 \rangle^2}, \tag{3.53}$$

which can be straightly obtained from Eq. 3.34 by setting $n = 3$.

In addition to the already mentioned tracer galaxies parameters (the already presented b, r, α_0, α_1 in the shot-noise models), there were more nuisance quantities: parameters describing photometric redshift and shear biases (see Table I from [1] for a complete list).

The posterior distribution was obtained through an MCMC sampling in the already mentioned parameter space. The final marginalized constraints of the cosmological parameters for the different shot-noise parametrizations of the DSS are summarized in Fig. 10 from [1], while the corresponding confidence limits are summarized in Table II from the same publication. Both shot-noise models described in a similar way the extension of the linear behavior (and the super-Poissonian shot-noise) as a function of the matter overdensity (see Fig. 8 from [2]). The constraints on Ω_m from both shot-noise models agrees ($\Omega_m = 0.26^{+0.04}_{-0.03}$ and $\Omega_m = 0.28^{+0.05}_{-0.04}$ for the (b, r) and (b, α_0, α_1) models respectively), but there is a divergence in the final estimation of σ_8 when analyzing DES data ($\sigma_8 = 0.97^{+0.07}_{-0.06}$ and $\sigma_8 = 0.80^{+0.06}_{-0.07}$ respectively).

One of the remarkable findings from the DSS is that *DES data tends to prefer shot-noise models beyond the linear bias*. Nevertheless, some improvements can be made in the prior modeling of the stochasticity parameters. In Chapter 7 we address the first of the listed future research tasks suggested in Sec. VII. from [1]: *The derivation of informed priors for the shot-noise nuisance parameters that decrease the uncertainties of the cosmological parameters*. To reach this goal, we do a thorough analysis of the probabilistic and statistical considerations that must be taken into account in Chapters 4 and 5 respectively.

Chapter 4

Basics of probabilities

Astronomy has been the discipline that has carried the primary advances in probability and statistics. From *Hipparchus* [190-120 BC], who suggested the use of the mid-range quantity to estimate the length of the year [71], to the modern use of artificial intelligence algorithms to classify real-time photometric alerts in unprecedented sky surveys. Cosmology has been one of the sub-fields that has caught more attention in the last hundred years due to advances in the conceptual background and the discovery of independent cosmological observables [72–74].

Until recently, independent cosmological proofs have tended to converge in the standard theory of ΛCDM Model that would have explained with great confidence the evolution and contents of our universe, as introduces in Chapter 2. Though the respective errors have decreased with the perfectionism of the analysis and augmentation of available data, there are systematics errors that have led to notable tensions (see Section 5.6 for a detailed example). One question that needs to be asked is whether our theory is inaccurate or how we extract and use the data is not optimum.

A cosmological model could not be sufficient to explain observations due to two reasons. First, the reality could be different from what the theory predicts: it could be more or less complex. The second reason would be that the theory could be correct, but we cannot overcome the present selection biases and systematics when collecting the data.

For millennia, astronomical understanding has oscillated between overcoming these barriers for some period to finding other inconsistencies due to the increase in the quantity and quality of data. In this continual seesawing among conciliation and decay of concepts, theories had to choose how to increase their complexity while avoiding over-fitting the data. Nowadays, there is a fundamental difference: the exponential rate of increment of data and the evolution of theories. In the 21st-century, researchers might need to change their scientific conclusions after a couple of years. The shrinking of these periods of acceptance-rejection requires us to change our conception of making cosmological inferences. In that line, we propose and choose to adopt the following perspectives:

We will first adopt a *pessimist* view of the statistical methodology in cosmological inference. For the following chapters, we will assume that models are anything but useful inventions: underlying causal relationships can not be discovered uniquely through the study of patterns

in the observable phenomena. We extract our conception from the following two statements:

"[Statistical] models can provide us with ideas which we test against data, and about which we build up experience. They can guide our thinking, lead us to propose courses of action, and so on, and if used sensibly, and with an open mind, and if checked frequently with reality, might help us learn something that is true. Some statistical models are helpful in a given context, and some are not. (...) What we do works (when it does) because it can be seen to work, not because it is based on true or even good models of reality." (Speed 1992, addressing a meeting of astronomers) [75]

"The object [of statistical inference] is to provide ideas and methods for the critical analysis and, as far as feasible, the interpretation of empirical data. (...) The extremely challenging issues of scientific inference may be regarded as those of synthesising very different kinds of conclusions if possible into a coherent whole or theory. (...) The use, if any, in the process of simple quantitative notions of probability and their numerical assessment is unclear..." (Cox 2006) [75]

In Sec. 4.1 we will define the fundamental axioms of the probability theory necessary to distinguish hereafter the types of probabilities at Sec. 4.2. Finally, we end up presenting the two main interpretations that have been used in the last centuries: the frequentist and Bayesian concepts in Sec. 4.3.

4.1. The basics of probabilities

Even if a deterministic physical arrangement governs the interactions inside the universe, we cannot access that information without some degree of uncertainty. The probability theory provides us with tools to model the uncertainty of a given experiment.

4.1.1. Outcome spaces and events

For example, we cannot predict the redshift of a galaxy before its measurement, although our prior knowledge of the photometric redshift distribution dispersion will constrain the prediction. In this *experiment*, in which the result cannot be predicted with certainty a priori, the set of possible redshifts will account for the *outcome space*. One possible *event*, in this example, would be the resulting galaxies that populate the redshift interval [0.8, 1.0]. In Table 4.1, we define these concepts in further detail.

Probabilities are concerned with assigning numerical values to the degree of likeliness of an event happening. These follow three primary axioms.

4.1.2. Axioms of probability

1. To each event A is assigned a non-negative real number $P(A)$, such that

$$0 \leq P(A) \leq 1, \tag{4.1}$$

where 0 indicates impossibility of the event and 1 indicates certainty.

Table 4.1: Definitions of basic statistics adapted from Feigelson & Babu (2012) [75]

Term	Definition
Random Experiment	Any action that can have a set of possible results where the actually occurring result cannot be predicted with certainty prior to the action.
Sample Space (Ω)	Set of all outcomes of an experiment.[1]
Event (A)	Subset of the outcome space ($A \subset \Omega$).
Random Variable (X)	It is function that assigns numerical values to each of an experiment's outcomes. It can take either discrete or continuous values.
Range of a Random Variable	It is the set of possible values that a X can take.[2]
Probability ($P(A)$)	Measurement of how likely the event A is to occur.

[1] These sample spaces can be either finite or infinite. One example of a discrete sample space would be the number of supernovae detected by a telescope. In contrast, a continuous sample space would be the possible measured temperature of the Cosmic Microwave Background.
[2] An event A is equivalent to a subset of a range of X.

2. The probability of occurrence of at least one of the elements in the sample space is

$$P(\Omega) = 1. \tag{4.2}$$

3. For mutually exclusive and pairwise disjoint events $A_1, A_2, ..., A_n$, the probability of either occurring is denoted as $P(A_1 \cup A_2 \cup ... \cup A_n)$ and:

$$P(A_1 \cup A_2 \cup ... \cup A_n) = P(A_1) + P(A_2) + ... + P(A_n). \tag{4.3}$$

From these, we can define the *complementary event* A^c, which is the event that A does not occur, and derive its probability. These are *mutually exclusive*, that is, they cannot both occur at the same time, and *exhaustive* (i.e. $A \cup A^c = \Omega$). Therefore, the probability of a complementary event must be given by:

$$P(A^c) = 1 - P(A). \tag{4.4}$$

Another quantity that one can define from these axioms is the *conditional probability* of one event given the occurrence of another. For example, let B be the known event, then the conditional probability of A given B is written as $P(A|B)$ and is equivalent to the fraction of probability B that intersects with A:

$$P(A|B) = \frac{P(A \cap B)}{P(B)}, \tag{4.5}$$

where $P(A \cap B)$ is called the intersection of two events A and B or the *joint probability* of both events happening simultaneously.

We suggest reading [76] for more implications of these axioms.

4.2. Types of probabilities

We can look further into the definition of probabilities depending on the nature of the considered random sample. A *probability distribution* is a function that assigns probabilities to the values that a random sample can take [77]. The random sample can be either discrete or continuous.

4.2.1. Discrete probability distributions

For a *discrete random variable* X_D we will have a finite[21] number of values $x_1, x_2, ..., x_n, ...$ The probability distribution that describes the probability of X_D taking the value x_n is given by:

$$P(X_D = x_n). \tag{4.6}$$

Another important function is the *discrete cumulative probability distribution*, which is the probability of obtaining a value equal to or smaller than x_i:

$$P(X_D \leq x_n) = \sum_{i=0}^{n} P(X_D = x_i). \tag{4.7}$$

Two classic examples of discrete probability distributions observed in cosmology are the Binomial and Poisson distributions.

Binomial distribution:

The *Binomial distribution* represents the number of successes n in a two possible outcomes experiment of success or failure in N trials. Each of these realizations yields success with probability p, and all the trials are independent. Let $X_D \sim \text{Bin}(n, p)$, its probability distribution is given by:

$$P(X_D = n) = \binom{N}{n} p^n (1-p)^{N-n}, \tag{4.8}$$

for $n = 0, 1, 2, ..., N$.

It is common to observe this distribution when selecting samples from survey data. One example would be determining the probability distribution of selecting galaxy clusters containing a dominant central galaxy.

Poisson distribution:

The *Poisson distribution* models the distribution of randomly distributed, independent, point-like events. It expresses the probability of some occurrences n of an event in a fixed interval (in observational cosmology, usually in time or space) if these events occur with a known

[21] Or countable infinite.

constant mean rate λ independent of the last occurrence. Let $X_D \sim \text{Poisson}(\lambda)$, its probability distribution is given by:

$$P(X_D = n) = \frac{\lambda^n}{n!} e^{-\lambda}, \tag{4.9}$$

for $n = 0, 1, 2, ..., N$.

It is possible to demonstrate that, under certain conditions, the Poisson distribution can be derived from the Binomial distribution when $p \longrightarrow 0$ and $Np \longrightarrow \lambda$. (See [78] for an explicit derivation)

The Poisson distribution is of particular importance in astronomy. It is present in various sub-fields, from the *photon noise* [79] to the *two-point correlation shot-noise of the distribution of galaxies at large scales*. (See Sec. 3.5 for an example of the Poissonian shot-noise parametrization of the galaxy stochasticity).

4.2.2. Continuous probability distributions

For a *continuous random variable* X_C, which can take any value in $\mathbb{R}$, the probability of obtaining a single value is always null: $P(X_C = x) = 0$ due to the density of X_C. However the probability between intervals is given by the integral of the *probability density function* $f(x)$:

$$P(a \leq X_C \leq b) = \int_a^b f(x)dx \tag{4.10}$$

for any a, $b \in \mathbb{R}$, such that $a \leq b$.

There is also a *continuous cumulative distribution* (analogue to Eq. 4.7), which is given by:

$$F(x) = P(X_C \leq x) = \int_{-\infty}^{x} f(x')dx'. \tag{4.11}$$

A probability density function will have the following properties:

- $f(x) \geq 0$, $\forall x \in \mathbb{R}$
- $\int_{-\infty}^{\infty} f(x)dx = 1$

The Uniform, Gaussian, and Chi-squared distributions are three essential examples of continuous probability distributions observed in cosmology. We also briefly present the Log-normal distribution.

Uniform distribution:

The *Uniform distribution* describes an experiment where the probability of obtaining an outcome from a fixed interval $[a, b]$ is constant. Let $X_C \sim \mathcal{U}(a, b)$, its probability density function is given by:

$$f(x; a, b) = \begin{cases} \frac{1}{b-a} & \text{for } a \leq x \leq b, \\ 0 & \text{otherwise} \end{cases}. \tag{4.12}$$

Gaussian distribution:

The *Gaussian distribution* represents one of the most common observed distributions in nature. It can be derived using the *central limit theorem* (see [77] for a derivation). Let $X_C \sim \mathcal{N}(\mu, \sigma)$, its probability density function is given by:

$$f(x; \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \tag{4.13}$$

where μ would be the mean and σ^2 the variance. It is also common to find its multivariate version, when the random variable sample is obtained from a vector ($\mathbf{X_C} = (X_1, \ldots, X_k)^{\mathrm{T}}$), such that $\mathbf{X_C} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, where $\boldsymbol{\mu}$ would be the mean vector and $\boldsymbol{\Sigma} = \langle \mathbf{x}\mathbf{x}^{\mathsf{T}} \rangle - \langle \mathbf{x} \rangle \langle \mathbf{x}^{\mathsf{T}} \rangle = \langle (\mathbf{x} - \boldsymbol{\mu})(\mathbf{x} - \boldsymbol{\mu})^{\mathsf{T}} \rangle$ the covariance matrix. Its probability density function is given by:

$$f(\mathbf{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{1}{\sqrt{\det(2\pi\boldsymbol{\Sigma})}} e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^{\mathsf{T}}\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})}. \tag{4.14}$$

The Gaussian distribution is often observed in the noise of the observations and in the features of the cosmological density fields (such as the cosmic microwave background, the large-scale structure, etc.).

Chi-squared distribution:

The *Chi-squared distribution* appears usually in science because it can be derived from the addition of k independent random variables $Z_1, \ldots, Z_k$ such that $Z_i \sim \mathcal{N}(0, 1)$, $\forall i = 1, 2, \ldots, k$. The addition of these random variables:

$$X = Z_1^2 + Z_2^2 + \cdots + Z_k^2 \tag{4.15}$$

$$= \sum_{i=1}^{k} Z_i^2 \tag{4.16}$$

will follow a Chi-squared distribution of k degrees of freedom ($X \sim \chi_k^2$). Its density function is given by the following expression:

$$f_X(x; k) = \frac{\left(\frac{1}{2}\right)^{\frac{k}{2}}}{\Gamma\left(\frac{k}{2}\right)} x^{\frac{k}{2}-1} e^{-x/2}, \tag{4.17}$$

where $x > 0$ and Γ is the Gamma function.

It is possible to develop further the underlying idea of this distribution. If we have a number of observations X_i (usually called *data vector*), such that $i = 1, ..., n$ is the number of the observables. Despite the distribution of the individual random variable, if n is large, the normalized joint distribution is well described by a Gaussian distribution in the observable space. The addition of these will be represented by

$$\chi_n^2 = \sum_{i=1}^{n} \left(\frac{X_i - \mu_i}{\sigma_i} \right)^2. \tag{4.18}$$

There is a generalization of the Chi-squared distribution that allows it to be *noncentral*. In this case, the k independent random variables $Z_1, \ldots, Z_k$ will be such that $Z_i \sim \mathcal{N}(\mu_i, 1)$, with μ_i their respective mean. Then, the random variable $X = \sum_{i=1}^{k} Z_i^2$ will follow a *noncentral Chi-squared distribution* with k degrees of freedom and a noncentrality parameter λ, which is related to the mean of the random variables Z_i by:

$$\lambda = \sum_{i=1}^{k} \mu_i^2. \tag{4.19}$$

The probability density function is given by

$$f_X(x; k, \lambda) = \sum_{i=0}^{\infty} \frac{e^{-\lambda/2}(\lambda/2)^i}{i!} f_Y(x; k + 2i), \tag{4.20}$$

where Y is distributed as Chi-squared with $k + 2i$ degrees of freedom.

Log-normal distribution:

The *log-normal distribution* corresponds to the logarithm of a variable has a normal distribution, i.e. $\log(X_C) \sim \mathcal{N}(\mu, \sigma)$. Its probability density function is given by:

$$f(x; \mu, \sigma) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left(-\frac{(\log x - \mu)^2}{2\sigma^2}\right) \quad x > 0. \tag{4.21}$$

One of the main features of this distribution is that its skewness ($(e^{\sigma^2}+2)\sqrt{e^{\sigma^2}-1}$) is directly related to its standard deviation ($\sqrt{e^{\sigma^2}(e^{\sigma^2}-1)}$) [80]. In Sec. 3.5 it is used its multivariate version, which can be obtained from its joint probability density function:

$$f(\mathbf{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{1}{(x_1 \cdot \ldots \cdot x_n)\sqrt{(2\pi)^n \det(\boldsymbol{\Sigma})}} e^{-\frac{1}{2}(\log(\mathbf{x})-\boldsymbol{\mu})^{\mathsf{T}}\boldsymbol{\Sigma}^{-1}(\log(\mathbf{x})-\boldsymbol{\mu})}, \tag{4.22}$$

where $\mathbf{x} = (x_1, ..., x_n)$ and the other parameters are the same as the ones shown in the Multivariate Gaussian distribution from Eq. 4.14.

4.3. Interpretation of the probabilities

We can interpret probabilities from two main approaches: the frequentist and the Bayesianist perspective of probabilities. The first one focuses on probabilities representing frequencies of repeatable random experiments, while Bayesians use probabilities to describe their partial knowledge of a fixed parameter.

4.3.1. Frequentist probability

Let A be an event in the outcome space Ω of some experiment. The probabilities are thought of as the number of times A occurs over the total number of trials N in the limit of an infinite series of equiprobable repetitions:

$$P(A) = \frac{\#A}{N}, \tag{4.23}$$

where $\#A$ denotes the number of event occurrences A. This relative frequency concept of probabilities aims to describe the properties of the underlying theory rather than the properties of the studied data.

Some of the main criticisms of the frequentist probabilities are the following [81]:

- *Sampling-size effect*: The number of repetitions that must be done in an experiment to achieve the asymptotic properties of the estimators is not precise.
- *Unrepeatable event*: The repetitions of an experiment might not be possible. One clear example is the universe: we only observe one realization.

Indeed, there are inaccuracies in the frequentist formalism that are interpretable and solved by the Bayesian methods, but they also have their own flaws.

4.3.2. Bayesian probability

If probabilities are conceptualized as a measure of the degree of belief about a proposition, Bayes' theorem arises almost naturally using the axioms of probability presented in Section 4.1.2.

Let A and B be two events. From Eq. 4.5, the joint probability of both events equals the conditional probability of A given B times the probability of B occurring on its own, and vice versa:

$$\begin{aligned} P(A \cap B) &= P(A|B)P(B), \\ P(B \cap A) &= P(B|A)P(A). \end{aligned} \tag{4.24}$$

Given $P(A \cap B) = P(B \cap A)$, it is straightforward to derive the Bayes' theorem:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}. \tag{4.25}$$

Let θ be the set of parameters describing a given model $\mathcal{M}$, and let $\mathcal{D}$ be the observed data. The corresponding Bayes' theorem would be:

$$P(\theta|\mathcal{D}) = \frac{\mathcal{L}(\mathcal{D}|\theta)P(\theta)}{P(\mathcal{D})}. \tag{4.26}$$

In Eq. 4.26, we have used $\mathcal{L}(\mathcal{D}|\theta)$ instead of $P(\mathcal{D}|\theta)$, the conditional probability of $\mathcal{D}$ as a function of θ. From Eq. 4.26, we identify four fundamental terms:

- *Likelihood function* $\mathcal{L}(\mathcal{D}|\theta)$ (also written as $\mathcal{L}(\theta)$): it is a function that assigns a probabilistic prediction to the observed data $\mathcal{D}$ to each parameter value θ in the parameter space.
- *Prior distribution* $P(\theta)$: it describes all the previous knowledge for the hypothesis of an uncertain quantity before some evidence is taken into account.
- *Bayesian evidence* (also known as *Marginal likelihood*) $P(\mathcal{D})$: it can be interpreted as the probability of the evidence (data for us) given all possible hypotheses of the model

$\mathcal{M}$ when we extend the *marginalization* as:

$$P(\mathcal{D}) = \sum_{\theta_i} P(\mathcal{D}|\theta_i)P(\theta_i), \tag{4.27}$$

for the discrete case, and

$$P(\mathcal{D}) = \int P(\mathcal{D}|\theta)P(\theta)d\theta, \tag{4.28}$$

for the continuous case for one dimension.

The Bayesian evidence is fundamental for model comparison purposes. Still, for Bayesian statistical inferences, if the data is already available, we do not compute it and simply normalize over the numerator. Therefore, we will focus on the following version of the Bayes' theorem:

$$P(\theta|\mathcal{D}) \propto \mathcal{L}(\mathcal{D}|\theta)P(\theta). \tag{4.29}$$

- *Posterior distribution* $P(\theta|\mathcal{D})$: it represents the state of belief of our model after considering the observed data.

The difference between Eq. 4.25 and Eq. 4.26 is our interpretation of the quantities involved. There is no logical indication of how and when to determine the prior. Therefore, it is subjected to our judgment.

Usually, in the area of cosmology, the likelihood functions, under Gaussian assumptions (see Eq. 4.18), will be:

$$\mathcal{L} \propto \exp\left[-\frac{1}{2}\sum_{i=1}^{n}\left(\frac{X_i - \mu_i}{\sigma_i}\right)^2\right] \equiv \exp\left[-\frac{1}{2}\chi_n^2\right], \tag{4.30}$$

with χ_n^2 being a Chi-squared distribution of n degrees of freedom.

While the Bayesian set has lead to enormous advances in science, it has one big challenge that is discussed below:

- *Subjective nature of priors*

  There is no strict and standard definition of the most suitable prior for a statistical problem. As the prior function aims to describe the knowledge about a quantity before additional evidence is taken into account, all the possible and reliable past information has to be considered and mathematically synthesized. If the priors keep inferences in a reasonable range, they are called *informative.* [22]

  It is also possible that the prior distribution is broad concerning the actual parameter values. In this case, the prior is called *non-informative*. Usually, a large uniform distribution is assigned to a quantity because *where* the actual value lies in the parameter space is unknown. By doing this, all values will be considered equally likely. However,

[22] There are cases in which prior information is not available, or it cannot be directly mapped in a function with *hyperparameters*. In cosmology, this can be overcome when there is an extensive study of the prior knowledge of the astrophysical theory and the empirical studies of the proposed models. However, even if the prior knowledge is available, it may not be obvious how to convert that into a direct prior distribution. It is a current challenge to determine the mapping of the prior knowledge from one model to another.

if a non-linear change of variables is performed, the uniform prior distribution will be uninformative in the second space. In addition, if we also do not have a valid argument for assigning the bounds of a uniform prior, then there would not be a clear argument to set one distribution over another. This can lead to mistaken inferences due to this *overestimation* of the prior distributions. [23]

Recent analyses of cosmological data have shown no consensus on how to constrain a parameter if the prior knowledge is not enough, even with a large data set. In Chapter 7 we provide an alternative solution to define a prior distribution that is updated after demanding to the posterior certain statistical requirements in a simulated likelihood analysis.

4.3.3. Which probabilistic approach should we use?

Both the frequentist and Bayesianist approaches are expected to be equivalent for simple problems and lead to the same results. However, there are cases in which they diverge significantly for more complicated situations. This difference usually relies on the philosophical nature of both families of probabilities. While the first one is fundamentally related to the derived frequencies of events, the second is inherently related to our prior knowledge. There are some examples in the literature in which we observe this divergence. (See [79] for two examples.)

In addition, Bayesian analysis deals with the data that is actually being observed. How can general statements be made if the knowledge is not complete or biased towards a high probability of occurring events? Is it possible to overcome these systematics using a Bayesian perspective? On the other hand, frequentist methods focus on the distribution of potential data that have not been obtained. Deductions should not be based on *what could have happened*. The answer to this crucial point might be hidden in combining these two methods: In [82] the necessity of using both interpretations of probability is described as necessary for future statistical analysis. In [83], and [84], we see cases in which frequentist consistency tests are used to measure the accuracy of the Bayesian methods, and in [85], there is a demonstration of the correspondence between frequentist and Bayesian tests. In Chapter 7 we combine these two approaches to derive more absolute and consistent inferences.

Prior information can also be derived from the underlying theoretical model or from the results of experiment simulations changing initial conditions. Regardless of how they were obtained, the lack of a standard prior function has led to not worrying about the effects these might have in the resulting analysis. There could be two main scenarios: It is possible that the prior considered is *excessively constrained* because the prior knowledge is biased, not enough, or limited. In these cases, priors would be underestimated. In a worst-case scenario, the parameter could be fixed, leading to inherent biases. (See the left panel from 7.1 for an illustrative example).

[23] One example of this is the *prior volume effect*, in which the projection in some subset of the parameter space can lead to biases due to a highly non-Gaussian distribution. (See the right panel from 7.1 for a simple two-dimensional example).

Chapter 5

Statistical inference

As cosmologists, we aim to reach conclusions about the premises and history of our universe. The most significant challenges are to quantify relationships between properties and test if the cosmological observations agree with the theories we have developed and presented throughout this work. Statistical inference uses data analysis to test hypotheses and measure the degree of agreement between models and data. It is also defined as the process by which we infer population properties from sample properties [86].

The statistical inference can be classified in *parametric* or *non-parametric*. Suppose the analysis uses a priori suppositions for the model. In that case, a *parametric* inference will be performed. In contrast, the *non-parametric* uses the data exclusively to infer unknown quantities of interest while making as few assumptions as possible.[24] During this theoretical exploration, we will concentrate on the parametric statistical inference (specifically on frequentist and Bayesian statistical inference to estimate relevant points and intervals) and suggest the reader see [75] for a detailed study on non-parametric inference methods.

Throughout this chapter, we will describe the fundamentals of statistical inference, starting with the Law of large numbers in Sec. 5.1. In Sec. 5.2 and 5.3 we will focus on estimators that allow us to do the statistical inference and we will discuss how to choose the most suitable in Sec. 5.4. We will also present the basics of the sampling methods in 5.5 to obtain more information about the probability distributions. We will finish examining the uncertainties and errors that can appear in statistical inference at Sec. 5.6.

5.1. Law of large numbers

The *law of large numbers* states that if the samples $X_1, X_2, ..., X_N$ of a random variable X are independent and identically distributed random variables, then the sample average converges to some true mean μ as:

$$\frac{1}{N}\sum_{i=1}^{N} X_i \longrightarrow \mu. \tag{5.1}$$

[24] There is also an intermediate approach, the *semi-parametric* inference methods, which uses techniques from both.

We will denote $\hat{\mu} = \frac{1}{N}\sum_{i=1}^{N} X_i$ to the the sample average.

5.2. Point estimators

Let $\boldsymbol{\theta} = (\theta_1, \theta_2, ..., \theta_p)$ be the vector of the parameters that describe a given model. When a measurement takes place, and there is a probability associated, it is possible to summarize information into these interpretable parameters that describe the outcome. This single value is a point in the parameter space and allows us to make inferences about the observed data and the tested model. The *point estimate*, which we will denote $\hat{\boldsymbol{\theta}}$, usually has a deviation with respect to the *true value*[25], which we will call $\boldsymbol{\theta}^*$.

There are many point estimates listed in the literature, but the selection should be made considering the scientific goal of the ongoing experiment. Here we list the three criteria used by Fisher in his famous work *On the Mathematical Foundations of Theoretical Statistics* [87] to determine the quality of a point estimator:

- **Consistency**: *"A statistic satisfies the criterion of consistency, if, when it is calculated from the whole population, it is equal to the required parameter."* (Fisher 1922) We will say that an estimator is consistent if the difference between the estimator and the target population parameter becomes smaller as we increase the samples. A consequence of consistency is *unbiasedness*: An estimator is considered *unbiased* if its expected value is equal to that parameter. We will study this in detail in Section 5.6.
- **Efficiency**: *"The criterion of efficiency is satisfied by those statistics which, when derived from large samples, tend to a normal distribution with the least possible standard deviation"* (Fisher 1922) Therefore, given two unbiased estimators for a parameter, the one with a smaller variance is more efficient.
- **Sufficiency**: *"A statistic satisfies the criterion of sufficiency when no other statistic which can be calculated from the same sample provides any additional information as to the value, of the parameter to be estimated."* (Fisher 1922).

We will now present well known and used point estimators in cosmology:

5.2.1. Method of moments

The *method of moments* is based on the law of large numbers, which allows us to derive equations that relate the *moments of a distribution* μ_r to the unknown parameters $\boldsymbol{\theta}$. Depending on the nature of the considered random variable, these would have a different derivation.

For the discrete case, the non-central moments are given by:

$$\mu_r = \mathbb{E}(X_D^r) = \sum_{i=1}^{N} x_i^r \cdot P(X_D = x_i). \tag{5.2}$$

We can derive the expected value (case $r = 1$), also called mean:

[25] If this last one is actually deductible.

$$\mathbb{E}(X_D) = \sum_{i=1}^{N} x_i \cdot P(X_D = x_i) = \mu_1. \tag{5.3}$$

The central moments are the ones calculated with respect to the mean μ_1:

$$\widetilde{\mu}_r = \mathbb{E}((X_D - \mu)^r) = \sum_{i=1}^{N} (x_i - \mu_1)^r \cdot P(X_D = x_i). \tag{5.4}$$

From these, we can derive the variance of a discrete random variable:

$$\widetilde{\mu}_2 = Var(X_D) = \sigma^2 = \sum_{i=1}^{N} (x_i - \mu_1)^2 \cdot P(X_D = x_i). \tag{5.5}$$

Similar to the discrete case, the moments for a continuous random variable X_C are given by:

$$\mu_r = \mathbb{E}(X_D^r) = \int_{-\infty}^{\infty} x^r f(x) dx. \tag{5.6}$$

The mean and variance are calculated similarly:

$$\mu_1 = \mathbb{E}(X_D) = \int_{-\infty}^{\infty} x f(x) dx, \tag{5.7}$$

$$Var(X_C) = \sigma_{X_C}^2 = \int_{-\infty}^{\infty} (x - \mu_1)^2 f(x) dx. \tag{5.8}$$

Then, to estimate the parameter vector θ, we must express the first p moments as:

$$\begin{aligned} \mu_1 &= \mathbb{E}(X) = g_1(\theta_1, \theta_2, ..., \theta_p), \\ \mu_2 &= \mathbb{E}(X^2) = g_2(\theta_1, \theta_2, ..., \theta_p), \\ &\vdots \\ \mu_p &= \mathbb{E}(X^p) = g_p(\theta_1, \theta_2, ..., \theta_p), \end{aligned} \tag{5.9}$$

where $g_1, g_2, ..., g_p$ are functions that can be inverted into the functions $h_1, h_2, ..., h_p$:

$$\begin{aligned} \theta_1 &= h_1(\mu_1, \mu_2, ..., \mu_p), \\ \theta_2 &= h_2(\mu_1, \mu_2, ..., \mu_p), \\ &\vdots \\ \theta_p &= h_p(\mu_1, \mu_2, ..., \mu_p). \end{aligned} \tag{5.10}$$

Using the law of large numbers, we can estimate the moments as:

$$\begin{aligned}\hat{\mu}_1 &= \frac{1}{N}\sum_{i=1}^{N} X_i, \\ \hat{\mu}_2 &= \frac{1}{N}\sum_{i=1}^{N} X_i^2, \\ &\vdots \\ \hat{\mu}_p &= \frac{1}{N}\sum_{i=1}^{N} X_i^p.\end{aligned} \tag{5.11}$$

We finally estimate the parameters substituting the sample averages $\hat{\mu}_j$ in the h_j functions:

$$\begin{aligned}\hat{\theta}_1 &= h_1(\hat{\mu}_1, \hat{\mu}_2, ..., \hat{\mu}_p), \\ \hat{\theta}_2 &= h_2(\hat{\mu}_1, \hat{\mu}_2, ..., \hat{\mu}_p), \\ &\vdots \\ \hat{\theta}_p &= h_p(\hat{\mu}_1, \hat{\mu}_2, ..., \hat{\mu}_p).\end{aligned} \tag{5.12}$$

See [88] for introductory examples and applications.

5.2.2. Method of least squares

The *method of least squares* was developed independently by Gauss, Legendre, and Adrain [77]. This method allows us to find a regression equation that relates a dependent variable with one or more independent variables. The regression can be linear or non-linear, and the optimization of the parameters is obtained by deriving the minimum of the sum of squares of the residuals (i.e., deviations from the regression).

Let us have n observations: $X_1, X_2, ..., X_n$, such that the number of parameters to estimate p is smaller than the number of observations ($p < n$). Each observation will have a random error (i.e., following a Gaussian distribution) associated as:

$$X_i = \alpha + \epsilon_i, \tag{5.13}$$

where X_i would be the i-th measurement, α the true value of a physical quantity, and ϵ_i the error in the i-th measurement. Our estimators for both α and ϵ_i are:

$$\begin{aligned}\hat{\alpha} = a &\longrightarrow \text{ estimate of } \alpha \\ \hat{\epsilon_i} = e_i &\longrightarrow \text{ estimate of } \epsilon_i\end{aligned} \tag{5.14}$$

This would be equivalent to have:

$$X_i = \alpha + \epsilon_i = a + e_i. \tag{5.15}$$

The method's goal is to estimate the regression coefficients (in this case α and ϵ_i) by minimizing the sum of the squares (or other quadratic forms). The sum of the squares is

$$\sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} (a - X_i)^2. \tag{5.16}$$

The minimum is obtained as

$$\frac{\partial(\sum_{i=1}^{n} e_i^2)}{\partial a} = 0 \Rightarrow a = \sum_{i=1}^{n} \frac{y_i}{n} = \overline{y}, \tag{5.17}$$

which the last term is equal to the arithmetic mean of the observations $\overline{y}$, returning a simple estimate of α. The corresponding estimation of the dispersion is the standard deviation from $\overline{y}$. The simplest case is when we apply a linear regression to a set of points $\boldsymbol{x}, \boldsymbol{y} = \{x_i, y_i\}_{i=1}^{n}$ such that they follow a linear relation:

$$\boldsymbol{y} = \alpha + \beta \boldsymbol{x}, \tag{5.18}$$

and the values x_i are exact, while the values y_i are subjected to some error ϵ_i. Following the same approach, the observational equations would be given as:

$$y_i = \alpha + \beta x_i + \epsilon_i = a + b x_i + e_i, \tag{5.19}$$

where a, b, and e_i are estimates of α, β and ϵ_i respectively. The optimization should be done on

$$\sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} (a + b x_i - y_i)^2. \tag{5.20}$$

We minimize both a and b by optimizing:

$$\begin{aligned} \frac{\partial(\sum_{i=1}^{n} e_i^2)}{\partial a} = 0 \Rightarrow n a + b \sum_{i=1}^{n} x_i = \sum_{i=1}^{n} y_i, \\ \frac{\partial(\sum_{i=1}^{n} e_i^2)}{\partial b} = 0 \Rightarrow a \sum_{i=1}^{n} x_i + b \sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} x_i y_i. \end{aligned} \tag{5.21}$$

The solution would be:

$$\begin{aligned} a = \frac{n \sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n \sum_{i=1}^{n} x_i^2 - (\sum_{i=1}^{n} x_i)^2}, \\ b = \frac{\sum_{i=1}^{n} x_i^2 \sum_{i=1}^{n} y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} x_i y_i}{n \sum_{i=1}^{n} x_i^2 - (\sum_{i=1}^{n} x_i)^2}. \end{aligned} \tag{5.22}$$

were a and b would be the y-intersection and the slope of the linear regression respectively.

One classic application in cosmology of this method is the linear regression of the *Hubble–Lemaître Law* done by Edwin Hubble in 1929 [6].

Application of the method of least squares: Hubble–Lemaître Law

At the end of the 1920s, both *Edwin Hubble* and *Georges Lemaître* discovered that galaxies farther away from us are moving faster than nearby galaxies [5, 6]. They derived a linear relationship between the recession velocity v and the distance from us d:

$$v = H_0 d. \tag{5.23}$$

This is the famous *Hubble–Lemaître Law*. Its importance relies on the fact that it was the first observational evidence to support an expanding universe.

Hubble proposed a value of $H_0 \sim 500$ [km s^{-1} Mpc^{-1}] [6]. As seen in Fig. 5.1, we obtained a value of $H_0 = 454.2$ [km s^{-1} Mpc^{-1}] using the method of least squares in the same data. Nowadays, the estimated value for H_0 differs by a factor of 7 from the ones found by Hubble. (See Fig. 5.11) The inconsistencies are solved with a better calibration of the distances. Despite having different coefficients, the linear approximation was correct.

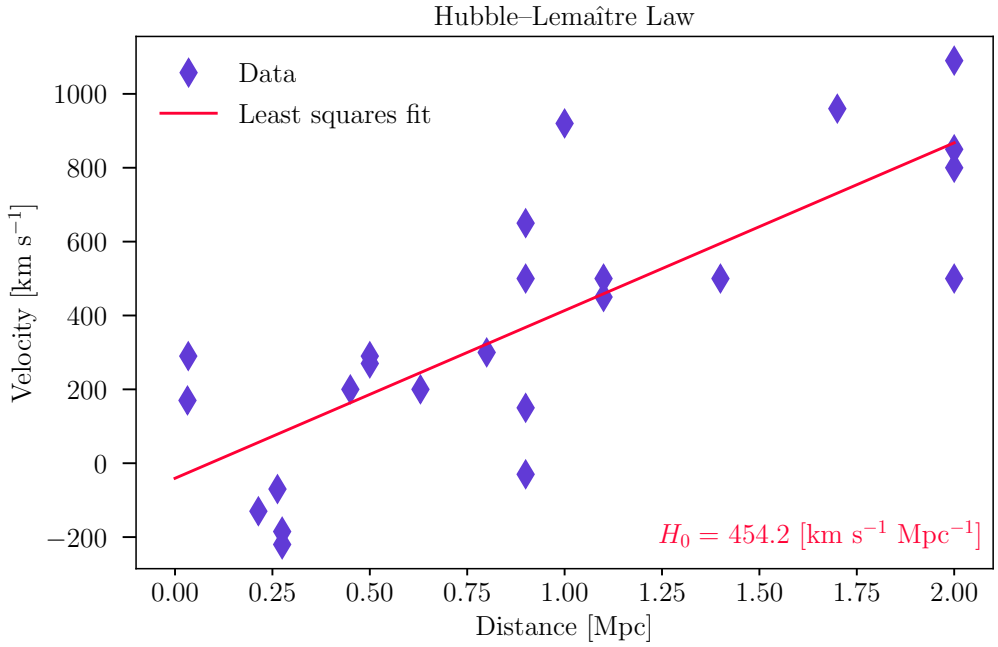


Figure 5.1: Least square approximation of Hubble's data [6].

5.2.3. Maximum likelihood method

The *maximum likelihood method* is almost immediately associated with Fisher.[26] His main contribution relies on the formality, definition, and study of properties of this method [91]. In Eq. 4.26, we introduced the concept of likelihood. Its first definition appeared a century ago in Fisher's work.[27]

[26] Nevertheless, the first records of statisticians using this method take us to 1760 with Lambert's *Photometria* [89]. Other mathematicians, throughout the years, used the methodology, such as Lagrange between 1770-1773 [90].

[27] In 1912, Fisher published the original statements that founded the maximum likelihood method [92], without defining (for the first time) the likelihood function until 1921 [93]. In the following year, he published [87] and proved the properties in a way we would consider not rigorous enough but still fundamental to the statistics of the 20th century.

Fisher's critics of the methods of moments and the method of least squares relied on the arbitrary choice of moment equations and the lack of invariance under a scale change in the variables. This motivated him to introduce the formalism we use nowadays.

Let $\boldsymbol{X} = (X_1, X_2, ..., X_n)$ be a random sample, $\boldsymbol{\theta} = (\theta_1, \theta_2, ..., \theta_p)$ a parameter vector and let $\mathcal{L}(\boldsymbol{\theta})$ be the likelihood function originally obtained through the joint probability distribution for the random variables. In the case the samples are independent and identically distributed, the likelihood is equivalent to the product of the individual probability density function:

$$\mathcal{L}(\boldsymbol{\theta}) = \prod_{i=1}^{n} f_{\boldsymbol{\theta}}(X_i). \tag{5.24}$$

We would say that any $\boldsymbol{\theta}_{MLE}$ such that maximizes $\mathcal{L}(\boldsymbol{\theta})$ is called the *maximum likelihood estimator* of the unknown true parameters $\boldsymbol{\theta}$ [77].

$$\boldsymbol{\theta}_{MLE} = \arg\max_{\boldsymbol{\theta}} \left\{ \prod_{i=1}^{n} f_{\boldsymbol{\theta}}(X_i) \right\}. \tag{5.25}$$

Usually, the $\log(\mathcal{L}(\boldsymbol{\theta}))$ function is maximized due to the invariant property of the estimator.

In the case of a pure n-dimensional multivariate Gaussian distribution in the parameters $\mathcal{L}(\boldsymbol{\theta}) \propto \exp(-\chi^2_n/2)$. Then, the maximum likelihood $\mathcal{L}_{\max}$ would be equivalent to the minimum Chi-squared $\chi^2_{\min}$. (Similar to Eq. 4.30.)

If we apply the maximum likelihood method to the data from the Hubble-Lemaître law from Section 5.2.2, we will obtain the same estimators as the ones found in Eq. 5.22. This is a special case when we assume that the observations follow a Gaussian distribution with unknown mean and variance:

$$f_{\boldsymbol{\theta}}(X_i) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(a+bx_i-y_i)^2}. \tag{5.26}$$

The optimization would be:

$$\log\left\{ \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(a+bx_i-y_i)^2} \right\} \propto \sum_{i=1}^{n} (a + bx_i - y_i)^2. \tag{5.27}$$

The RHS value is the same result as the one observed in Eq. 5.20. If the errors from the observations are non-Gaussian, this proportionality is no longer true.

5.2.4. Maximum a posteriori

The method of *maximum a posteriori* aims to estimate *the most likely* $\boldsymbol{\theta}$ *giving the data* and our prior knowledge [94]. It is a common mistake to assume that this estimate is the same as the maximum likelihood encountered value (see Section 5.2.3). The last one determines the value of parameters that *makes the data most likely*.

Mathematically, we would say that any $\boldsymbol{\theta}_{MAP}$ such that maximizes the posterior $P(\boldsymbol{\theta}|\mathcal{D})$ is called the maximum a posteriori estimator of the unknown true parameters $\boldsymbol{\theta}$:

$$\boldsymbol{\theta}_{MAP} = \arg\max_{\boldsymbol{\theta}} \Big\{ P(\boldsymbol{\theta}|\mathcal{D}) \Big\}. \tag{5.28}$$

If we use the Bayes' theorem (Eq. 4.26) and in the case the samples are independent and identically distributed, Eq. 5.28 would be equivalent to:

$$\boldsymbol{\theta}_{MAP} = \arg\max_{\boldsymbol{\theta}} \left\{ \frac{\prod_{i=1}^{n} f_{\boldsymbol{\theta}}(X_i) P(\boldsymbol{\theta})}{P(\mathcal{D})} \right\}. \tag{5.29}$$

The logarithm of the above expression is:

$$\boldsymbol{\theta}_{MAP} = \arg\max_{\boldsymbol{\theta}} \Big\{ \sum_{i=1}^{n} \log(f_{\boldsymbol{\theta}}(X_i)) + \log(P(\boldsymbol{\theta})) - \log(P(\mathcal{D})) \Big\}. \tag{5.30}$$

When we optimize Eq. 5.30, the Evidence $P(\mathcal{D})$ would not be considered because it does not depend on $\boldsymbol{\theta}$:

$$\boldsymbol{\theta}_{MAP} = \arg\max_{\boldsymbol{\theta}} \Big\{ \sum_{i=1}^{n} \log(f_{\theta}(X_i)) + \log(P(\boldsymbol{\theta})) \Big\}. \tag{5.31}$$

If the prior is non-informative (i.e. $P(\boldsymbol{\theta}) = 1$), then $\boldsymbol{\theta}_{MAP} = \boldsymbol{\theta}_{MLE}$.

5.2.5. Other point estimators

We invite the reader to study other point estimator techniques, such as

- **Uniformly minimum-variance unbiased estimator**: see chapter 2 from [95] for a large review of this estimators based on the unbiasedness property.
- **Minimum mean squared error estimator**: also known as the posterior mean:

  $$\widehat{\theta}(x) = E[\theta|x] = \int \theta \, p(\theta|x) \, d\theta.$$

  See [96] for more comments on this.
- **Median unbiased estimator**: see [97] for an introduction of the best median-unbiased estimators in simple standard problems.

5.3. Interval estimators

An *interval estimator* bounds the possible values of a parameter in a *range* with some degree of sense. Estimating an interval allows us to draw inferences about a population of interest. They aim to comprise the true underlying parameter.

The interval estimators contrast to the already studied point estimators from Section 5.2 because they assign an interval of possible values rather than a specific one. As in the point estimates, interval estimation can be complex, and they are not uniquely defined. Their definition is even more entangled when considering many dimensions.

In many areas of science, and in particular, in cosmology, the most used interval estimators are the *confidence intervals*. However, as we will see in the following subsections, it is a common mistake to understand them from a Bayesian perspective, while they came from the frequentist theory counterpart.

5.3.1. Confidence Intervals

The *confidence intervals* was first presented by Neyman in 1937 [98], and, currently, they are the most used interval estimators. These are the resulting intervals estimated using frequentist reasoning and summarize the concept of *coverage* of the true parameter in independent realizations of an experiment. In other words, the confidence intervals ensures an interval in which, after independent realization of the experiment, some fraction will fall inside of this interval with some probability.

The interval is determined at a certain *confidence level*, which represents the long-run frequency of confidence intervals that contain the true value of the parameter.[28] Common confidence levels are $68.3\,\%, 95.4\,\%$ and $99.7\,\%$ (these are the $1-\sigma$, $2-\sigma$ and $3-\sigma$ ranges away from a common Gaussian distribution) and it might not necessarily be always *symmetric* (or centered with respect to a point estimate) [77]. The extension of this interval in many dimensions is the *confidence region*.

Let X be a sample from some probability distribution. Let θ be the parameter of interest. A confidence interval for the θ, with confidence level $1-\alpha$, is an interval with random endpoints $[\theta_{min}(X), \theta_{max}(X)]$ such that in

$$P(\theta_{min}(X) < \theta < \theta_{max}(X)) = 1 - \alpha. \tag{5.32}$$

We will present a simple illustration of the idea of confidence intervals in this work: let θ be the parameter of interest and $[\theta_{min}, \theta_{max}]$ the confidence interval for its estimate at a $1-\alpha = 75\,\%$ confidence level. If we sample 20 times independently from the original population, we would expect that the truth parameter θ^* lies within the confidence interval a total of 15 times out of 20 approximately. Fig. 5.2 represents the example.

The probability distribution that describes the coverage will correspond to a binomial distribution if the experiment is repeated a fixed number of times.[29] We will call this a *binomial proportion confidence interval*.

The confidence level and the sample size will affect the width of the estimated confidence interval. Still, it would be often desired to have a *valid* estimates (the coverage probability should be achieved approximately) [100].

In many dimensions, the concept of the 1-σ, 2-σ, and 3-σ makes no sense anymore because

[28] See [99] and Lecture 17 from [88] for a classical approach to derive confidence intervals and page 13 from [81] to derive them from the expansion of the log-likelihood distribution. These last approximations of the distributions assume a Gaussian-shaped probability distribution, and in consequence, the well-known *empirical rule* of the confidence intervals of the mean with the standard deviations. (See pages 505-508 from [99].)

[29] Each trial of the experiment has two possible outcomes (inside or not of the interval), the probability of success is the same for each trial, and the trials are independent.

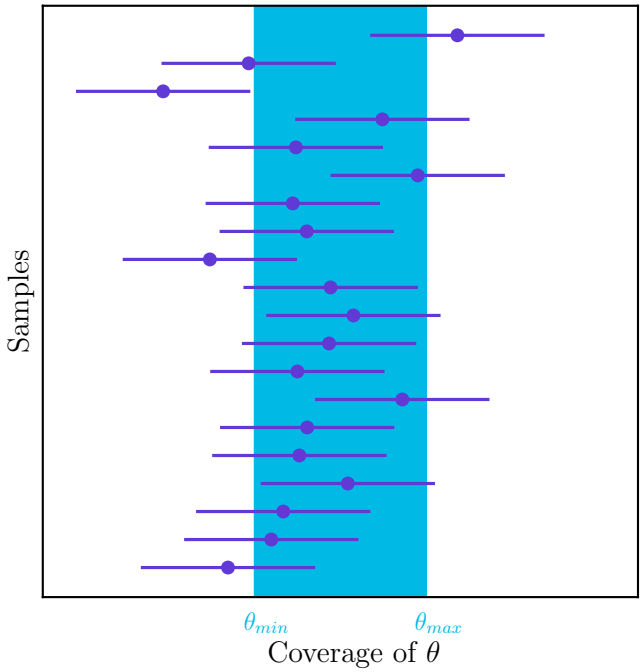


Figure 5.2: Example coverage of 20 samples of θ. A total of 15 relied inside the confidence interval $[\theta_{min}, \theta_{max}]$, which is expected for a confidence level of $1 - \alpha = 75\,\%$.

there is no addition of standard deviations, but the corresponding probabilities are still used to construct the confidence levels. An usual approach is to first calculate the maximum likelihood estimator $\mathcal{L}_{\max}$ from Sec. 5.2.3 and then select the bounds x_1 and x_2 such that:

$$\int_{x_1}^{x_2} \mathcal{L}(x)dx = 0.683 \int_{\infty}^{\infty} \mathcal{L}(x)dx. \tag{5.33}$$

The same can be done with $95.4\,\%$ and $99.7\,\%$ in the RHS part of Eq. 5.33. Some other attributes can be demanded, such as symmetry with respect to the $\mathcal{L}_{\max}$.

It is fundamental to reiterate the meaning of the confidence intervals: *if an experiment is repeated, the truth parameter will lie a certain fraction of times inside of it.* In cosmology, we only have *one realization of the observable universe*. This state opens philosophical discussions and a fundamental point when analyzing the confidence intervals of cosmological parameters. Many articles express that the true cosmology is bounded inside the confidence interval limits, but this is a misconception of its frequentist nature. There is a whole family of interval estimators that accounts for this Bayesian perspective that we will review hereafter.

5.3.2. Credible Intervals

A *credible interval* describes the probability of a parameter being contained in an interval. It is calculated in the one-dimensional posterior probability distribution, and it is the com-

monly estimated interval of the Bayesian statistics. It can be determined considering a point estimate. The extension of this interval in many dimensions is the *credible region*.

In simple words, the subjective probability that θ lies between θ_{min} and θ_{max} is $1-(\alpha_1+\alpha_2)$, being $\theta_{min} \leq \theta \leq \theta_{max}$ the corresponding $1-(\alpha_1+\alpha_2)$ credible interval. See Fig. 5.3 for an illustration.

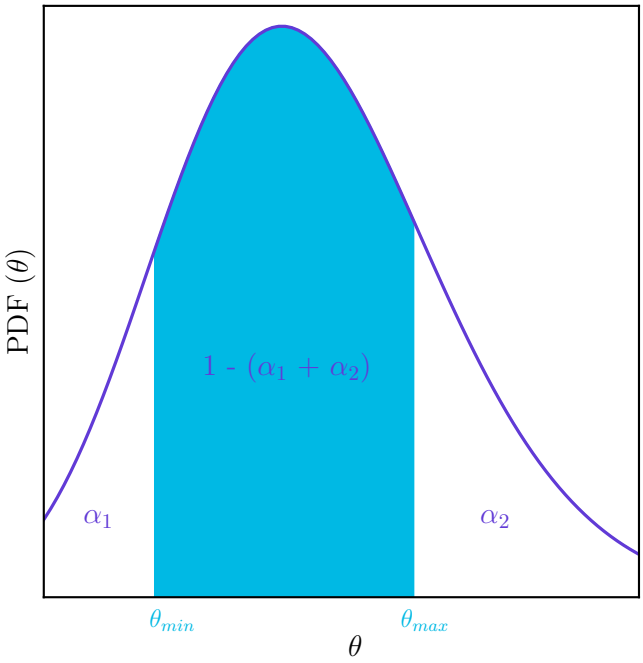


Figure 5.3: Illustration of the credible interval $[\theta_{min}, \theta_{max}]$ in which the true parameter θ lies inside with a is $1-(\alpha_1+\alpha_2)$ probability.

In the case $\alpha_1 = \alpha_2$, we would call it the *centered credible interval*, while $\alpha_1 \neq \alpha_2$ would be called the *non-centered credible interval*.

The spaces that can account for this are more complex to define in many dimensions. It is common to use the isodensity lines in two dimensions, which are lines that join points with the same probability density value. The hyper-volume encapsulated accounts for the credible region. In Fig. 5.4 we show the projection of the isodensity lines for two bivariate Gaussians. It is important to recall that the credible intervals *are not unique*. It is possible to center the intervals on a posterior distribution with respect to point estimators. That includes the maximum a posteriori (*highest posterior density interval*), the median (*equal-tailed interval*) and the mean.

5.3.3. Other interval estimators

We invite the reader to study other estimator techniques, such as

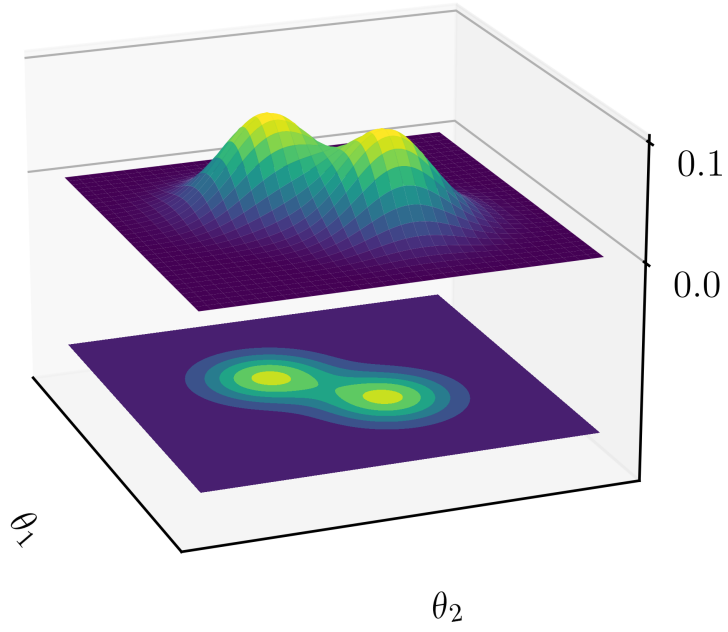


Figure 5.4: Two bivariate Gaussian probability density distributions and the corresponding isodensity curves. Each isodensity line represents the combination of the parameters θ_1 and θ_2 with the same probability density. These are used as an extension of the credible intervals to visualize credible regions for the true parameters aimed to estimate.

- **Bootstrap intervals**: see pages 51-54 from [99] for a friendly introduction.
- **Feldman-Cousins intervals**: this interval is determined through the likelihood ratio comparison across potential hypotheses:

$$\Lambda_\theta(x) = \frac{\mathcal{L}(x|\theta)}{\mathcal{L}(x|\theta_{best})}, \tag{5.34}$$

  where x is the measured value, θ is the true value, and θ_{best} is the maximum likelihood given the data. See Feldman and Cousins' original paper [101] for more details.

5.4. Discussion on selecting estimators

There are open questions in statistical inference when analyzing cosmological data. Some of them are the following:

- *Which point and interval estimators should we use when doing a cosmological analysis?*
- *How can we relate point and interval estimators in many dimensions and in their marginalized spaces?*

A common approach is a mid-point between the minimum variance and the highest probability of a point estimate. However, this is not evident when the underlying distribution is

unknown, which is the typical case in cosmology.

One example of this tension is comparing the use of the MAP in many dimensions versus the marginalized 1D mean. In Figure 7.1, we illustrate how much the point estimators may vary if the shape of the distribution is highly non-Gaussian. On the other hand, if the distribution follows a multivariate Gaussian distribution, then the MAP would be equivalent to the projected 1D mean:

$$\boldsymbol{\theta_{MAP}} = [\theta_{MAP,1}, \theta_{MAP,2}] = [\overline{\theta_1}, \overline{\theta_2}] = \overline{\boldsymbol{\theta}}. \tag{5.35}$$

In DES Y3 results [102, 103], the point estimates are presented as:

$$\text{Parameter} = \text{1D mean}^{+\text{upper 34\% bound}}_{-\text{ lower 34\% bound}}(\text{MAP value})$$

Another case where determining confidence intervals is not clear is when the number of realizations of an experiment is small. This leads to conflicts, mainly because of divergent results on the estimates. (E.g. some particle physics experiments were identified to have this problem; see Biller & Oster (2015) [104]).

An interdisciplinary study of how estimators vary in different experiments can help solve the current statistical challenges of observational cosmology due to the enormous variety of scale-dependent physics present in the universe.

5.5. Sampling methods

Sampling methods are widely used to infer properties about a distribution and can be used to solve integration and optimization problems in large dimensional spaces. They also allow us to simulate physical systems. There is not a unique sampling method, and all of them have limitations. Some of them are random walks (such as the *Gibbs sampling* [105] and the *Metropolis Hastings* [106, 107]), *rejection sampling* [108], *sequential Monte Carlo* (like the *importance sampling* [109]), and many others. We will focus on two: Markov chains Monte Carlo and Sequential Monte Carlo.

5.5.1. Markov chains Monte Carlo

To obtain a general sample of points that represents a probability distribution, the *Markov Chain Monte Carlo* (MCMC) methods are a family of practical algorithms for that purpose [110]. They mainly construct a *Markov chain* that will have an equilibrium distribution equivalent to the sampled one, as the number of steps on it increase.

The intuition behind this family of sampling methods is to simulate a random thermal fluctuation of a system from state to state throughout an experiment [111]. The chain will assign a weight to each state (or step) that will be equal to the probability of having that particular state. The complete information would allow us to obtain estimators of the underlying probability distribution and all the possible states of the system. The MCMC is obtained by setting a transition dynamic from step to step following the *Markov property*: the future state probability would only depend on the present state (with some uncertainty), but not on the past states [110].

There are many ways to construct a Markov chain according to the Markov property, but its quality (and therefore convergence) will depend on other factors such as the number of steps, the starting points, and anothers.

Metropolis-Hastings algorithm:

Introduced by Metropolis (1953) [107] and Hastings (1970) [106], this algorithm returns a Markov chain in which each state $\boldsymbol{x}^{t+1}$ depends only on the previous state $\boldsymbol{x}^t$ using a proposal density distribution $Q(\boldsymbol{x}^{t+1}|\boldsymbol{x}^t)$. The step will be selected with some random factor. The algorithm is described as:

1. Initialize with an starting point $\boldsymbol{x}^0$.
2. Choose a proposal function $Q(\boldsymbol{x}^{t+1}|\boldsymbol{x}^t)$ that determines how to transit from one state to another[30].
3. For $i = 0, ..., N-1$:
   - Make a proposal step $\boldsymbol{x}^p$.
   - Calculate the ratio of the likelihood distribution evaluated at the proposal step $\boldsymbol{x}^p$ and the current state $\boldsymbol{x}^t$:
   $$r = \frac{\mathcal{L}(\boldsymbol{x}^p)}{\mathcal{L}(\boldsymbol{x}^t)}.$$
   - If $r \geq 1$: $\boldsymbol{x}^{t+1} = \boldsymbol{x}^p$
   - If $r < 1$:
     - Sample a random number $u \sim U[0, 1]$.
     - If $u \leq r$: $\boldsymbol{x}^{t+1} = \boldsymbol{x}^p$
     - If $u > r$: $\boldsymbol{x}^{t+1} = \boldsymbol{x}^t$

Finally, the number of times the chain remained in the same state (*weight*) is proportional to the probability distribution that wants to be sampled.

A common approach is to use an MCMC sampler to estimate cosmological parameters from observations of *Supernovae Type Ia* (SNIa). Since the end of the twentieth century, these have been used by cosmologists to estimate the acceleration of the universe's expansion due to being standard candles. (See Chapter 2 for a theoretical explanation of the underlying physics.)

From the observed magnitude, redshift, and covariance of a total of 40 supernovae type Ia[31] obtained from the Pantheon catalogue[32], we will consider a simple version of the ΛCDM model, fixing Ω_r, and studying three parameters: Ω_m, Ω_Λ and $\mathcal{M}$. From these, the first two are

[30] A common approach is to use a multivariate Gaussian with a covariance matrix comparable to the underlying errors of the parameters.

[31] Data originally provided by Prof. Dragan Huterer at the ICTP-Trieste/ICTP-SAIFR School on Cosmology 2021.

[32] https://archive.stsci.edu/prepds/ps1cosmo/scolnic_datatable.html

cosmological (the matter and dark energy density parameters, respectively) and one nuisance (the absolute magnitude).

The apparent magnitude m from Eq. 2.50 would be the theoretical model μ:

$$m = 5\log_{10}(H_0 d_L) + \mathcal{M} = \mu(\Omega_m, \Omega_L, \mathcal{M}). \tag{5.36}$$

We can then have a final likelihood distribution to infer the posterior distributions around our parameters of interest $\boldsymbol{p} = (\Omega_m, \Omega_L, \mathcal{M})$:

$$\mathcal{L}(\boldsymbol{x}) \propto \exp\Big\{-\frac{1}{2}(\boldsymbol{x} - \mu(\boldsymbol{p}))^T C^{-1}(\boldsymbol{x} - \mu(\boldsymbol{p}))\Big\}. \tag{5.37}$$

One fundamental ingredient from the Metropolis-Hastings algorithm is the use of a proposal distribution $Q(\boldsymbol{x}^{t+1}|\boldsymbol{x}^t)$. We will show the resulting chains when considering three different Gaussian-distributed proposal distributions, which will return a random jump from a current position in parameter space:

$$Q(\boldsymbol{x}^{t+1}|\boldsymbol{x}^t) = \boldsymbol{x}^t + \mathcal{N}(\boldsymbol{0}, \boldsymbol{\Sigma}_i), \tag{5.38}$$

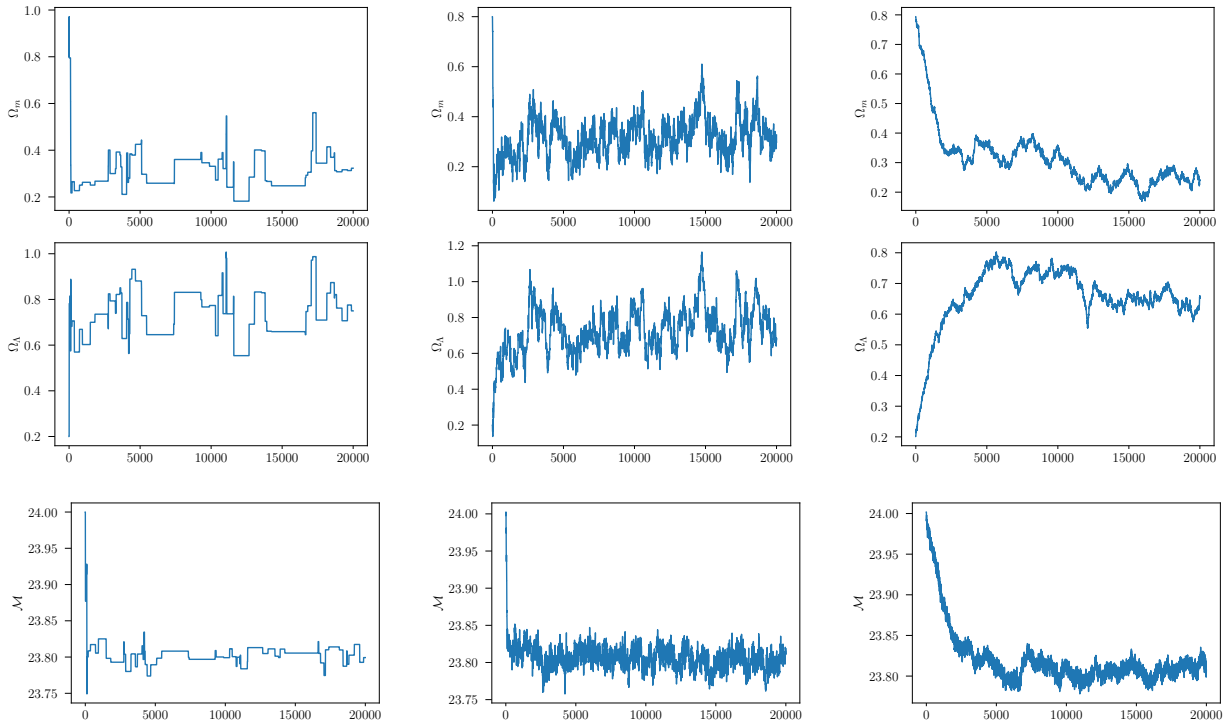
where $\boldsymbol{\Sigma}_i$ would be a vector with the different standard deviations: $\boldsymbol{\Sigma}_i = (\sigma_{\Omega_m,i}, \sigma_{\Omega_\Lambda,i}, \sigma_{\mathcal{M},i})$. In Table 5.1 we present the different proposal distributions and in Figs. 5.5, 5.6 and 5.7 we show three sampled chains with a total of 20,000 points each.

Table 5.1: Different cases for the steps of the proposal distributions, including the acceptance rate of the tested MCMC for the inference of the parameters Ω_m, Ω_L and $\mathcal{M}$ in the cosmological analysis of a sample of SNIa.

	Case 1	**Case 2**	**Case 3**
	Large step	**Efficient step**	**Short step**
σ_{Ω_m}	0.1	0.01	0.002
σ_{Ω_Λ}	0.1	0.01	0.002
$\sigma_{\mathcal{M}}$	0.1	0.025	0.005
Acceptance rate	0.01	0.16	0.65

It is clear that cases 1 and 3 are not efficient chains. This is not only showed in the corresponding acceptance rates, but in the performance of Figs. 5.6 and 5.7. The case in which the step is large (case 1), the sampler rapidly reaches the convergence, but due to the far proposed points, it usually reject most of the proposal points when it reaches high probability zones. (See left panels from 5.5.)

The other extreme, the case in which the step is too short (case 3), we observe the opposite phenomena. The close proposal points can make the sampler to get stuck in local minima, taking a lot of time to reach the zone of convergence. (See right panels from 5.5.) The case in which the step is considered efficient, shows an harmonic concentration towards the zone of convergence. (See central panels from 5.5.)



(a) Case 1: Large step (b) Case 2: Efficient step (c) Case 3: Short step

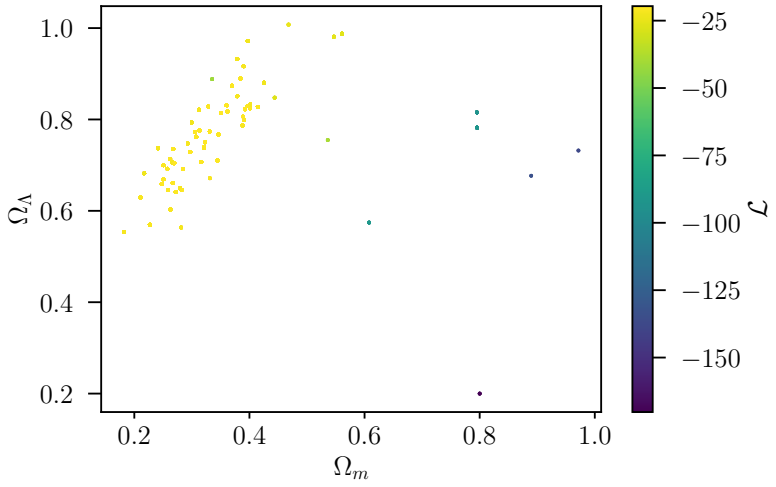
Figure 5.5: Resulting MCMC chains projected in each of the dimensions for different proposal distributions (see Table 5.1) for the cosmological example of SNIa. In the left panels, the proposal distribution reaches the actual values in a few iterations, but it does not maintain the required stability. In the center, the chain that is efficiently sampled is shown. At the right panels, the steps are small, and therefore, the chain takes longer to reach the best values.

The advantage this method offers is its mathematically simplicity: it is only a few lines of coding. Nevertheless, there are a few aspects that must be taken into consideration:

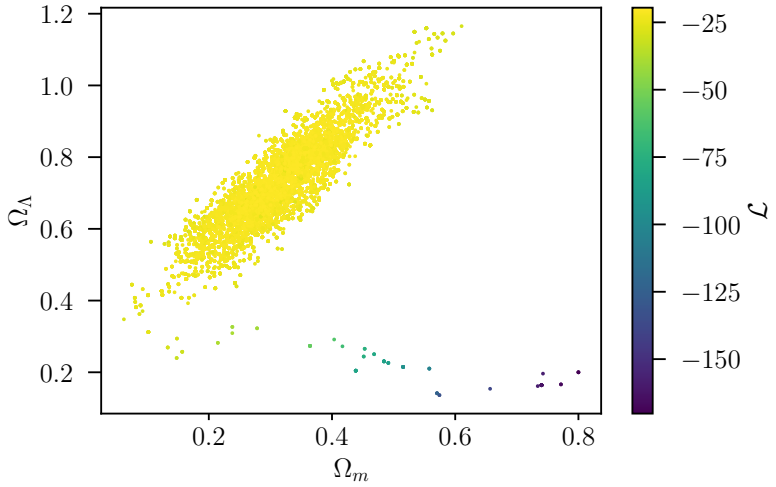
- Initially, it is necessary to drop some fraction of the chain. If the distance between the starting point and the zone where the sampler converges is too large, keeping it will bias the results. This is known as *burn-in* stage.

- As shown in 5.1, the size of the step of the proposal distribution, and therefore the acceptance rate, will determine how efficient our chain is. Usually, an acceptance rate should be between 0.1 and 0.3.

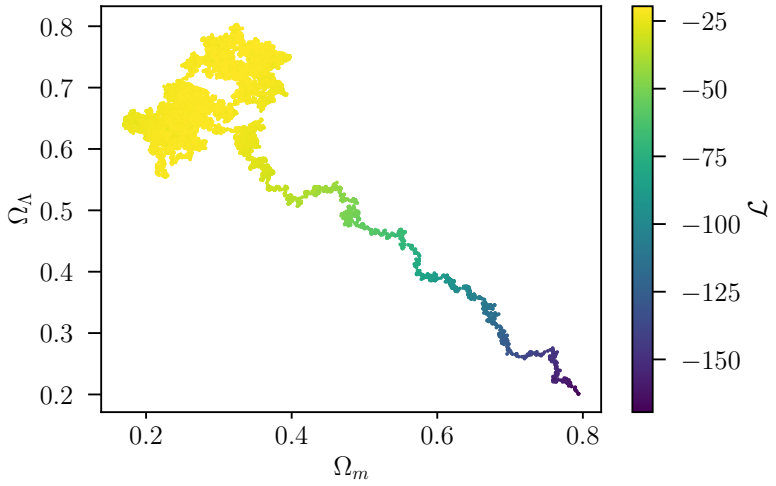
For more comments on the limitations of the MCMC method, see [112].



(a) Case 1: Large step

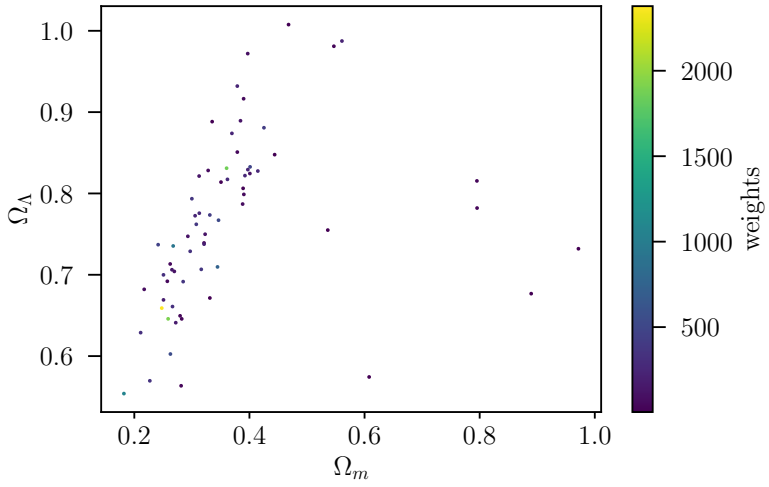


(b) Case 2: Efficient step

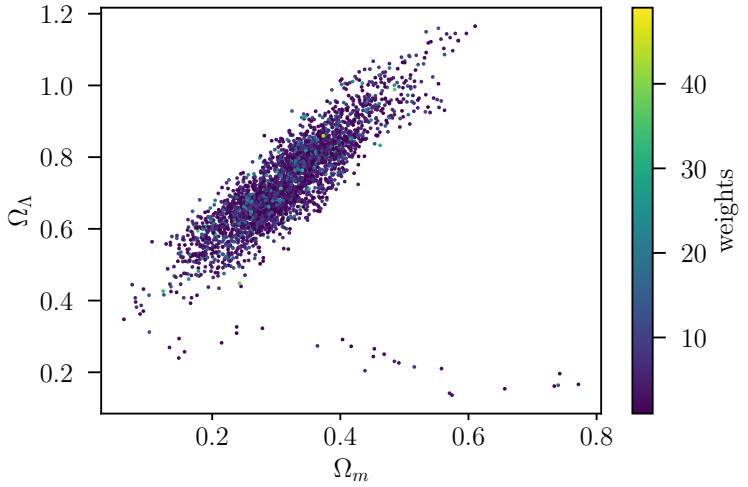


(c) Case 3: Short step

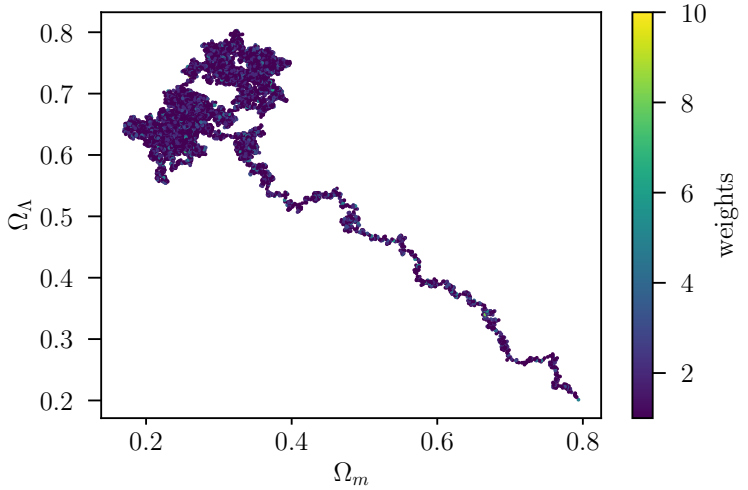
Figure 5.6: 2D Marginalized likelihood result chains for Ω_m and Ω_L.



(a) Case 1: Large step



(b) Case 2: Efficient step



(c) Case 3: Short step

Figure 5.7: 2D weighted result chains for Ω_m and Ω_L.

5.5.2. Sequential Monte Carlo

The *sequential Monte Carlo* is a method that approximates a probability distribution $p(\boldsymbol{x})$ using samples from a proposal distribution $\pi(\boldsymbol{x})$. The goal is to estimate the posterior density of some state variables given some observation variables [113].

Usually, sequential Monte Carlo problems consist of both hidden and observable variables. Let $\boldsymbol{x}$ represent some of those hidden states and $\boldsymbol{y}$ represent an observation. The observations are related to the hidden states by some known functional form. The probability of observing $\boldsymbol{y}$ given $\boldsymbol{x}$ would be given by the conditional probability $p(\boldsymbol{y}|\boldsymbol{x})$, also known as *filtering distribution*, and is the one we will sample. Similarly, the dynamical system describing the evolution of the state variables is also known probabilistically. The expectation of an observable $\mu(\boldsymbol{x})$ would be given by:

$$\mathbb{E}[\mu(\boldsymbol{x})|\boldsymbol{y}] = \int \mu(\boldsymbol{x}) p(\boldsymbol{x}|\boldsymbol{y}) d\boldsymbol{x} = \frac{\int \mu(\boldsymbol{x}) p(\boldsymbol{x}|\boldsymbol{y}) d\boldsymbol{x}}{\int p(\boldsymbol{x}|\boldsymbol{y}) d\boldsymbol{x}}. \tag{5.39}$$

The goal will be to sequentially estimate some hidden states $\boldsymbol{x}^1, ..., \boldsymbol{x}^N$, given the values of the observation process $\boldsymbol{y}^1, ..., \boldsymbol{y}^N$, at any time step.

This methodology provides an approximation of the conditional probability $p(\boldsymbol{x}^i|\boldsymbol{y}^i)$ by calculating weights obtained with a type particle algorithm. Importance sampling can be used to model the full posterior $p(\boldsymbol{x}^1, ..., \boldsymbol{x}^N|\boldsymbol{y}^1, ..., \boldsymbol{y}^N)$. To do this, we would use an alternative and simpler function $\pi(\boldsymbol{y}|\boldsymbol{x})$.

Importance sampling:

Instead of using a grid over the space of parameters to observe $\boldsymbol{y}$, we can draw weights $w(\boldsymbol{x}) = \dfrac{p(\boldsymbol{x}|\boldsymbol{y})}{\pi(\boldsymbol{x}|\boldsymbol{y})}$ (called *importance weights*) that represent the probabilities using an *importance sampling* [114] technique.

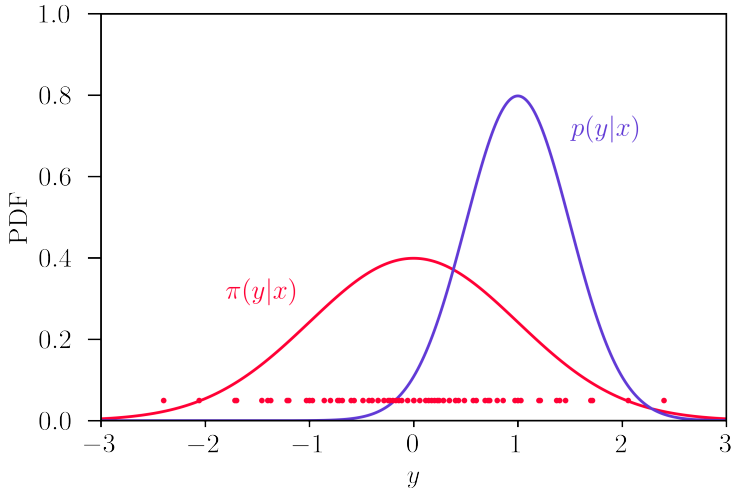
The expectation of $\mu(\boldsymbol{x})$ would be as follows:

$$\mathbb{E}[\mu(\boldsymbol{x})|\boldsymbol{y}] = \frac{\int \mu(\boldsymbol{x}) w(\boldsymbol{x}) \pi(\boldsymbol{x}|\boldsymbol{y}) d\boldsymbol{x}}{\int w(\boldsymbol{x}) \pi(\boldsymbol{x}|\boldsymbol{y}) d\boldsymbol{x}}. \tag{5.40}$$

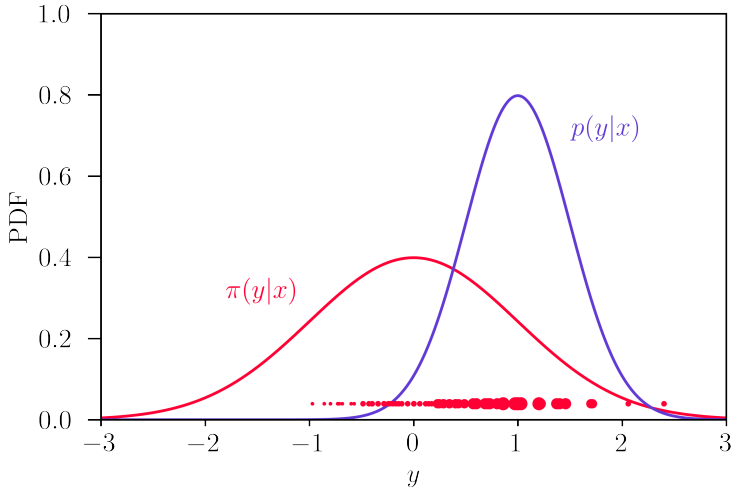
Then, to approximate $\mathbb{E}[\mu(\boldsymbol{x})|\boldsymbol{y}]$, we just need to draw samples $\{\boldsymbol{x}^i\}_{i=1}^N$ from $\pi(\boldsymbol{x}|\boldsymbol{y})$ and calculate the discrete representation of the underlying probability:

$$\mathbb{E}[\mu(\boldsymbol{x})|\boldsymbol{y}] = \frac{1/N \sum_{i=1}^N \mu(\boldsymbol{x}^i) w(\boldsymbol{x}^i)}{1/N \sum_{i=1}^N w(\boldsymbol{x}^i)} = \sum_{i=1}^N \mu(\boldsymbol{x}^i) \overline{w}(\boldsymbol{x}^i), \tag{5.41}$$

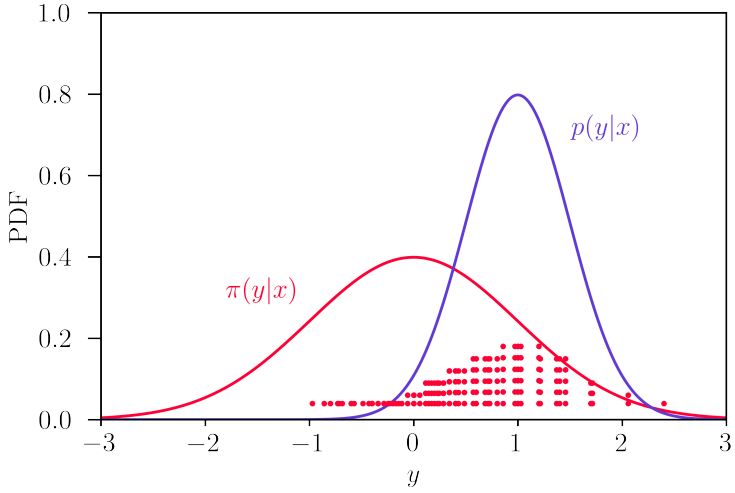
where $\overline{w}$ are the normalized importance weights. Those weights can be transformed, as done in Sect. 5.5.1 to obtain the final posterior. (See Fig. 5.8 for a sketch that summarizes the idea behind this method).



(a) Sampling



(b) Weighting



(c) Re-sampling

Figure 5.8: Illustration of the three stages of importance sampling: *sampling*, *weighting* and *re-sampling*.

The sampling technique can be done iteratively. As the method follows the Monte Carlo principle, the different results will return a sequence.

The main challenge of the method is to propose a distribution $\pi(\boldsymbol{x}|\boldsymbol{y})$ that *supports*[33] the filtering distribution $p(\boldsymbol{x}|\boldsymbol{y})$. (See Fig. 5.8, where the points drawn from $\pi(\boldsymbol{x}|\boldsymbol{y})$ are significant). This characteristic will ensure that the samples are being obtained from zones of high probability. With a sufficient number of points, it would be possible to make statements about $p(\boldsymbol{x}|\boldsymbol{y})$, such as calculating point and interval estimators.

A common approach is to sample a *Gaussian cloud* of points using some representative point estimator (such as the MAP or projected mean) as the multivariate mean and some covariance matrix. Another approach is to use a *Sobol sampling*.

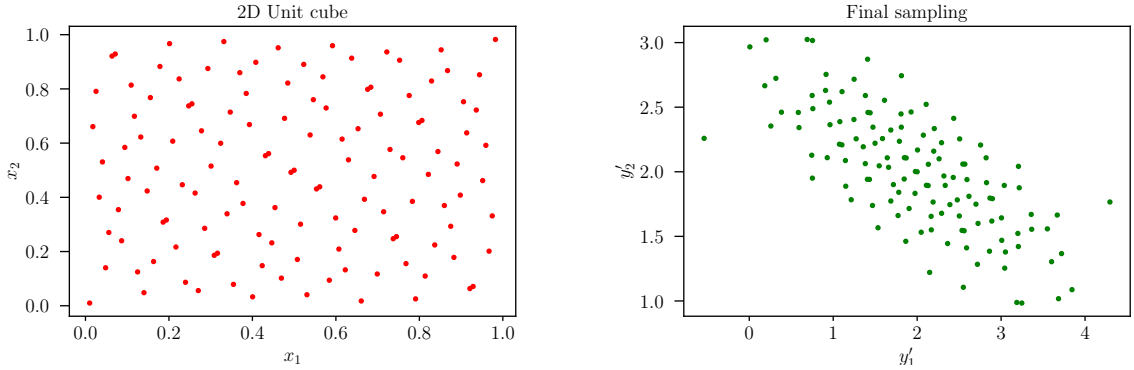
Sobol and Inverse transform sampling methods:

A Gaussian random cloud will have most points centered around the mean. An alternative method is to cover a wider area, keep the elliptical shape, and use a *Sobol sequence* to sample points.

Sobol sampling generates quasi-random numbers covering the space more evenly than a Gaussian random sample. Rather than generating random numbers, Sobol sampling generates a uniform distribution in probability space. It is a distribution which *fills in* unsampled regions. In Section A we describe the mathematical background of the sequence.

It is possible to transform the Sobol sequence into N (number of dimensions) standard normal distributions using the CDF (Cumulative Density Function) and then rotate it to the coordinate system with the parameter covariance. This procedure is called *Inverse transform sampling* [115], and it is illustrated in Fig. 5.9.

[33] Their centers are near and have a comparable width.



(a) 2D unit cube composed with a total of $2^7 = 128$ points sampled using the Sobol sampling method for the dimensions x_1 and x_2.

(b) Final sampled ellipse-shape distribution using the Sobol sampling as a base for the inverse transformation to y_1 and y_2.

Figure 5.9: Example of an inverse transform sampling using a Sobol sampled 2D unit cube as a base for a multivariate Gaussian with mean $\boldsymbol{\mu} = (2.0, 2.0)$ and covariance $\boldsymbol{\Sigma} = \begin{pmatrix} 1.0 & -0.25 \\ -0.25 & 0.5 \end{pmatrix}$.

The inverse transformation sampling takes uniform samples of a number u between 0 and 1, interpreted as a probability, and then returns the largest number from the sampled distribution such that its probability is less than u. In our case, u would be a number from one of the dimensions of the Sobol sequence.

Both the Sobol and inverse transform methods were combined in Chapter 7 to sample in a more efficient way the desired and described posterior.

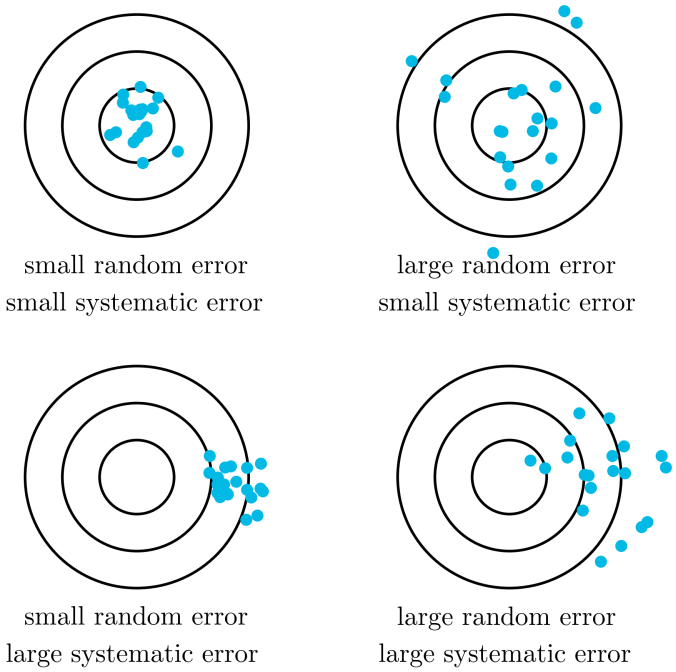
5.6. Errors and uncertainties

Errors synthesize the *uncertainties* in an experiment. If we assume that there are no measurements free of uncertainty, our goal should be to keep their magnitude to the minimum and understand their origin. This is fundamental to make reliable inferences about the universe when analyzing cosmological observations that are inherently noisy. There are two main types of errors when estimating parameters: the *random* and the *systematic errors*, which are illustrated in the top panel of Fig. 5.10.

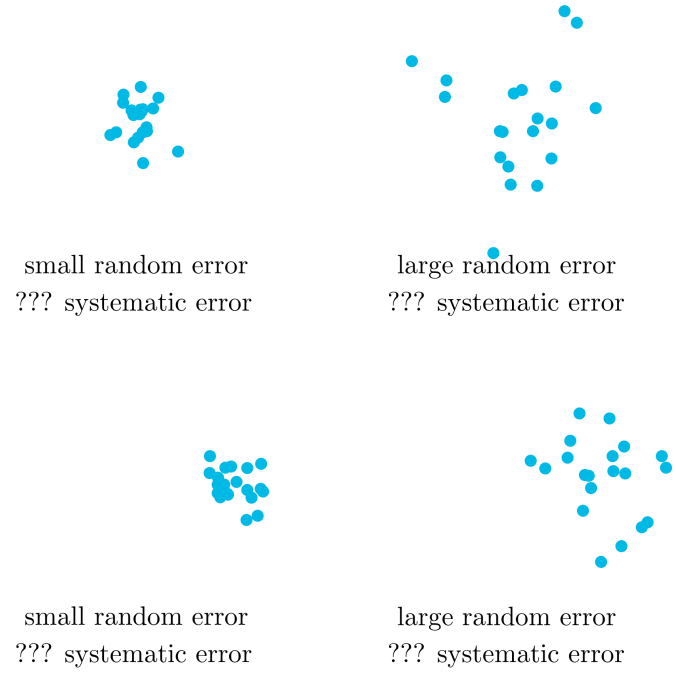
When a measurement can be repeated, the distribution's spread contributes to the uncertainty of a measured quantity. These are called *random errors*, which are the random deviations of the observations from a true value. An estimator is *consistent* if the error tends to zero as the sample become larger [116], but we must ensure that the measurements are the same each time we reproduce the experiment. One example is the Gaussian noise.

A large difference between two measurements decreases our *reliability* on the estimation. There are cases in which, even if the same quantity is measured each time, repeated samples do not always decrease the uncertainty. Those cases are called *systematic errors*. These are given by inaccuracies due to the nature of the compilation of data and cannot be removed by increasing the sample size or choosing a different type of estimator [116]. Therefore, the relative magnitude of the error does not decrease as the number of observations increases.

In reality, it is not possible to know if models correctly describe phenomena, so systematic errors are hard to evaluate and detect, unless their effect is evident. In the bottom panel of Fig. 5.10, we illustrate this by taking out the target practice. This problem has been observed in cosmology: In the last decade, the Hubble constant (H_0) parameter estimate has decreased its random errors significantly due to the augmentation of low and high redshift independent observables. Late universe observations suggest that the expansion rate is now larger than in the early universe. This inconsistency is called the *Hubble tension*, and it has been intensely debated. It is not clear if it is produced by a misconception of the underlying physics, or if it is produced by systematic errors in the collection and treatment of data. In Fig. 5.11, we illustrate two independent proofs of H_0 that decrease its error through time.



(a) Random and systematic errors in target practice.



(b) Random and systematic errors without target practice.

Figure 5.10: The size of random errors depends on the distance between the points. Therefore, they do not depend on the position of the circles, i.e., the model. On the other hand, systematic errors can only be determined if the points are systematically centered or not with respect to some reference. (Image adapted from [117].)

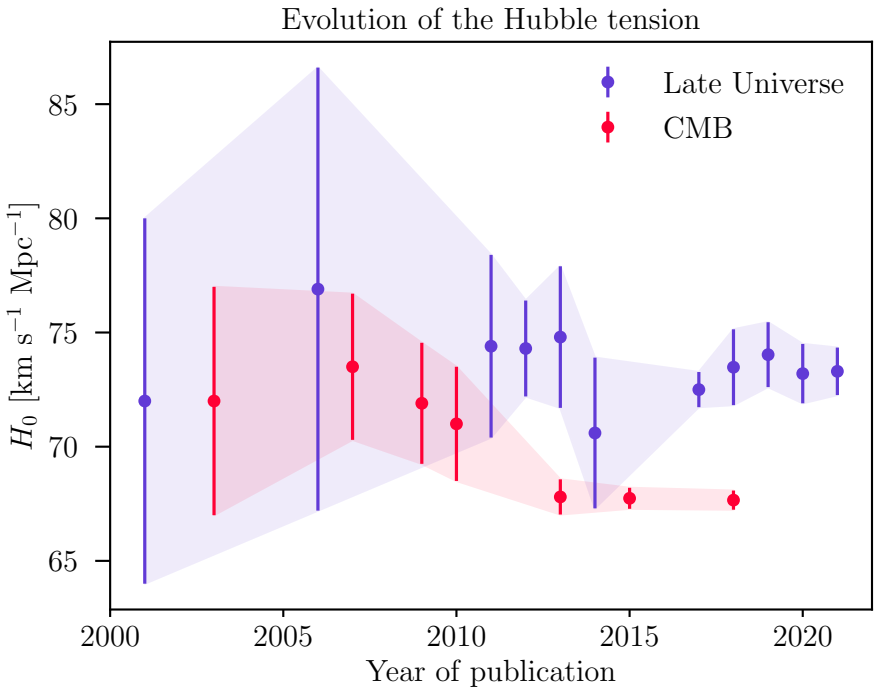


Figure 5.11: Hubble constant (H_0) estimate and uncertainties changes over time in the ΛCDM model 2.4. Observations of the early universe, corresponding to the Cosmic Microwave Background (CMB) [44, 118–123], are shown in red, while the purple measurements are given by late universe standard candles, such as Supernovae and Cepheid variable stars [124–134]. Error bars are 1σ.

Fig. 5.11 illustrates a clear *discrepancy*. During the first decade (2000-2010), the estimated intervals of the measurements appeared to converge and overlap, allowing cosmologists to believe that the ΛCDM theory of the cosmos described the observations accurately. Nevertheless, during the last decade (2010-2020), the difference between measured values for H_0 has become even more significant.

A second tension has increased the community's interest: when comparing Planck's data with weak lensing measurements and redshift surveys, the value of the matter energy density Ω_m, and the amplitude or rate of growth of structure (σ_8, $f\sigma_8$). See Fig. 5 from [135] for quantification of this tension through the parameter S_8, using different probe combinations.

In the last US Community Study on the Future of Particle Physics event (SnowMass 2021), an extensive review of the cosmological tensions and anomalies associated with Particle Physics, Astrophysics, and Cosmology was released [136]. One of the critical questions cosmologists should answer in the following decade is

- *Can these tensions, individually or together, be systematic errors in the current measurements? Are these tensions statistical flukes, or are they pointing to new physics?*
- *If not due to systematics, what is the origin of the sharpened tension in the observed and inferred values of H_0, $f\sigma_8$, and S_8?*

- *Are these tensions uncorrelated, or connected and different manifestations of a single tension?*

Solving these questions is fundamental to make substantial statements about ΛCDM.

5.6.1. Measuring the systematic errors

In 1-Dimension, the difference between the estimated point $\hat{\theta}_i$ and the true value θ_i^* will be quantified by the bias $B[\hat{\theta}_i]$:

$$B[\hat{\theta}_i] = \hat{\theta}_i - \theta_i^*. \tag{5.42}$$

In the case it is zero for all values of θ_i, the estimator is said to be unbiased. [137] In Fig. 5.12 we illustrate an example.

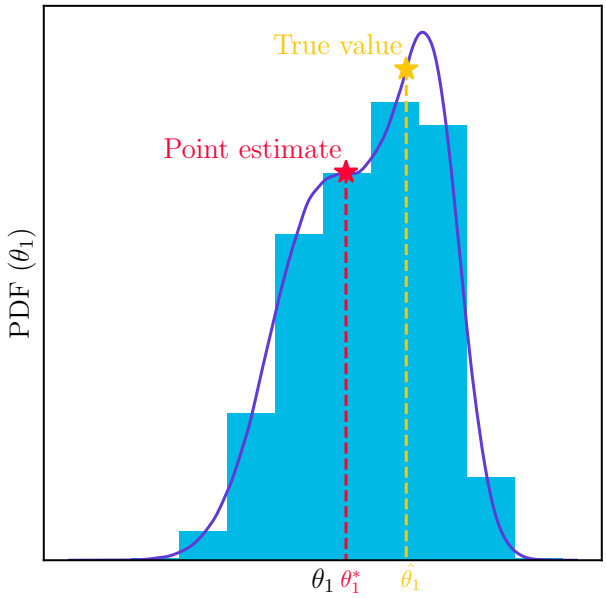


Figure 5.12: In this simple plot, the bias would be equivalent to the difference between $\hat{\theta}_1$ and θ_1^*.

In many dimensions, this *difference* is not trivial to calculate. To define distances between two points in a multivariate space, it is fundamental to have a clear notion of a *metric*. Let x_1 and x_2 be two points in a n-dimensional space. A usual approach to define distances is to use a Minkowski metric of order p [138]:

$$d_p(x_1, x_2) = \left(\sum_{k=1}^{n} |x_{1k} - x_{2k}|^p \right)^{1/p}, \tag{5.43}$$

where p is an integer. Depending on the value of p, we would have specific distances:

Table 5.2: Minkowski distances for different positive integer values of p.

p	$d_P(x_1, x_2)$	Name
$p = 1$	$\sum_{k=1}^{n} \lvert x_{1k} - x_{2k} \rvert$	Manhattan distance
$p = 2$	$\left(\sum_{k=1}^{n} (x_{1k} - x_{2k})^2 \right)^{1/2}$	Euclidean distance
...	...	...
$p \longrightarrow \infty$	$\lim_{p\to\infty} \left(\sum_{i=1}^{n} \lvert x_{1k} - x_{2k} \rvert^p \right)^{1/p} = \max_{i=1}^{n} \lvert x_{1k} - x_{2k} \rvert$	Chebyshev distance

In the context of multivariate statistical analysis, if the errors are known, it is possible to determine weighted distances to account for biases. One possible approach is to use the metric distance between two points, scaling them by the covariance matrix. The most used error-weighted distance is the *Mahalanobis distance*. This expresses the distance between an observation x and an estimator, for example, the mean vector $\boldsymbol{\mu}$:

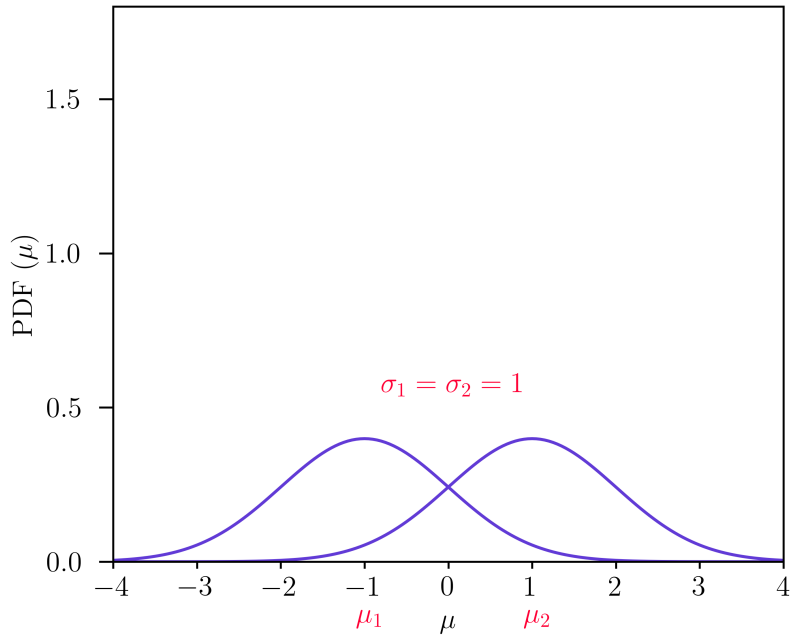
$$d_M(\mathbf{x}, \boldsymbol{\mu}) = \sqrt{(\mathbf{x} - \boldsymbol{\mu})^\intercal \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu})}, \tag{5.44}$$

where $\boldsymbol{\Sigma}$ is the covariance matrix of the sample data $\mathbf{x}$. It was first introduced by P. C. Mahalanobis in 1936 [139]. In 1D, the equation would be:

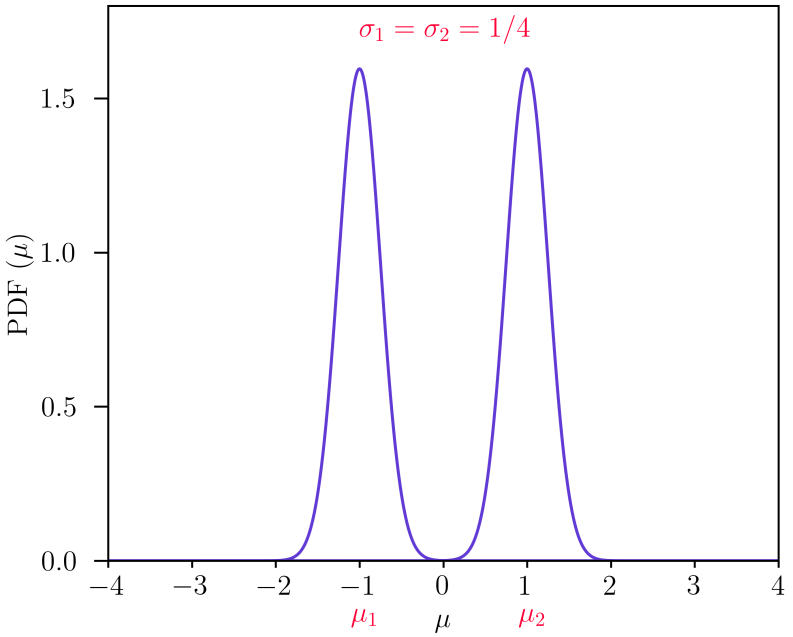
$$d_M(x, \mu) = \frac{|x - \mu|}{\sigma}, \tag{5.45}$$

where σ is the standard deviation of x.

To estimate biases in statistical inferences, it is advantageous to use error-weighted distances rather than the common Minkowski distances. This would change the shape of the Euclidean space to one that accounts for the correlation between the dimensions. In Fig. 5.13 we show a 1D example:



(a) Two Gaussian distributions $\mathcal{N}(\mu_1, \sigma_1) = \mathcal{N}(-1, 1)$ and $\mathcal{N}(\mu_2, \sigma_2) = \mathcal{N}(1, 1)$.



(b) Two Gaussian distributions $\mathcal{N}(\mu_1, \sigma_1) = \mathcal{N}(-1, 1/4)$ and $\mathcal{N}(\mu_2, \sigma_2) = \mathcal{N}(1, 1/4)$

Figure 5.13: Example of two pairs of 1D Gaussians shifted two units away, but with different standard deviations. It is clear that there is a larger inconsistency in the lower panel, which is illustrated with the Mahalonobis distance in Eq. 5.47. The Euclidean distance is not able to capture this effect.

For the classic Euclidean distance, the difference between the means μ_1 and μ_2 of the two

Gaussians from Fig. 5.13 is simple the absolute difference:

$$|\mu_1 - \mu_2| = 2 \tag{5.46}$$

in both cases. On the other hand, the Mahalanobis distance would be:

$$d_M(\mu_1, \mu_2) = \frac{|\mu_1 - \mu_2|}{\sigma_{1,2}} = \begin{cases} 2 & \text{if } \sigma_1 = \sigma_2 = 1 \\ 8 & \text{if } \sigma_1 = \sigma_2 = 1/4 \end{cases}. \tag{5.47}$$

The basic interpretation that can be done is that μ_1 and μ_2 are $2\sigma_{1,2}$ and $8\sigma_{1,2}$ away from the other, respectively.

In many dimensions, it is direct to relate the Mahalanobis distance with the Chi-squared distribution (see Fig. 1 from [140] for visualization purposes):

$$\chi^2_n = \sum_{i=1}^{n} \left(\frac{X_i - \mu_i}{\sigma_i} \right)^2 = (\mathbf{x} - \boldsymbol{\mu})^\intercal \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu}) = d_M(\mathbf{x}, \boldsymbol{\mu})^2. \tag{5.48}$$

Besides quantifying the biases, this geometrical relationship allows us to calculate sample probabilities using the Chi-squared cumulative probability distribution given some confidence level. We will use this in Chapter 7.

Chapter 6

Revision of the German tank problem

Throughout history, scientific knowledge had been used to both save lives but also to produce suffering. During wars, the espionage techniques based in secret intelligence increases enormously. Statistical inference is not free from such acts.
During World War II, by 1943, the Allies tried to estimate the number of tanks produced by the Germanic forces through the Economic Warfare Division of the American Embassy in London [141]. To accomplish this, the Western forces used the printed serial numbers of destroyed or captured tanks that were sequentially numbered. This was a key point to get prepared for the confrontations: an over-estimation of the number of tanks imply committing too many resources, while a sub-estimation would lead to a rapid defeat.

The Allies assumed that the numbers from the printed tanks were sequential and that the first one is numbered 1. Then, estimating the total number of produced tanks would be equivalent to determining the largest number of the considered series. The formula derived for the estimator by the Allies used a frequentist approach, being the results successful.

In this Chapter we present a review analysis of the *tank problem*, that was originally displayed in the weekly seminar of the *Astrophysics, Cosmology, and Artificial Intelligence Group* of Professor Daniel Gruen at the *Universitaets-Sternwarte Muenchen* from the *Ludwig-Maximilians-Universität*. We first list the presuppositions of the problem in Sec. 6.1. We then address it using frequentist inference theory in Sec. 6.2 and then solve the same problem using a Bayesian approach and compare their results in Sec. 6.3. We finally conclude in Sec. 6.4.

6.1. Presuppositions

Let N be the unknown total number of tanks produced, and let us assume each of them had labeled a natural serial number. If we observe n serial numbers from a random sample $(X_1, ..., X_n)$, being X_{max} the maximum observed value $X_{max} = \max(X_1, ..., X_n)$, which would be the best estimate for N?

For this problem, there are $\binom{N}{n}$ ways to choose n numbers from N without any order. To count the number of sets of n serial numbers such that the maximum serial number is x, we have to consider all the possible ways of having x as the maximum number and additional $n-1$ tanks.

Given that one of the n captured tanks must have serial number x, the remaining $n-1$ tanks

must have been selected from the $x-1$ tanks preceding the tank with the max serial number x. Therefore, the subset of sample space such that $x = X_{max}$ is $\binom{x-1}{n-1}$.

Finally, the probability for observing $X_{max} = x$ in a sample of size n given N tanks is:

$$P(X_{max} = x|N) = \frac{\binom{x-1}{n-1}}{\binom{N}{n}}, \text{ if } n \leq x \leq N, \text{ and 0 otherwise.} \tag{6.1}$$

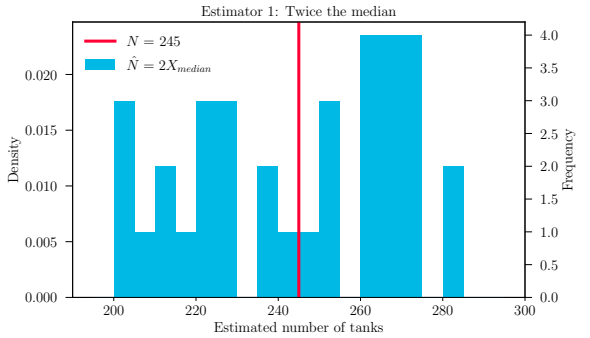
We will first show how distinct frequentist estimators return different values for N and then compare it with the final unbiased estimator (used by the Allies).

The monthly production of tanks through the months of June 1940, June 1941 and August 1942 was of 122, 271 and 342 tanks respectively. [141] The mean production is equivalent to 245 tanks per month. We will do a frequentist and a Bayesian analyses for a total of 245 tanks, from which we will draw 25 random serial numbers from 50 independent observations.

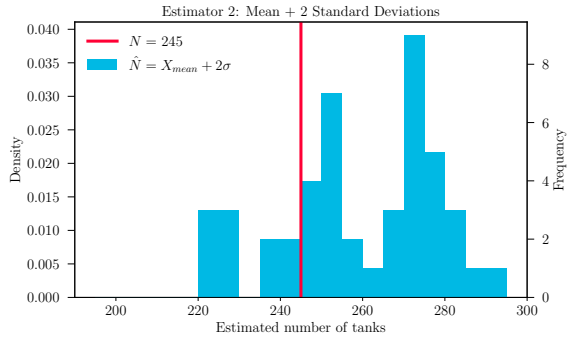
6.2. Frequentist approach

Only using our intuition, it is possible to list different point estimators for $\hat{N}$. In Fig. 6.1 we show the performance of following candidates:

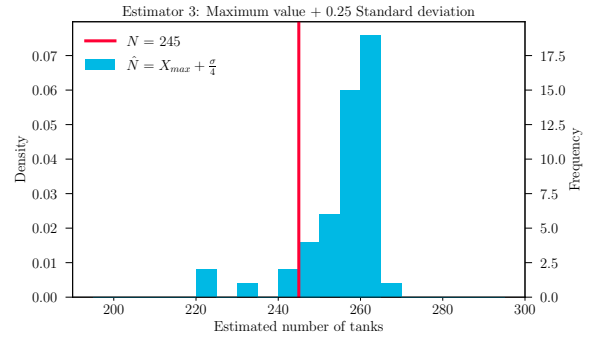
1. Twice the median: $\hat{N} = 2X_{median}$
2. Mean + 2 Standard Deviations: $\hat{N} = X_{mean} + 2\sigma$
3. Maximum value + 0.25 Standard deviation: $\hat{N} = X_{max} + \frac{\sigma}{4}$
4. Maximum value + Minimum value: $\hat{N} = X_{min} + X_{max}$
5. 1.1 times de Maximum value: $\hat{N} = 1.1X_{max}$
6. Median + Inter Quartil Range: $\hat{N} = X_{median} + (Q_3 - Q_1)$



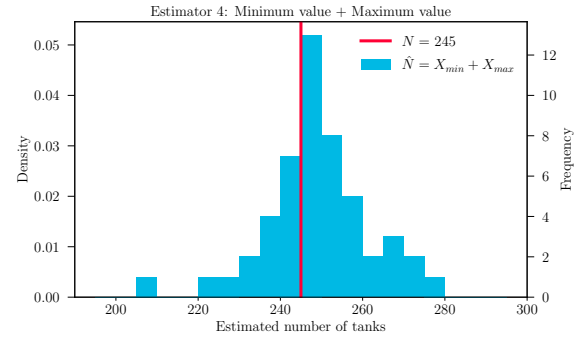
(a) Twice the median



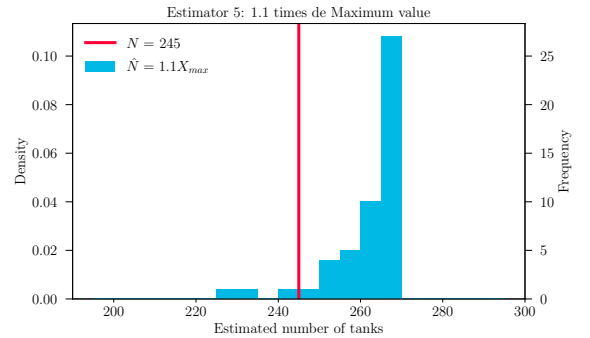
(b) Mean + 2 Standard Deviations



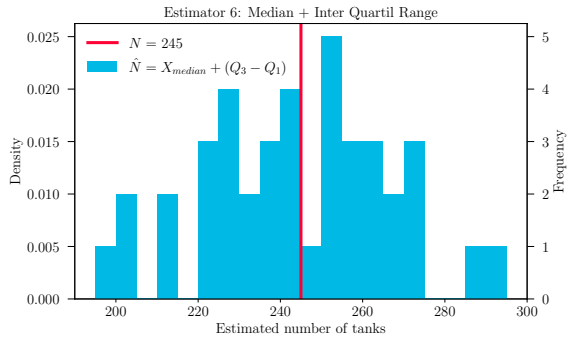
(c) Maximum value + 0.25 Standard deviation



(d) Maximum value + Minimum value



(e) 1.1 times de Maximum value



(f) Median + Inter Quartil Range

Figure 6.1: Different histograms of the proposed frequentist estimators of $\hat{N}$ for a real number of $N = 245$ and a total of 25 random serial numbers of 50 independent observations.

The first case, the twice the median estimate $\hat{N} = 2X_{median}$ lead to the worst estimate results. The deviation from the true value $N = 245$ is large, while there is no clear center. The second, third and fifth cases $\hat{N} = X_{mean} + 2\sigma$, $\hat{N} = X_{max} + \frac{\sigma}{4}$ and $\hat{N} = 1.1X_{max}$ respectively, tend to over-estimate N. On the other hand, the fourth and last cases, $\hat{N} = X_{min} + X_{max}$ and $\hat{N} = X_{median} + (Q_3 - Q_1)$ respectively, are not biased, but their variances are large.

It is possible to infer that the fourth case $\hat{N} = X_{min} + X_{max}$ is the one that shows the best performance. This relies on the fact that *our sample is random and made over an uniform discrete distribution of serial numbers*. This implies that the lowest and higher values will

approximate to the real values as we increase the sample, $X_{min} \sim 1$ and $X_{max} \sim N$, and therefore, $X_{min} + X_{max} \sim N + 1$. From this mathematical fact, we can derive the most successful and unbiased estimator.

Final unbiased estimator

Let us find the best estimate for X_{max}. We calculate the expected value:

$$\mathbb{E}(X_{max}) = \sum_{x=n}^{N} x \cdot P(X_{max} = x). \tag{6.2}$$

We substitute the probability:

$$\mathbb{E}(X_{max}) = \sum_{x=n}^{N} x \frac{\binom{x-1}{n-1}}{\binom{N}{n}}, \tag{6.3}$$

and expand the binomial coefficients:

$$\begin{aligned}
\mathbb{E}(X_{max}) =& \sum_{x=n}^{N} x \frac{(x-1)!}{(n-1)!(x-n)!} \frac{n!(N-n)!}{N!} \\
=& \sum_{x=n}^{N} \frac{x!n}{(x-n)!} \frac{(N-n)!}{N!} \\
=& \frac{n(N-n)!}{N!} \sum_{x=n}^{N} \frac{x!}{(x-n)!} \\
=& \frac{nn!(N-n)!}{N!} \sum_{x=n}^{N} \frac{x!}{(x-n)!n!} \\
=& \frac{nn!(N-n)!}{N!} \sum_{x=n}^{N} \binom{x}{n} \\
=& \frac{nn!(N-n)!}{N!} \frac{(N+1)!}{(n+1)!(N-n)!} \\
=& n \frac{(N+1)}{(n+1)}.
\end{aligned} \tag{6.4}$$

We then isolate N:

$$N = \frac{(n+1)}{n} \mathbb{E}(X_{max}) - 1. \tag{6.5}$$

As our best estimate for $\mathbb{E}(X_{max})$ is actually the observed X_{max}, then

$$\hat{N} = \left(\frac{n+1}{n} \right) X_{max} - 1, \tag{6.6}$$

and the variance of $\hat{N}$ is given by

$$\mathbb{V}(\hat{N}) = \frac{1}{n} \frac{(N-n)(N+1)}{(n+2)}. \tag{6.7}$$

Let us see its performance for the case in which $N = 245$ and the samples are 25 for 50 independent observations:

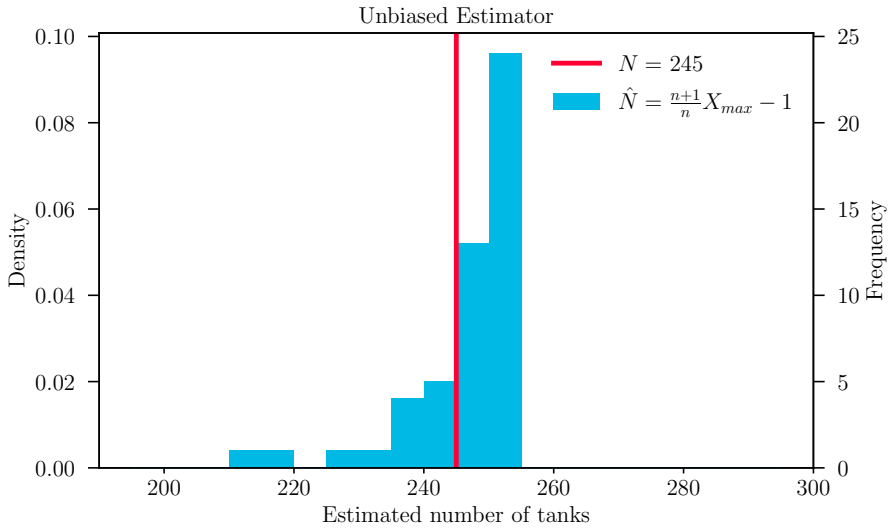


Figure 6.2: Histogram of the unbiased frequentist estimator from Eq. 6.6 for the same quantities as the listed in Fig. 6.1.

The resulting histogram is both, centered and with low variance, about the estimate of N. For a low number of samples, e.g. $n = 1$, the estimate would be equivalent to $\hat{N} \sim 2X_{max} - 1$. This can be interpret that the estimates minimizes the bias using the fact that the sample is uniform and the expected value should rely in the middle of the distribution. Therefore, twice that quantity would be, statistically, close to the true value but with high chances of error. We corroborate the statements here listed augmenting the number of observed serial number per realization in Appx. B.1.

6.3. Bayesian approach

For any Bayesian analysis, it is necessary to assign a prior distribution to the parameters. The posterior distribution can be calculated using the discrete version of the Bayes' theorem from Sec. 4.3.2:

$$\begin{aligned} \mathbb{P}(N|x) =& \frac{\mathbb{P}(x|N)\mathbb{P}(N)}{\mathbb{P}(x)} \\ =& \frac{\mathbb{P}(x|N)\mathbb{P}(N)}{\sum_{N'=x}^{\infty} \mathbb{P}(x|N')\mathbb{P}(N')} \end{aligned} \tag{6.8}$$

for $N = x, x+1, ...$ and 0 otherwise.

The prior knowledge we have of this problem is that the total number of tanks must be a positive integer. Besides of the logical considerations around the German capacity to produce

tanks, it is not easy to define a possible maximum quantity considering how dangerous would be to define an upper bound that is lower to the real value. The simplest case would be then to consider an *improper uniform prior*, i.e., an uniform discrete prior without upper limit. See Appx. B.2 for examples and a discussion about using other priors.

Improper uniform prior

The improper uniform prior is

$$P(N) = 1 \text{ for } N = \{1, 2, ..., \infty\}. \tag{6.9}$$

This would give us the following posterior for Eq. 6.8:

$$P(N|x) = \frac{n-1}{x}\frac{\binom{x}{n}}{\binom{N}{n}} \text{ if } N = x, x+1, x+2, ..., \text{ and 0 otherwise.} \tag{6.10}$$

The demonstration is described below:

$$\begin{aligned} P(N|x) =& \frac{P(x|N)P(N)}{\sum_{N'=x}^{\infty} P(x|N')P(N')} \\ =& \frac{\frac{\binom{x-1}{n-1}}{\binom{N}{n}} \cdot 1}{\sum_{N'=x}^{\infty} \frac{\binom{x-1}{n-1}}{\binom{N'}{n}} \cdot 1} \\ =& \frac{1}{\binom{N}{n}} \frac{1}{\sum_{N'=x}^{\infty} \frac{1}{\binom{N'}{n}}} \end{aligned} \tag{6.11}$$

$$
\begin{aligned}
\sum_{N'=x}^{\infty} \frac{1}{\binom{N'}{n}} =& \sum_{i=0}^{\infty} \frac{1}{\binom{i+x}{n}} \\
=& \sum_{i=0}^{\infty} \frac{n!(x+i-n)!}{(x+i)!} \\
=& \sum_{i=0}^{\infty} \frac{n!(x+i-n)!}{(x+i)!} \cdot \frac{x!(x-n)!}{x!(x-n)!} \\
=& \frac{n!(x-n)!}{x!} \sum_{i=0}^{\infty} \frac{(x+i-n)!x!}{(x+i)!(x-n)!} \\
=& \frac{1}{\binom{x}{n}} \sum_{i=0}^{\infty} \frac{x!}{(x+i)!} \frac{(x+i-n)!}{(x-n)!} \\
=& \frac{1}{\binom{x}{n}} \sum_{i=0}^{\infty} \frac{\Gamma(x+1)}{\Gamma(x+1+i)} \frac{\Gamma(x-n+1+i)}{\Gamma(x-n+1)} \\
=& \frac{1}{\binom{x}{n}} \sum_{i=0}^{\infty} \frac{(x-n+1)_i}{(x+1)_i},
\end{aligned}
\tag{6.12}
$$

where we use that $\Gamma(n) = (n-1)!$ and $(\ )_i$ is the Pochhammer symbol.

$$
\begin{aligned}
\sum_{N'=x}^{\infty} \frac{1}{\binom{N'}{n}} =& \frac{1}{\binom{x}{n}} \sum_{i=0}^{\infty} \frac{(x-n+1)_i}{(x+1)_i} \\
=& \frac{1}{\binom{x}{n}} \sum_{i=0}^{\infty} \frac{(x-n+1)_i}{(x+1)_i} \cdot \frac{i! \cdot 1^i}{i!} \\
=& \frac{1}{\binom{x}{n}} \sum_{i=0}^{\infty} \frac{(x-n+1)_i (1)_i}{(x+1)_i} \cdot \frac{1^i}{i!}, \text{ because } (1)_i = i! \\
=& \frac{1}{\binom{x}{n}} {}_2F_1(x-n+1, 1; x+1; 1),
\end{aligned}
\tag{6.13}
$$

where we used the definition of the hypergeometric series: ${}_2F_1(a, b; c; z) = \sum_{n=0}^{\infty} \frac{(a)_n (b)_n}{(c)_n} \frac{z^n}{n!}$

$$
\begin{aligned}
\sum_{N'=x}^{\infty} \frac{1}{\binom{N'}{n}} =& \frac{1}{\binom{x}{n}} {}_2F_1(x-n+1, 1; x+1; 1) \\
=& \frac{1}{\binom{x}{n}} \frac{\Gamma(x+1)\Gamma(x+1-(x-n+1)-1)}{\Gamma(x+1-(x-n+1))\Gamma(x+1-1)} \\
=& \frac{1}{\binom{x}{n}} \frac{\Gamma(x+1)\Gamma(n-1)}{\Gamma(n)\Gamma(x)} \\
=& \frac{1}{\binom{x}{n}} \frac{x}{n-1},
\end{aligned}
\tag{6.14}
$$

where we use that ${}_2F_1(a,b;c;1) = \frac{\Gamma(c)\Gamma(c-a-b)}{\Gamma(c-a)\Gamma(c-b)}$ and $\Gamma(n+1) = n\Gamma(n)$.

Therefore, the posterior for a improper uniform prior is:

$$P(N|x) = \frac{n-1}{x}\frac{\binom{x}{n}}{\binom{N}{n}} \text{ if } N = x, x+1, x+2, ..., \text{ and 0 otherwise.} \tag{6.15}$$

The posterior mean, calculated by [142], is

$$\mathbb{E}(N|x) = \frac{n-1}{n-2} \cdot (x-1) \text{ for } n > 2, \tag{6.16}$$

and the posterior variance:

$$\mathbb{V}(N|x) = \frac{(n-1)(x-1)(x-n+1)}{(n-2)^2(n-3)} \text{ for } n > 3. \tag{6.17}$$

Due to the layer of mathematical complexity around the prior assignments, we will compare the two approaches (frequentist and Bayesian) for the unbiased estimate and the improper uniform prior respectively. In Fig. 6.3, we show the expected values $\mathbb{E}(N|x)$ for 50 independent cases in which $n = 25$. The results are similar to the ones observed in Fig. 6.2.

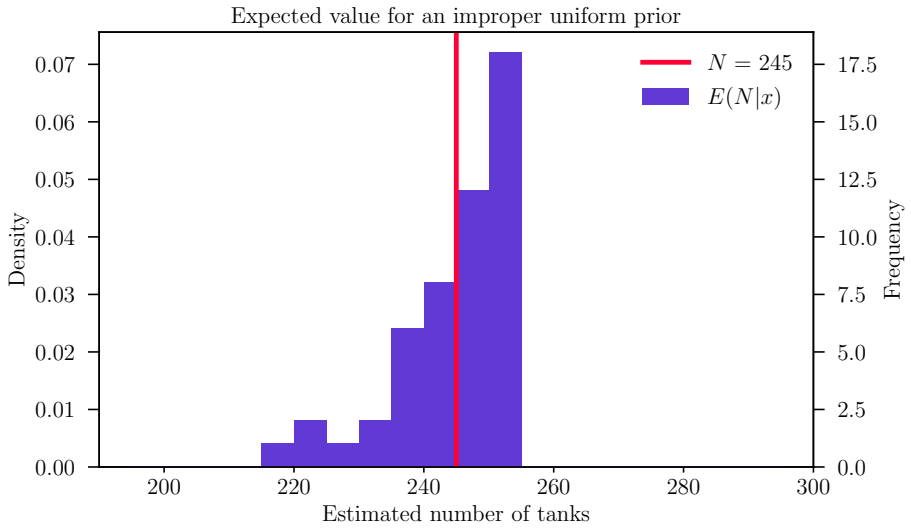


Figure 6.3: Histogram of the Bayesian expectation value of N for 50 independent observations of 25 samples from a total number of 245 for an improper prior.

The two results converge with large values of n because $(n-1)/(n-2) \sim 1$, and therefore, $\mathbb{E}(N|x) \sim X_{max} - 1$. The variance would also vanish with $n \gg 1$. However, for $n = 1, 2$ we find that it is not possible to estimate $\mathbb{E}(N|x)$, while for $n = 3$ we can only access to the point estimate, but not to the associated uncertainty $\mathbb{V}(N|x)$. This divergent behaviour is illustrated in the lower part of Fig. 6.4.

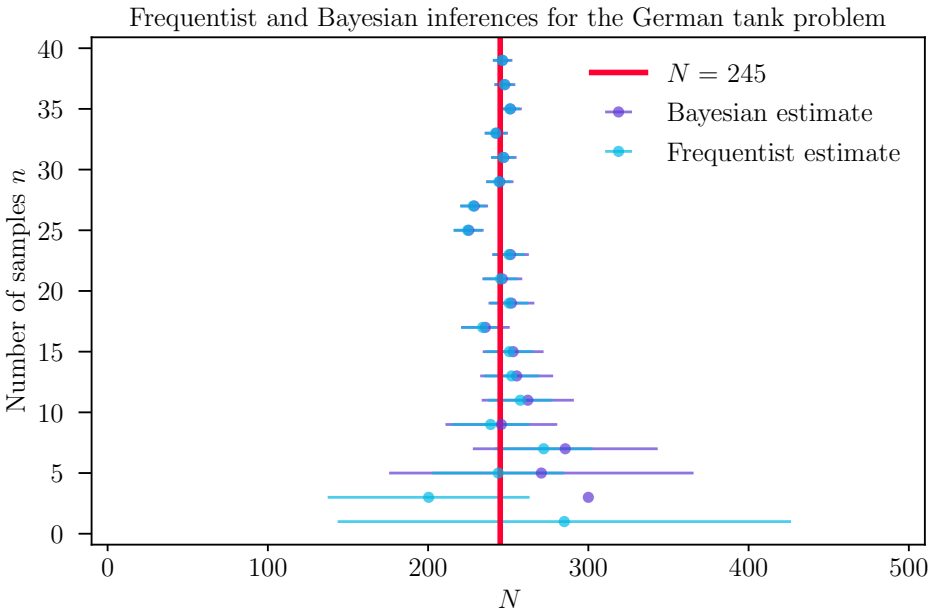


Figure 6.4: Estimates of N from the the frequentist and Bayesian approaches in cyan and purple respectively for different sample sizes n. The plotted values for n are odd numbers, i.e. $n = 1, 3, 5, ...$, and the intervals correspond to the standard deviation centered in each point estimate.

There is a clear general convergence towards the true value N as the number of samples increases. Nevertheless, the case $n = 1$ does not have any estimate in the Bayesian framework, while the case for $n = 3$ only has the corresponding expected value, but non associated uncertainty. This implies that in these cases there is either not available constrain information, or the estimate is certain and there is no way to measure a Mahalonobis distance (see Sec. 5.6.1) to determine how bias is the estimate with respect to its error. Therefore, we have proved that the tank problem can be approached using either frequentist inference or Bayesian inference with success, however leading to different results when $n = 1, 2, 3$. We will discuss the consequences of this in Sec. 6.4.

6.4. Discussions and conclusion

We have showed how to implement the simplest frequentist and Bayesian approximation of the total number of a series from the tank problem, specifically, for the unbiased estimate and the improper uniform prior respectively. These had similar (and converging) results as the number of samples increases and are such that $n \geq 4$, but the frequentist approach was mathematically simpler than the Bayesian one.

On the other side, for smaller samples, the results clearly diverge, and in some cases, the estimates are not even be determinable, leading to an inconsistency between theories when

a improper uniform prior is assumed. Therefore, this is an example in which Bayesian and frequentist analysis of probabilities are not equal.

This problem has a direct impact on what we have discussed throughout chapters 4 and 5. The frequentist perspective of probabilities states that probabilities arise from the law of large numbers when an experiment is repeated enough times to approximate the ratio of occurrences to the total number of cases. The Bayesian interpretation, in the counter part, concepts probabilities as the measure of degree of belief about a proposition. One would expect that under a small amount of repetition of an experiment, the Bayesian framework has a better performance, however, if there is a lack of prior information, which one should one use?

If we assign a proper discrete uniform distribution (i.e. setting an upper bound N_u) to $P(N)$, how sure we can be that we are not underestimating N and, therefore, assigning a null probability to the true value? Or, in a more related discussion to Chapter 7, being non-informative when there is no prior information available?

In the particular case of the tank problem, the frequentist approach is more desirable because, in the absence of prior information, is able to derive a value and a degree of its uncertainty for all the possible number of estimates in a more straightforward way. This case opens questions regarded the analysis of our cosmos: *If we are observing only one specific realization of the universe, which interpretation of probability should we use? Are we using the prior information accurately if we are inside of the same and do not account for random variations? Is it possible to use both approaches complementary to increase our cosmos' understanding?*

This fundamental questions should be part of the itinerary of observational cosmology, due to the strong impact their answers have in our models and in the treatment of systematics. In Chapter 7 we do propose a standardized method to derive informed prior distributions to nuisance parameters enforcing their validation considering, e.g. a validation of a frequentist interpretation of the posterior constraints. We believe this will help to set a direction to overcome the tensions that are arising in the area of cosmology.

Chapter 7

Informed Total-Error-Minimizing (ITEM) Priors: Interpretable cosmological parameter constraints despite complex nuisance effects

The following chapter presents a framework that proposes prior distributions that control projection effects when marginalizing the dimensions of the parameters of interest. We test the method in the shot-noise parametrization from the density split statistics [1, 2], and derive priors to the nuisance parameters that decreased the total error and were informative concerning the defined initial conditions.

The present chapter is under peer revision, and the author list includes Bernardita Ried, Oliver Friedrich, Daniel Gruen, and Dylan Britt. Both OF and DG have provided the data vectors, the covariance matrix, and the projected density maps used in this work to simulate the data from the DSS.

7.1. Abstract

While Bayesian inference techniques are standard in cosmological analyses, the average audience of these analyses interprets resulting parameter constraints with a frequentist intuition. It is increasingly understood that this intuition can fail miserable when marginalising high-dimensional parameter spaces onto subsets of parameters because of what has come to be known as *projection effects* or *prior volume effects*. We present the method of Informed Total-Error-Minimizing (ITEM) priors to address this effect. ITEM priors are prior distributions on a set of nuisance parameters intended to enforce the validity of a frequentist interpretation of the posterior constraints derived for a set of target parameters (e.g. cosmological parameters). Our method works as follows: We split the vector of model parameter into target and nuisance parameters. For a set of realistic data realisations we generate target parameter posteriors using several different candidate priors for the nuisance parameters. Next, we remove candidate priors that do not accomplish minimum requirements of bias (of point estimates) and coverage (of 1-σ confidence regions) for the target parameters. Of the priors that survive this cut we select the ITEM prior as the one that minimizes the total error of the marginalized posteriors of the target parameters.

As an example, we apply our method by re-analyzing the Density Split Statistics (DSS) measured in Dark Energy Survey Year 1 data. We provide ITEM priors needed to model the shot-noise of tracer galaxies that enter the DSS. We demonstrate that the ITEM priors substantially reduce prior volume effects that arise when marginalising over these shot-noise parameters.

7.2. Introduction

Present and future wide-area surveys of galaxies will provide an unprecedented volume of data over most of the extragalactic sky. Examples of these are the Dark Energy Survey (DES)[34] [58, 59], the Vera C. Rubin Observatory's Legacy Survey of Space and Time (LSST)[35] [64, 65], the Dark Energy Spectroscopic Instrument (DESI) survey[36] [63], the space mission EUCLID[37] [143], the 4m Multi-Object Spectroscopic Telescope (4MOST)[38] [60], the Kilo-Degree Survey (KiDS)[39] [60] and the Hyper Suprime-Cam (HSC)[40] [61]. The promise of precision cosmology with these data can only be realized by accurately accounting for nuisance effects – both astrophysical and in the calibration uncertainty of data – with increasingly complex models.

Examples of astrophysical effects include intrinsic alignments of galaxies as a nuisance for weak lensing analyses (see [144] for a recent review and [145–147] for the latest analyses), baryonic physics that impact the statistics of the cosmic matter density field such as star formation, radiative cooling and feedback (e.g. [148–150]), galaxy bias and other parameters describing the galaxy-matter connection (see [151] for a recent review, [135, 152–156]), systematic effects on the baryon acoustic oscillations signal [157–159], or redshift-space distortions [160]. Calibration-related nuisance effects include measurement biases on galaxy shapes (see [161] for a review, [147, 162, 163]), the estimation of redshift distributions of photometric galaxy samples (see Newman and Gruen in prep. for a review and [164–168]), or systematic clustering of galaxies induced by observational effects [169].

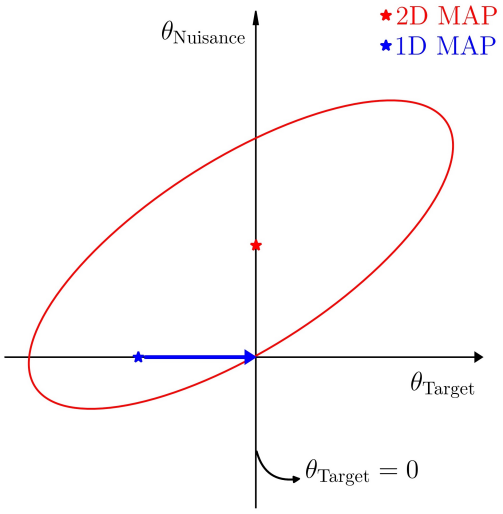
A key step of the cosmological inference from survey data is to perform likelihood analyses that results in a many-dimensional joint posterior distributions of cosmological and nuisance parameters. A marginalized posterior for one or several parameters can be derived from this joint posterior by projecting out the remaining parameters. We might desire for these marginalized posteriors to have certain properties:

- *Unbiasedness:* An exercise that can fail in cosmological analyses with high-dimensional parameter spaces goes as follows. Predict a data vector at a set of fiducial parameters, using a theoretical modelling pipeline. Then derive parameter constraints from that data vector - as if it was a noisy measurement - using the same modelling pipeline. We demand, that the maximum posterior point (MAP) of each cosmological parameter in that exercise have a bias w.r.t. the fiducial parameters that is smaller than some fraction

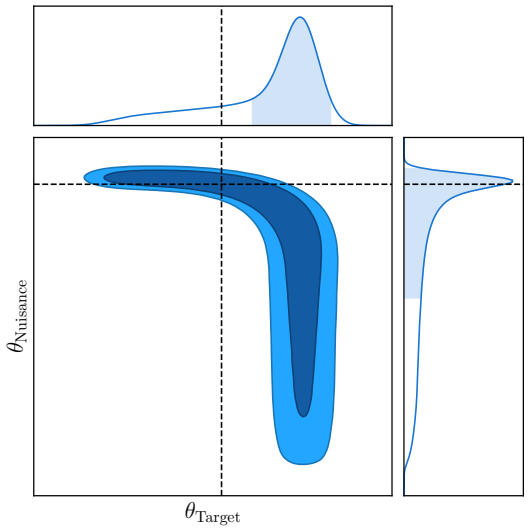
[34] https://www.darkenergysurvey.org/
[35] http://www.lsst.org/lsst
[36] https://www.desi.lbl.gov/
[37] www.cosmos.esa.int/web/euclid, www.euclid-ec.org
[38] https://www.4most.eu/cms/
[39] http://kids.strw.leidenuniv.nl/
[40] https://hsc.mtk.nao.ac.jp/ssp/

of the width of the corresponding confidence region.

- *Reliability:* We demand that in repeated trials the true value of a set of target parameters is within a derived confidence region of those parameters at least some fraction of times.
- *Robustness:* We demand that the previous two requirements hold independent of which of a list of plausible nuisance and physical realizations of the universe is actually realized.
- *Precision:* Among all the possible inference procedures whose confidence regions fulfill the above requirements statements, we prefer the one which, on average, results in the smallest total uncertainty on the target parameters.



(a) **Bias due to underfitting a nuisance effect**: When θ_{Nuisance} is a free parameter of the 2D Gaussian model presented in the panel, the true value of the target (e.g.∼cosmological) parameter $\theta_{\text{Target}} = 0$ is recovered with the 2D Maximum a Posteriori (MAP). With a simpler nuisance model that fixes $\theta_{\text{Nuisance}} = 0$ and thus under-fits the nuisance, the 1D MAP (blue mark) of θ_{Target} can be biased.



(b) **Prior volume effect**: In this toy illustration, the plot shows the 2D and 1D marginalized posterior of a distribution in which the nuisance parameter θ_{Nuisance} presents a large peak. This shape can have the MAP inside the inner contour plots, as it does in the intersection of the black dashed lines. The problem arises when the Probability Density Function (PDF) is projected over θ_{Target} as shown in the right. Even if the posterior in the entire parameter space is centered around the correct parameters, the posterior marginalized over the nuisance parameters can be off of the correct target parameters. Such posteriors can, e.g., result from a model that is non-linear in the nuisance parameters. One way of fixing this projection effect is setting a prior that decreases this effect by constraining the nuisance parameter.

Figure 7.1: Sketches on how the treatment of nuisance parameters can cause systematic errors on the target parameter posteriors.

There are two ways in which the treatment of nuisance parameters can cause *systematic* errors in cosmology and, therefore, violate these statements. The first one is to assume a nuisance

parameter model that is overly simplistic or constrained, e.g.∼fixing a nuisance parameter that indeed needs to be free in order to accurately describe observed data. In the left panel of Fig. 7.1, we present a simple sketch of this situation, in which the assumption that the nuisance parameter $\theta_{\rm Nuisance} = 0$ leads to a bias on the cosmological parameter $\theta_{\rm Target}$. We will call this effect *underfitting* of the nuisance, following [170].

The second effect is the opposite: assuming prior distributions over the nuisance parameters that are too large with respect to the likelihood. When nuisance parameters are degenerate with the cosmological parameters of interest, the projection of the posterior onto the parameters of interest can lead to an offset from their true value in the presence of large prior volumes. This effect is known as *prior volume effect* and is illustrated in the right panel of Fig. 7.1. A characteristic feature of this effect is that the marginalized posterior will have a peak in a region of parameter space where the likelihood profile is flat. For real-life examples of such projection effects, see e.g.∼[145, 171] on intrinsic alignments, [155, 172] on galaxy bias or [173] on galaxy cluster mass profiles.

Another consequence of the prior volume effect is that the *coverage fraction* (the fraction of marginalized confidence intervals in repeated trials that encompass the true target parameters) is not consistent with the associated confidence level. Assuming for the sake of argument that there are indeed true values for the model parameters θ, this can be rephrased as follows: the marginalized 1σ confidence contours obtained with repeated Bayesian analyses in an ensemble of universes do not necessarily contain the true configuration $\theta_{\rm Target}$ of these universes with a frequency of $\sim 68\,\%$. One can indeed construct examples where seemingly reasonable analysis choices lead to vast discrepancies between formal confidence levels and their actual success frequency [174–176]. This may not constitute a problem from a purely Bayesian standpoint since Bayesian statistics do not claim to satisfy frequentist expectations. However, this does not change the fact that one may want to make probabilistic statements about the true value of a physical parameter and that marginalized parameter constraints quoted in cosmological publications are interpreted in a frequentist way by a large fraction (if not the majority) of the cosmology audience. Projection effects in high-dimensional nuisance parameter spaces may hence cause a buildup of wrong intuitions in the inferred parameters.

This situation is further complicated by the fact that a precise model for nuisance effects, with a finite set of parameters and well-motivated Bayesian priors, is often not available. Commonly, at best, a number of plausible configurations of the nuisance effect are known. These could e.g.∼be a set of summary statistics measured from a range of plausible hydrodynamical simulations ([177, 178]) or a compilation of different models and parameters that have been found to approximately describe the nuisance, as is the case for intrinsic alignment [179–181], the galaxy-matter connection ([182–184], Britt et al. in prep.) or in the calibration of redshift distributions [165].

In this paper, we adopt the simple philosophy that the paramount objective of the analysis is to be able to make statements about the target parameters that fulfill the above criteria of: *unbiasedness*, *reliability*, *robustness* and *precision*. Since the nuisance priors we construct to satisfy these criteria is informed by knowledge of possible configurations of the nuisance effects and minimizes the combination of systematic and statistical error, we will call them *Informed Total-Error Minimizing (ITEM) priors*. How we construct these priors has two

important consequences for the interpretation of the resulting parameter constraints:

- In the presence of limited information about the nuisance effect, the considered realizations can be used to tune the priors of the nuisance parameters to ensure certain properties of the resulting constraints on the target parameters. We do not assume the final prior to directly represent information about the nuisance effect. It is merely a tool for the interpretable inference of the target parameters. As a consequence, we also give up on the ability to derive posterior constraints on the nuisance parameters.
- The coverage fraction of any procedure to derive parameter constraints in a set of repeated experiments will be sensitive to the true values of cosmological and nuisance parameters that underlie those experiments. Given a set of plausible realizations of cosmology and nuisances, we can hence only ensure a *minimum coverage* of the marginalized constraints derived from ITEM priors.

The paper is structured as follows: We start by describing the philosophy and general methodology for obtaining ITEM priors in Sec. 7.3. In Sec. 7.4 we investigate the performance of the method in a test case: cosmological parameters constraints marginalized over models for non-Poissonian shot-noise in Density Split Statistics (DSS, [1, 2]). We explore the change in simulated and DES Y1 results of these statistics when using ITEM priors. Finally, discussion and conclusions are given in Sec. 7.5.

7.3. Methodology

Assume that some observational data can be described in a data vector $\hat{\boldsymbol{\xi}}$ of N_d data points. Let $\boldsymbol{\xi}[\boldsymbol{\theta}]$ be a model for this data vector that depends on a parameter vector $\boldsymbol{\theta}$ of N_p parameters, and let Cov be the covariance matrix of $\hat{\boldsymbol{\xi}}$. In many cases, it is reasonable to assume that the likelihood $\mathcal{L}$ of finding $\hat{\boldsymbol{\xi}}$ given the parameters $\boldsymbol{\theta}$ is that of a multivariate Gaussian distribution:

$$\ln \mathcal{L}(\hat{\boldsymbol{\xi}}|\boldsymbol{\theta}) = -\frac{1}{2}\sum_{i,j}[\hat{\xi}_i - \xi_i[\theta]]\mathrm{C}_{ij}^{-1}[\hat{\xi}_j - \xi_j[\theta]] + \mathcal{C} \ , \tag{7.1}$$

where C_{ij}^{-1} are the elements of the inverse of the covariance matrix, and $\mathcal{C}$ is a constant as long as the covariance does not depend on the model parameters. We derive the posterior distribution $\mathcal{P}(\hat{\boldsymbol{\xi}}|\boldsymbol{\theta})$ using Bayes' theorem:

$$\mathcal{P}(\boldsymbol{\theta}|\hat{\boldsymbol{\xi}}) \propto \mathcal{L}(\hat{\boldsymbol{\xi}}|\boldsymbol{\theta})\Pi(\boldsymbol{\theta}) \ , \tag{7.2}$$

with $\Pi(\boldsymbol{\theta})$ being a prior probability distribution incorporating a priori knowledge or assumptions on $\boldsymbol{\theta}$.

In many situations the parameters $\boldsymbol{\theta}$ can be divided into two types: target parameters $\boldsymbol{\theta}_\mathcal{T}$ and nuisance parameters $\boldsymbol{\theta}_\mathcal{N}$. This distinction is usually subjective and depends on which parameters one would like to make well-founded statements about (i.e., robust, valid, reliable, and/or precise estimates) and which parameters are needed to describe the complexities of the experiment but are not considered intrinsically of interest. For example, in a cosmological experiment, the target parameters are usually those describing the cosmological model, such

as ΛCDM and its variations. This could, e.g., include the present-day density parameters (Ω_x), the Hubble constant (H_0), and the equation of state parameter of dark energy (w). On the other hand, the nuisance parameters would be used to account for observational and astrophysical effects or uncertainties in the modeling of the data vector. Nevertheless, this target-nuisance classification is not fixed but depends on the focus of the analysis.

Recent cosmological analyses have proceeded by choosing an approximate nuisance model with a limited number of free parameters for all these complex effects and either assuming a wide flat prior distribution over them or a prior distribution set by simulation or calibration measurements [15, 44, 146, 185].

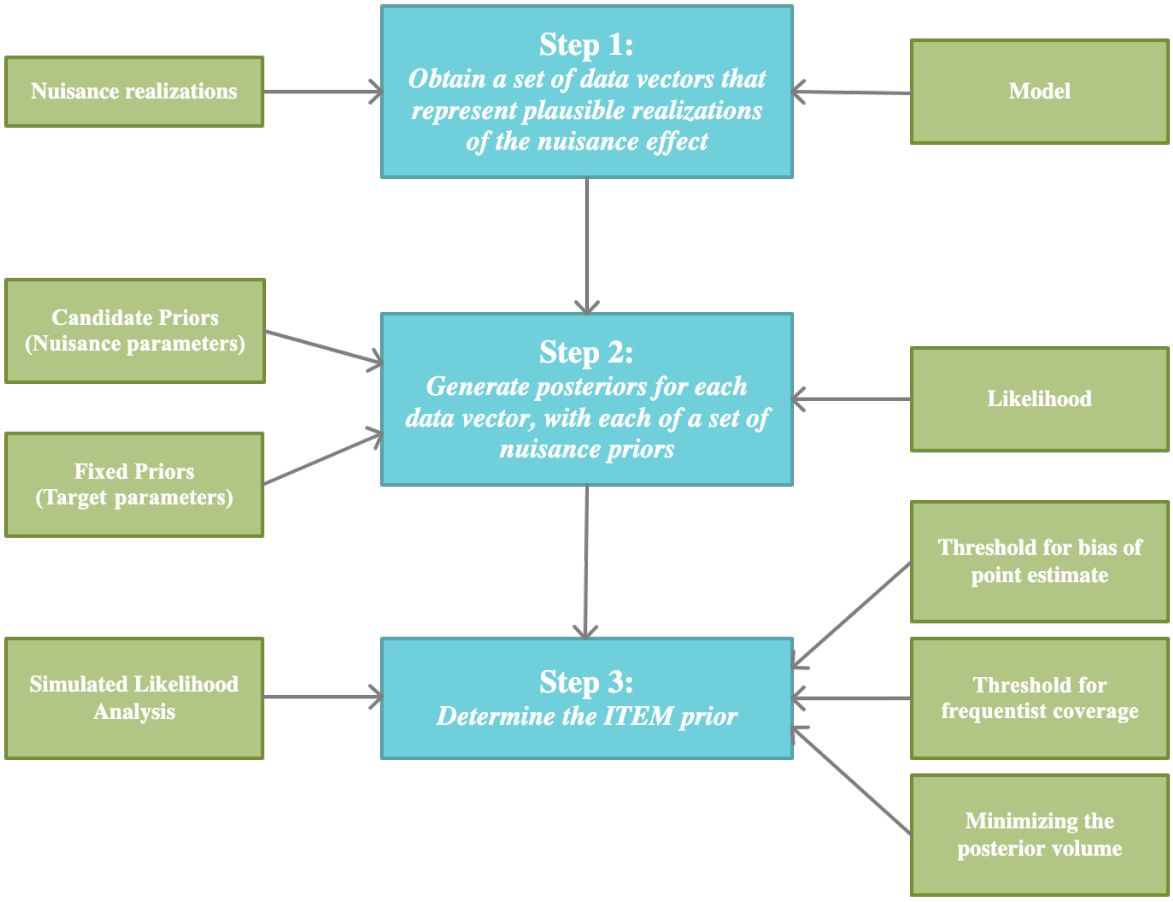


Figure 7.2: Basic steps and concepts present in our methodology. In light blue, we summarize the steps of our pipeline and highlight other ingredients of the methodology with green.

As we explained in Sec. 7.2, precarious modeling of nuisance parameters can cause systematic errors, leading to the following consequences:

- *Limited modeling of the nuisance effects*: The assumed model may not sufficiently describe the nuisance effects actually realized. For example, there could be an oversight of one underlying source of the nuisance, the model of a nuisance effect may be truncated

at a low order, or an effect entirely neglected. This could result in an intrinsic bias on the inferred target parameters, as would e.g.∼be revealed by repeating the analysis on simulated data that include the full nuisance effect.

- *Prior-volume effects*: A wide prior over one or multiple nuisance or target parameters can lead to projection effects when analyzing the 1D marginalized posteriors of target parameters if the model does not depend linearly on the parameters. Even if the prior is physically motivated and reflects our knowledge on a nuisance effect, it could still cause a significant projection effect. And a different nuisance prior could result in a less uncertain yet more reliable inference of the target parameters.

We will now introduce a procedure aimed to identify prior distributions on nuisance parameters which ensure the criteria we presented in Sec. 7.2 for marginalized constraints on target parameters (see also Fig. 7.2). We will call the resulting priors *Informed Total-Error Minimizing* priors, or ITEM priors. As explained in Sec. 7.2, our procedure degrades nuisance priors to mere tools that ensure interpretability of the marginalised constraints on target parameters in a Bayesian likelihood analysis. They do not represent actual prior knowledge on the nuisance parameters and we also give up on the ability to derive posterior constraints on them.

7.3.1. Step 1: Obtain a set of data vectors that represent plausible realizations of the nuisance effect

Assume that at fixed target (e.g.∼cosmological) parameters $\boldsymbol{\theta}_{\mathcal{T}} = (\theta^1_{\mathcal{T}}, \theta^2_{\mathcal{T}}, ..., \theta^{N_t}_{\mathcal{T}})$, the range of plausible realisations of a nuisance effect is well represented by a set of N different resulting data vectors

$$\hat{\boldsymbol{\xi}}^i = \boldsymbol{\xi}[\boldsymbol{\theta}_{\mathcal{T}}, \text{nuisance realization } i], \text{ for } i = 1, ..., N \ ; \tag{7.3}$$

These could, e.g.∼be based on a model for the nuisance effect with different nuisance parameter values. Alternatively, the $\hat{\boldsymbol{\xi}}^i$ could result from a set of simulations with different assumptions that produce realizations of the nuisance effect (e.g.∼for the impact of baryonic physics on the power spectrum as measures in a variety of hydrodynamical simulations, see for example [186]). In some cases, direct measurement of the nuisance effect may be used, and e.g.∼a bootstrapping of those measurements could yield a set of possible realizations.

For the sake of concreteness, let us assume that the nuisance realizations are given in terms of particular nuisance parameter combinations $\boldsymbol{\theta}^i_{\mathcal{N}} = (\theta^{1,i}_{\mathcal{N}}, \theta^{2,i}_{\mathcal{N}}, ..., \theta^{N_n,i}_{\mathcal{N}})$ of the considered nuisance model:

$$\hat{\boldsymbol{\xi}}^i = \hat{\boldsymbol{\xi}}[\boldsymbol{\theta}_{\mathcal{T}}, \boldsymbol{\theta}^i_{\mathcal{N}}] \ . \tag{7.4}$$

In principle, it is possible to consider more than one realization of the target parameters as well (e.g., more than one cosmology for the same nuisance parametrization). We will discuss this extra layer of complexity that can be added to the pipeline in Sec. 7.5, but as a first pass, we will limit ourselves to a fixed target parameter vector for simplicity.

For the following tests to be meaningful, we require these data vectors to contain no noise, or noise that is negligible compared to the covariance considered in the likelihood analysis.

7.3.2. Step 2: Devise a family of nuisance priors among which to choose the ITEM priors

We will choose our ITEM prior among the *space of nuisance priors*, i.e.∼the space of the parameters that describe a family of considered prior distributions $\Pi(\boldsymbol{\theta}_{\mathcal{N}})$. The details of that family may depend on one's preferred functional form of the priors, on consistency relations that are a priori know, and on practical considerations like limits in the computational power that is available to find the ITEM prior. As mentioned before, the distinction between target and nuisance is also subjective, and these labels may change between iterations of the pipeline (i.e., a prior on a target parameter can be fixed for one analysis, but it can be studied and optimized as a nuisance in another).

For example, one may search for an ITEM prior among the set of uniform prior distributions for one particular nuisance parameter, $\Pi(\theta_{\mathcal{N}}) \sim \mathcal{U}(a, b)$. In that case, we propose to sample uniformly and independently over both a and b, subject to the consistency relation that $a < b$.

In practice, we can e.g. sample n lower bounds ($a_i \in [a_1, a_2, ..., a_n]$) and m upper bounds ($b_j \in [b_1, b_2, ..., b_m]$). Besides requiring that $a_i < b_j$ for all tested pairs one could demanded that other requirements are met, e.g. a minimum and maximum value for the prior width $b_j - a_i$, or that some nuisance parameter values are contained within the considered priors. When combining all these configurations, we will have a total of $M \leq n \cdot m$ possible priors with different widths and midpoints.

If we consider uniform priors for a set of nuisance parameters $\theta^l_{\mathcal{N}}$, with $l = 1, \ldots, N_n$, then we would obtain sets of M_l possible priors per nuisance parameter:

$$\Pi_k(\theta^l_{\mathcal{N}}) = \mathcal{U}(\theta^l_{\mathcal{N}}|a^l_k, b^l_k), \text{ for } k = 1, ..., M_l \ . \tag{7.5}$$

The total number of considered priors would then be

$$M = \prod_{l=1}^{N_n} M_l \ . \tag{7.6}$$

For any given measurement of the data $\hat{\boldsymbol{\xi}}$ this will result in M different posteriors on the full set of parameters $\boldsymbol{\theta}$.

$$\mathcal{P}_j(\boldsymbol{\theta}|\hat{\boldsymbol{\xi}}) \propto \mathcal{L}(\hat{\boldsymbol{\xi}}|\boldsymbol{\theta})\Pi_j(\boldsymbol{\theta}_{\mathcal{N}}) \text{ for } j = 1, ..., M \ . \tag{7.7}$$

In the next steps of our pipeline we will generates such sets of posteriors for different data vector realisations in order to determine, which priors meet the criteria we outlined in Sec. 7.2. To obtain all these posteriors, instead of running a large number of Monte-Carlo-Markov-Chains (MCMCs), one can use importance sampling [187, 188] performed over the posterior that was derived from the widest prior (i.e., the prior on the nuisance that encloses all the other priors). While we have focused here on a family of uniform priors, it is straightforward to generalise our pipeline to e.g. Gaussian prior distributions, or other functional forms that are commonly used in cosmological analyses.

7.3.3. Step 3: Determine the ITEM prior

Our goal is to find a prior on the nuisance parameters that returns robust, valid, reliable, and precise posteriors for the target parameters. To determine this ITEM prior, we apply a number of filters to the considered family of prior distributions. Among them are two optional binary selections – setting an upper limit for the bias and a lower limit for the coverage fraction of the posterior (cf. below for precise definitions of bias and coverage). The third is an optimization that minimizes the total error, i.e.∼the combination of statistical uncertainty and bias, on the target parameters. In Fig. 7.3 we synthesize the process in a diagram.

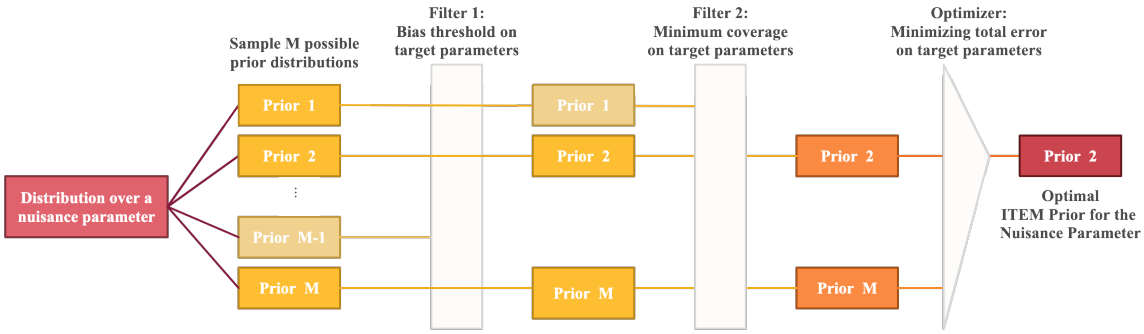


Figure 7.3: Flowchart of the procedure to obtain the ITEM prior. First, a set of M candidate prior distributions should be proposed over the nuisance parameters. Then, the first criterion will filter some of the priors by requiring that the bias on the target parameters in a simulated likelihood analysis be less than x^* times its posterior's statistical uncertainty (see Eq. 7.9). Secondly, we will require that the remaining priors result in a minimum coverage y^* (see Eq. 7.11) over the target quantities. This leads to fewer candidate priors, which are refined finally by minimizing the error they produce on the estimation of the target parameters, including statistical and systematic error contributions (see Eq. 7.14). The optimal one is what we call the ITEM prior.

- **Criterion 1: Maximum bias of point estimate**

  Constraints on target parameters are commonly reported as confidence intervals around a point estimate for those parameters. For example, that point estimate may be the maximum of the marginalized posterior (MAP) or its projected mean. To devise a measure for the bias of this point estimate w.r.t. the true values of the target parameters let us return to the exercise that was already mentioned in Sec. 7.2: Consider a noise-free realisation $\boldsymbol{\xi}$ of a data vector, and a data vector model $\boldsymbol{\xi}_m(\boldsymbol{\theta})$, which we assume to model the data perfectly, if the correct parameters $\boldsymbol{\theta}$ are known. Now use this model to run an MCMC around $\boldsymbol{\xi}$, as if the latter was a noisy measurement of the data. If the model $\boldsymbol{\xi}_m(\boldsymbol{\theta})$ is non-linear in the nuisance parameters, then e.g. the MAP of the marginalised posterior for the target parameters obtained from that MCMC can be biased w.r.t. to

the true parameters (even though $\boldsymbol{\xi}$ is noiseless and the model $\boldsymbol{\xi}_m(\boldsymbol{\theta})$ is assumed to be perfect).

In Sec. 7.3.1 we have obtained a set of realisations $\hat{\boldsymbol{\xi}}^i$ ($i = 1, \ldots, N$) of the data, corresponding to N different realisations of the considered nuisance effects. We will perform the above exercise for each of those data vectors and derive point estimates $\hat{\boldsymbol{\theta}}^i_T$ from each of the resulting chains. The difference between these estimates and the true target parameters underlying our data vectors will then serve as a measure for how much projection effects (or prior volume effects) bias our inference procedure. For now, we leave the exact point estimator used in this procedure unspecified. But in our example of Sec. 7.4 we will assume the $\hat{\boldsymbol{\theta}}^i_T$ to be the MAPs of the marginalised posteriors.

If $\boldsymbol{\theta}_\mathcal{T}$ are the true target parameters and $\hat{\boldsymbol{\theta}}^i_\mathcal{T}$ are the point estimate from nuisance realisation i, then we will use the Mahalanobis distance [139, 140] to quantify the bias between the two, i.e.

$$x_i = \sqrt{(\boldsymbol{\theta}_\mathcal{T} - \hat{\boldsymbol{\theta}}^i_\mathcal{T})^\top \mathbf{C}_i^{-1} (\boldsymbol{\theta}_\mathcal{T} - \hat{\boldsymbol{\theta}}^i_\mathcal{T})} \; . \tag{7.8}$$

Here $\mathbf{C}_i^{-1}$ is the inverse covariance matrix of the parameter posterior, which we directly estimate from the MCMC for nuisance realisation i.

The value of x_i measures how much the point estimate deviates from the truth compared to the overall width (and orientation) of the posterior constraints. The first filter we apply to our family of potential nuisance priors is now to demand that all x_i be smaller than a maximum threshold, which we denote with x^*. This means that we calculate x_i for each realization of the nuisance. We then determine the maximum

$$x = \max_i \{x_i\}, \tag{7.9}$$

of those Mahalanobis distances, and all candidate priors that do not meet the requirement

$$x \leq x^*, \tag{7.10}$$

will be excluded. In other words, we demand that the prior results in less than x^* bias for any realization of the nuisance with respect to the target parameter uncertainties. That should leave us with a total of $M' \leq M$ candidate priors. The ratio M'/M will depend on the considered threshold quantity x^*. In the DES year-3 2-point function analyses, this threshold was chosen as 0.3 for the joint, marginalised posterior of S_8 and Ω_m [171], and we will choose the same value for most of our example case in Sec. 7.4.

- **Criterion 2: Minimum frequentist coverage**

In our quest to find the ITEM prior we now apply a second filter to the candidate priors which is concerned with the coverage probability of the 1-σ confidence region of the marginalised posterior distribution of the target parameters (i.e. the smallest sub-volume of target parameter space that still encloses about 68 % of posterior). This coverage probability measures, how often the true target parameters would be found within that confidence region in repeated trials. In a Bayesian analysis it can in general not be expected that this coverage matches the confidence level of the considered region (in our case 68 %). This has e.g. be extensively discussed by [176].

To measure the coverage probability, we proceed as follows. For each noise-less nuisance realisation of the data, $\boldsymbol{\xi}_i$, we generate a number of $N_{\rm noise}$ noisy realisations by adding multi-variate Gaussian noise according to our fiducial data covariance. For each of these realisation we generate an MCMC and determine the 1-σ confidence region. The fraction y_i of these $N_{\rm noise}$ confidence regions that contain the true target parameters is then an estimate for the coverage probability that is to be expected, if the nuisance model i was realised. In a spirit similar to our previous filter we consider the minimum coverage per prior,

$$y = \min_i \{y_i\} \ , \tag{7.11}$$

and then require that

$$y \geq y^*, \tag{7.12}$$

in order for a candidate nuisance prior to pass our coverage criterion. In other words, we will demand that all priors result in a coverage of at least y^*, regardless of which nuisance realisation is actual. That should leave us with a total of $M'' \leq M'$ candidate priors.

Note that part of any mismatch between coverage probability and confidence level will be caused by the target parameter bias that we have considered in the previous subsection. We would like to factor out this contribution by taking into account the bias as a source of systematic uncertainty in the width of our target parameter posteriors. To motivate a practical way to do so, let us assume that our target posteriors are multivariate Gaussian distributions. In repeated trials, the Mahalanobis distance (cf. Eq. 7.8) between the point estimates and the true parameters would follow a non-central χ^2 distribution, with an offset given by the target parameter bias x_i (i.e. the bias measured also in the Mahalanobis distance). The widths of our posteriors do not know about this systematic error, and correspondingly the 68 % confidence regions will have a coverage of $< 68\,\%$. To counteract this, we can simply raise the confidence level used to derive those regions. To calculate the required confidence level, we proceed ass follows:

- calculate, for which value of χ^2 the cumulative distribution function of a non-central χ^2-distribution, centered at x_i, takes a value of 68 % ,
- at that value, calculate the value of cumulative distribution function of a central χ^2-distribution (with the same number of degrees of freedom). This will generally be higher than 68 %.

It is this higher confidence level which we use to calculate the 1-σ confidence region for the noisy realisations of our data vectors $\boldsymbol{\xi}_i$ and to measure the coverage probabilities y_i. For a bias of $x_i = 0.3$ and two degrees of freedom, we would e.g. require a confidence level of 69.6 % to calculate a confidence region for which we would expect a 68 %. This expectation is however only valid if the target posteriors were indeed Gaussian; non-Gaussianity of the posteriors can lead to mismatches between confidence level and coverage probability beyond this mere biasing of the marginalised posterior.

- **Criterion 3: Minimizing the posterior volume**

The candidate priors that have passed our previous two filters exhibit our definitions of *unbiasedness*, *reliability* and *robustness*. The ITEM prior will now be the one of the remaining priors that yields the smallest posterior uncertainty on the target parameters. This final selection step is necessary because it is in fact easier to meet our previous criteria with wide posteriors that result from unconstrained nuisance priors.

We quantify the target parameter uncertainty by the volume that the 1-σ confidence contours (68 %) enclose. As done in Sec. 7.3.3, we considered the bias and calculate an expanded volume using a larger confidence region. This allows us to optimize both, the random and systematic error. For a Gaussian posterior, the volume would be given in terms of the parameter covariance matrix $\mathbf{C}$ as (cf. [189] for a derivation)

$$V(N) = \pi \det(\mathbf{C})^{1/2} \ , \tag{7.13}$$

and for simplicity we will use that formula also for our (potentially non-Gaussian) posteriors.

Let $V_i(N_t)$ be the volume enclosed by the bias-expanded 1-σ ellipsoid of the ith data vector in the posterior distributions that were obtained in Sec. 7.3.3. Then, for each candidate prior, we propose to consider the mean of these volumes[41], over all data vectors,

$$\overline{V(N_t)} = \frac{\sum_{i=1}^{N} V_i(N_t)}{N} \ . \tag{7.14}$$

We will then choose the prior that has the minimum $V(N_t)$ at the ITEM prior. This would be the nuisance prior distribution that will return, statistically, the combined most unbiased, precise, and certain estimates in simulations for the target parameters.

7.4. ITEM Priors in the Density Split Statistics

In this section, we examine the determination and performance of the ITEM prior FOR a higher-order statistic of the galaxy-matter density called Density Splits Statistics.

[41] Other quantities, such as the maximum volume calculated per data vector, could be used alternatively.

7.4.1. Density Splits Statistics

We test our ITEM prior methodology on the galaxy-matter connection model employed for the Density Split Statistics (DSS) [1, 2], a framework that studies counts-in-cells (CiC) in combination with lensing-around-cells. DSS obtains data vectors that are compressed versions of the full joint PDF of matter density and galaxy count. For an overview of DSS, see Appx. C.

7.4.2. Modeling of the nuisance quantities in DSS

In this subsection, we demonstrate the necessity of proposing alternatives to the prior modeling of the nuisance of the DSS due to the presence of prior volume effects.

Our first analysis is limited to the five quantities listed in Table 7.1. The cosmological parameters are Ω_m[42] and σ_8[43]. The parameters describing the galaxy-matter connection are the linear bias b[44], and the stochasticity parameters α_0 and α_1. These two last parameters were introduced in the DSS modeling to have a higher-order statistic on the galaxy-matter connection and do not have physically motivated prior distributions.

Table 7.1: The parameters used in the simulated likelihood analyses for the α *Model* are presented here. We list their original prior ranges (using $\mathcal{U}$ to denote a uniform prior), and the fiducial values used to simulate the synthetic data for the two cases of stochasticity. The priors and the nuisance realizations are chosen to be either the same or the derived results from [1, 2]. For the ITEM prior derivation, the cosmological parameters would be the target ones, while the stochasticity parameters would be the nuisance ones. Ω_m, S_8 and b will have fixed priors and unchanged values for the parameter vectors.

	DSS Original Prior Distribution	Parameter Vector 1 *Non-Stochasticity*	Parameter Vector 2 *Buzzard Stochasticity*
Target Parameters			
Cosmological Parameters			
Ω_m	$\mathcal{U}$[0.1,0.9]	0.286	0.286
σ_8	$\mathcal{U}$[0.2,1.6]	0.82	0.82
Nuisance Parameters			
Tracer Galaxies			
b	$\mathcal{U}$[0.8,2.5]	1.54	1.54
Stochasticity			
α_0	$\mathcal{U}$[0.1,3.0]	1.0	1.26
α_1	$\mathcal{U}$[-1.0,4.0]	0.0	0.29

[42] The fractional energy density of matter, obtained when the total density of matter/energy needed for the universe to be spatially flat, is normalized. It includes baryonic and dark matter.

[43] The present-day linear theory root-mean-square amplitude of relative matter density fluctuations in spheres of radius 8 h^{-1} Mpc.

[44] The square root of the ratio of the galaxy and matter auto-correlation functions.

As a derived parameter, we also consider S_8[45]:

$$S_8 = \sigma_8 \sqrt{\Omega_m / 0.3} \ . \tag{7.15}$$

Both Ω_m and S_8 will be the target parameters, while α_0 and α_1 are the nuisance parameters whose priors we optimize. The prior on b will remain unchanged.

A fundamental ingredient to the ITEM prior method is using different realizations of the nuisance. We limit our analysis by considering only two realizations of the stochasticity listed in the third and fourth column of Table 7.1.

The first one corresponds to a configuration of the Universe in which there is no stochasticity in the distribution of galaxies, a common assumption on sufficiently large scales. That is equivalent to set $\alpha_0 = 1.0$ and $\alpha_1 = 0.0$. These two values recover the Poissonian distribution from Eq. C.5. We will call this the *Non-Stochasticity* case, and it will be the one usually shown in figures in this paper unless otherwise indicated.

The second one will be the best-fit shot-noise parameters found by [2] on REDMAGIC mock catalogs [193], constructed given realistic DES Y1-like survey simulations called Buzzard-v1.1 [194]. These Buzzard mock galaxy catalogs[46] have been used extensively in DES analyses [196]. The values for the stochasticity parameters are $\alpha_0 = 1.26$ and $\alpha_1 = 0.29$. We will call this the *Buzzard Stochasticity* case.

These two realizations should be considered a simple test case for the methodology. This work emphasizes that the resulting ITEM prior is limited because of the few considered nuisance realizations. We believe a wide range of plausible stochasticities configurations could be taken into account. In a companion paper, we will add diverse realizations of the shot-noise, motivated by different halo occupation distribution (HOD) simulations (Britt et al. in prep.).

The priors presented on the second column of Table 7.1 are the ones used originally in the DSS papers [1, 2]. We ran four MCMC chains, from which two were simulated fixing the stochasticity parameters, while the remaining considered them free parameters with the original broad priors. In Fig. 7.4 we find that in the cases in which we considered free stochasticity parameters, the posterior returns more significant uncertainties over the cosmological parameters Ω_m and S_8.

To test the errors of our simulated likelihood analysis, we consider a 2D Gaussian distribution centered at the 2D marginalized MAP and with the covariance matrix from the original chains (this would give us a point estimator and confidence contours of the target parameters). The estimators return considerable biases for the two shot-noise configurations: Fig. 7.5 illustrates this for each chain with their original large priors. Both cases are highly biased: in the case in which there is no stochasticity, the bias reaches a value of $0.66\sigma_{2D}$, more than two times the DES threshold of $0.3\sigma_{2D}$.

[45] The parameter S_8 has been widely used these last years because it breaks degeneracies that Ω_m with σ_8. It has also been manifested that there is a recent tension between the estimations of those quantities measured in the late universe [44, 135, 190] and the CMB [146, 191, 192].

[46] For a recent 2.0 version of the Buzzard see [195].

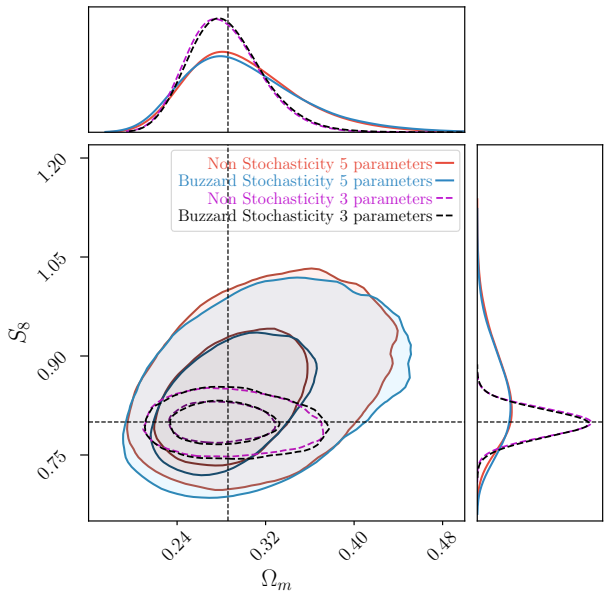


Figure 7.4: Marginalized (Ω_m, S_8) posteriors with 1σ and 2σ contours from simulated DSS α *Model* analyses of a synthetic, noiseless baseline data vector using the fiducial values and priors from Table 7.1. The red and blue contours show the performance of the model with 5 degrees of freedom, while for the purple and black contours, the stochasticity values were fixed to the input $\alpha_0 = [1.0, 1.26]$ and $\alpha_1 = [0.0, 0.29]$ respectively.

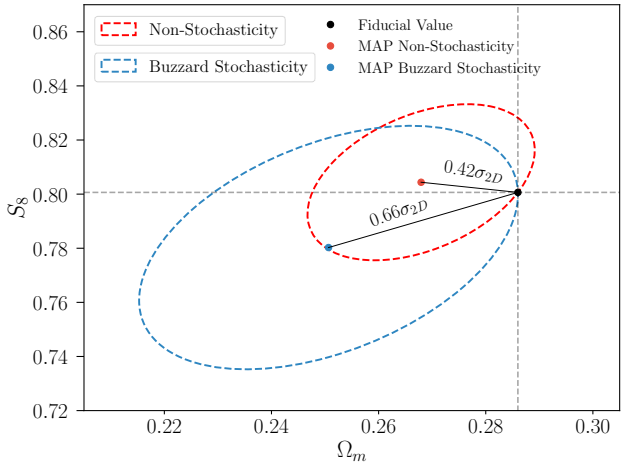


Figure 7.5: Parameter biases in simulated likelihood analyses: in red and blue, we show the ellipses for the 2D marginalized constraints of the non-stochasticity and the buzzard stochasticity configurations, respectively. These are centered on their corresponding 2D MAP. Because of prior volume effects, the marginalized constraints are not centered on the input cosmology.

In a second analysis we assumed the r *Model* from the DSS (detailed at Appx. C) as basis, yet still use the α parameters detailed at Table 7.1 to generate the studied data vectors. With this approach, we study how much information the ITEM prior can recover when there is explicitly an intrinsic systematic.

This model is more straightforward than the α *Model* because it only considers one correlation parameter r for the effect of stochasticity. One interesting aspect of this test thus is that it checks whether the conditions demanded by the ITEM formalism can be met with such a model that did not produce the analyzed data.

The original DSS prior used on r follows an uniform distribution: $\Pi(r) \sim \mathcal{U}[0.0, 1.0]$. Mathematically, we would require $r \leq 1$, but this sharp upper bound will intrinsically result in a prior volume effect. In Fig. 7.6 it is possible to appreciate the projection effect generated by this asymmetry of the prior. We explore the ITEM priors obtained in these situations in the following subsections.

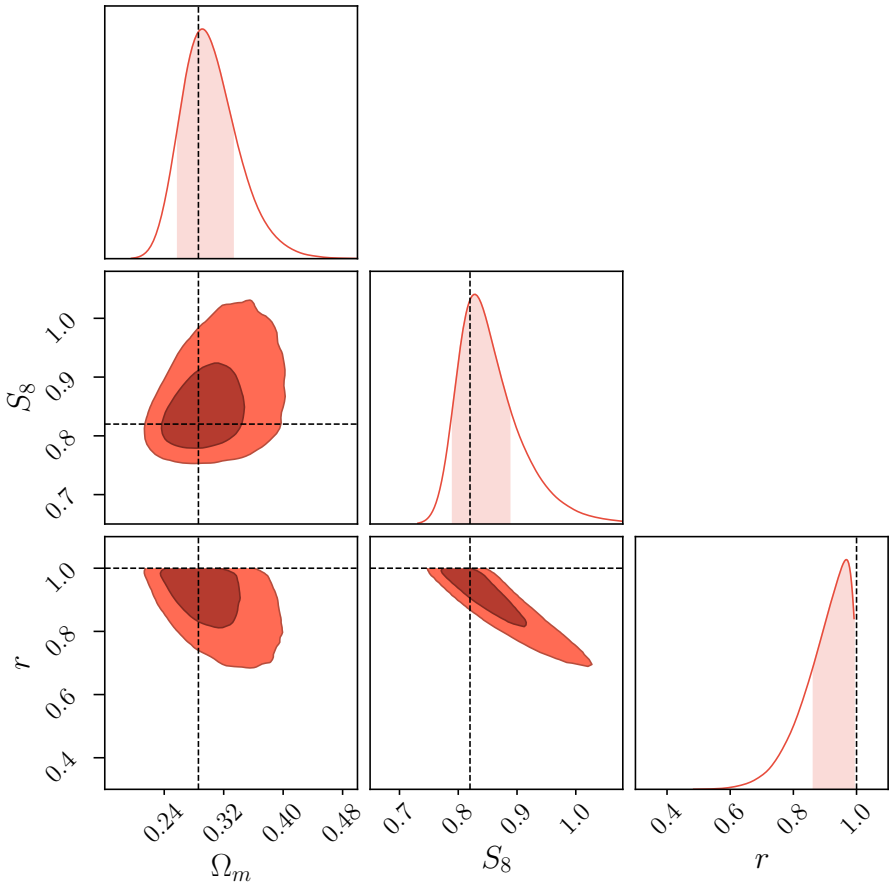


Figure 7.6: Marginalized (Ω_m, S_8, r) constraints posteriors from simulated DSS r *Model* using a non-stochastic parameter vector from the α *Model* marked with the dashed line. The peak on the 1D projection posterior on r indicates a bias that is explained on the sharp upper bound from the prior on r.

7.4.3. Obtaining the ITEM priors on the Density Splits Statistics

In this subsection, we detail the procedure we followed to obtain the ITEM priors for two nuisance models: first, a simple approach considering the α *Model* and a second, when the data is still simulated following the α *Model* but is analyzed assuming the r *Model*.

ITEM prior for the α *Model*

In DSS, the nuisance parameters α_0 and α_1 were assigned to uniform priors. We make the same choice here and only vary the limits of these flat priors. We note that one could also vary, e.g., the mean and width of a Gaussian prior.

We start by sampling the candidate priors as done in the example at Sec. 7.3.2. Then, we narrow the bounds of the flat priors of the stochasticity parameters due to an excess of prior volume with respect to the posteriors in the simulations for both nuisance realizations (i.e., the final posteriors on the stochasticity parameters were concentrated in a reduced fraction of the original prior distributions because they were not informative enough). Finally, we shrink the extremes ($\min[\alpha_0]$, $\max[\alpha_0]$, $\min[\alpha_1]$ and $\max[\alpha_1]$) with a step of 0.1 units per cut until they reach the closest rounded fiducial value of the data vector.

$$\min[\alpha_0] \in \{0.1, 0.2, ..., 0.9\}$$
$$\max[\alpha_0] \in \{1.4, 1.5, ..., 2.7\}$$

$$\min[\alpha_1] \in \{-1.0, -0.9, ..., -0.1\}$$
$$\max[\alpha_1] \in \{0.3, 0.4, ..., 3.0\}$$

This approach gives a total of 35,280 possible priors when combining each different bound. We then apply an importance sampling to the original posteriors and follow the three steps described in Sec. 7.3 to obtain the ITEM priors:

We first estimate, for each candidate prior, the maximum Mahalanobis distance x from Eq. 7.9 for the different posteriors. We decide to set the threshold to $0.3\sigma_{2D}$ for the first filter of the ITEM priors, following DES standards. There were, in total, 231 priors that had a bias of less than $0.3\sigma_{2D}$ ($\sim 0.65\,\%$ of the original set).

Secondly, we demand the total error confidence region (the one re-scaled by the bias in Sec. 7.3.3) to enclose the true parameter values on the 2D contour plot of Ω_m and S_8 at least 68 times when simulating 100 noisy independent realizations per nuisance configuration. These noisy realizations were simulated by adding multivariate Gaussian noise with the covariance matrix from [1] to the noiseless data vectors of our two nuisance scenarios.

To save computational power, instead of running a total of 200 MCMC chains, we opt to compute an importance sampling, assuming that a Gaussian sampling distribution would fit approximate the actual shape. This allows us to derive approximations of the desired contours and obtain the coverage per prior to sample specifically the 1-σ zone of interest.

Per noisy realization, we draw an elliptical cloud of points centered at the optimized MAP estimate of the noisy data vector and use a $5-\sigma$ extended covariance matrix of the original noiseless MCMC to determine the area. The distribution of the points is obtained using an inverse sampling of 2^{10} points from a 5-dimensional unit cube following a Sobol sequence [197] to be enough spread in the space of parameters. This is statistically more advantageous than using a Gaussian cloud of points centered in the same MAP because in the case the sampled distribution is highly non-Gaussian, the Sobol cloud will *support* with more certain these cases.

We calculate the corresponding importance weights representing the underlying distribution for each of these points and obtain the $1-\sigma$ extended regions (i.e., considering the bias) of confidence corresponding to all the proposed priors. We integrate the discrete weights until they reach the minimum probability. From these, a total of 111 out of the 231 candidate priors return a minimum total error coverage, for both nuisance realizations, over the extended 68 % confidence region.

To complement the previous result, we explore the frequentist probability distribution representation of the coverage and the expected Bayesian counterpart. If both families of probabilities match, the coverage X_{coverage} will follow a Binomial distribution $X_{\text{coverage}} \sim \text{Bin}(n, p)$, where n is the number of noisy realizations and $p = 0.68$. The uncertainty would be given by the inherently noisy nature of the considered noisy realizations.

We explore other minimum values for the coverage threshold keeping inside the a confidence zone of $\text{Bin}(100, 0.68)$. The mean and standard deviation would be $\mu = np = 68$ and $\sigma = \sqrt{np(1-p)} \approx 4.665$ respectively. From the cumulative distribution function of X_{coverage}, we will have that: $P(X_{\text{coverage}} \geq \mu - 2\sigma) \approx 0.986$, which is consistent with the one obtained from the data, illustrated in Table 7.2.

Table 7.2: Difference between the expected and the resulting fraction of priors that had a minimum required coverage x for a extended confidence region of 68 % for the α model.

x	$P(X_{\text{coverage}} \geq x)$	Final result	Difference
μ	0.547	0.484	0.063
$\mu - \sigma$	0.879	0.852	0.027
$\mu - 2\sigma$	0.986	0.939	0.047

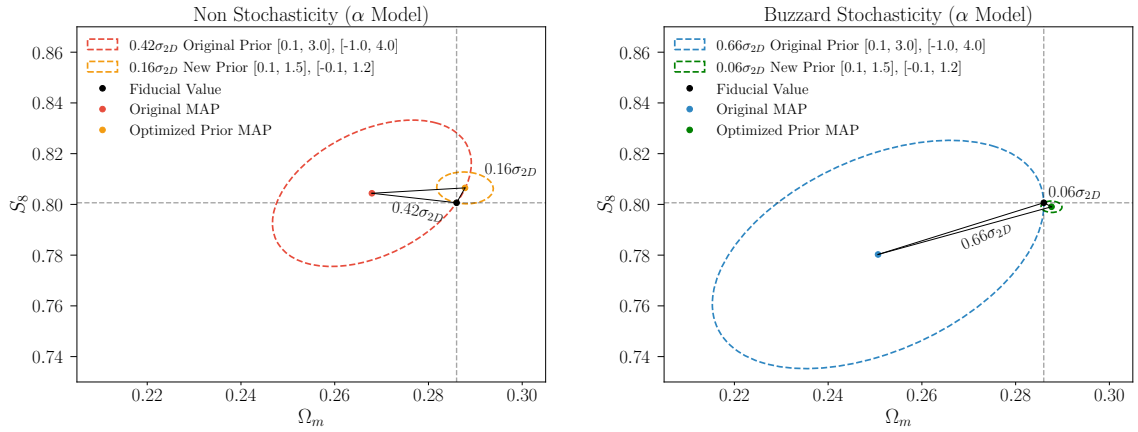
The three different possible minimum thresholds are consistent with the expected values if there were a match between both probabilities, meaning that the priors are congruous with the corresponding levels. We determine to set the threshold to $68 - 2\sigma$, filtering the number of candidate priors from 231 to 228. We also checked the ITEM priors derived with other thresholds (68 and $68-\sigma$), and we did not find substantial differences. We then proceed with the final step: the minimization of the error from Eq. 7.14 to obtain the corresponding ITEM prior.

The resulting ITEM priors for α_0 and α_1 are listed in the second row of Table 7.3. We first

note that the prior volumes of both parameters were reduced considerably. In particular, α_1 was narrowed to a final width of 24 % times its original value. The new configurations return substantially less biased estimated and competitive coverages for both nuisance realizations. They both had a bias of $0.16\sigma_{2D}$ and $0.06\sigma_{2D}$ for the target parameters as shown in Fig. 7.7, respectively.

Table 7.3: The original and ITEM priors ranges for the α *Model* for the case in which there is a $0.3\sigma_{2D}$ maximum required bias and the minimum total error coverage is $68\,\% - 2\sigma$. The bias and error coverage quantities from both nuisance realizations are illustrated for each prior as *non-stochasticity/buzzard stochasticity* cases respectively. In the final three columns, we included the average error of the final estimate for Ω_m and S_8 from both nuisance realizations and the mean volume of the confidence ellipse extended with the systematic bias from Eq. 7.14.

Parameter	Prior	Prior Width	Bias $[\sigma_{2D}]$	Extended Coverage	$\overline{\sigma}_{\Omega_m}$	$\overline{\sigma}_{S_8}$	$\overline{V}$
Original Priors							
α_0	$\mathcal{U}$[0.1, 3.0]	2.9	0.42 / 0.66	63 %/72 %	0.052	0.068	0.027
α_1	$\mathcal{U}$[-1.0, 4.0]	5.0					
ITEM Prior							
α_0	$\mathcal{U}$[0.3, 1.5]	1.2	0.16 / 0.06	62 %/68 %	0.036	0.037	0.009
α_1	$\mathcal{U}$[-0.1, 1.1]	1.2					



(a) Simulated likelihood analysis for the data vector without shot-noise ($\alpha_0 = 1.0$, $\alpha_1 = 0.0$).

(b) Simulated likelihood analysis for the data vector with buzzard stochasticity ($\alpha_0 = 1.26$, $\alpha_1 = 0.29$).

Figure 7.7: Parameter biases of the data vectors with original and ITEM priors: the red and blue ellipses show wide contours, while the yellow and green ones the $0.3\sigma_{2D}$ required threshold for the 2D marginalized constraints. Both are centered in their respective 2D projected MAP. Due to parameter volume effects, the marginalized constraints from the baseline prior analysis are not centered on the input cosmology. The dashed horizontal and vertical lines indicate the fiducial parameter values.

In the last columns of Table 7.3, we appreciate the average standard deviations and the total error of the target parameters being reduced. This can also be interpreted from the narrowed

marginalized distributions shown in Fig. 7.8. In particular, we find that the marginalized 1D distributions of Ω_m and S_8 obtained with the ITEM prior are less correlated than the original ones. The general constraint works for both 1D and 2D marginalized cases. These results fall into what we would expect from an ITEM prior because it is built to accomplish the requirements. Later in this section, in Table 7.5 and in Fig. 7.10, we present the impact of the ITEM priors on the cosmological constraint from DES Y1 real data.

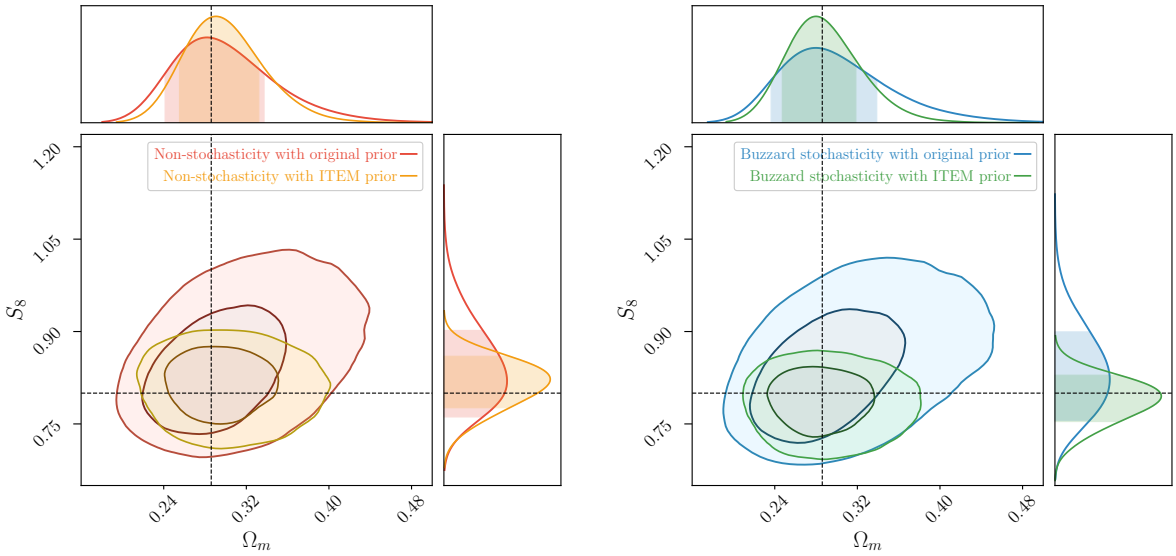


Figure 7.8: The 1D and 2D marginalized posterior distributions for the simulated likelihood chains from the α *Model* from DSS with the ITEM prior. In the top and bottom panel we illustrate the non-stochasticity ($\alpha_0 = 1.0$, $\alpha_1 = 0.0$) and the buzzard stochasticity ($\alpha_0 = 1.26$, $\alpha_1 = 0.29$) realizations of the nuisance respectively. The larger contours correspond to the posterior simulated using the original DSS α Model priors over the stochasticity parameters α_0 and α_1. In contrast, the narrow contours illustrate the performance of the ITEM prior methodology for a threshold of $0.3\sigma_{2D}$ and a minimum total error coverage of 68 %.

ITEM prior for r *Model* with a different basis model

In this subsection, we investigate how the ITEM priors can help us in cases where our assumed nuisance models are not the generators of the data. We used the same two-parameter vectors (the non-stochasticity and the Buzzard stochasticity) from the previous analysis, listed in Table 7.1. But instead, we now use the r *Model* to obtain posteriors. The parameters considered were Ω_m, σ_8, b and r. We use the same prior distribution over r as in the DSS papers: $\Pi(r) \sim \mathcal{U}[0.0, 1.0]$. We then obtained two posteriors (the non-stochasticity and the Buzzard realization of the nuisance).

As before, we sampled different priors for r, but in this case, we only change the lower bound $\min[r]$ and keep $\max[r]$ fixed to 1.0:

$$\min[r] \in \{0.000, 0.005, ..., 0.990\}$$

Table 7.4: The original and best fit prior ranges for r *Model* for the case in which there is a $0.5\sigma_{2D}$ maximum required bias and the minimum total error coverage is $68\,\% - 2\sigma$. As in Table 7.3, the bias and coverage quantities from both nuisance realizations are separated with a backslash.

Parameter	Prior	Prior Width	Bias[σ_{2D}]	Extended Coverage	$\overline{\sigma}_{\Omega_m}$	$\overline{\sigma}_{S_8}$	$\overline{V}$
Original Prior							
r	$\mathcal{U}$[0.00, 1.00]	1.00	0.33 / 0.55	62 %/63 %	0.038	0.059	0.017
ITEM Prior							
r	$\mathcal{U}$[0.75, 1.00]	0.25	0.47 / 0.46	61 %/60 %	0.035	0.045	0.013

We repeat the three steps of the ITEM prior pipeline:

We first reduce the number of candidate priors from 199 to 45 by setting a threshold equivalent of $0.5\sigma_{2D}$ ($\sim 22.6\,\%$ of the original set). The threshold is larger than the one used in the α *Model* analysis because the sharp upper bound of the prior on r leads to systematic biases with any prior (see Fig. 7.6). We also ensure that the total error coverage passes the minimum threshold with this higher value.

We secondly calculate the total error coverage of the noisy realizations required over Ω_m-S_8 to be more than $68\,\% - 2\sigma$ for the cosmological target parameters when simulating 200 noisy realizations as done for the α model. This second filter reduces the number of candidate priors from 45 to 39 ($\sim 87\,\%$ of selection). The original prior had a minimum coverage of 62 for a $68\,\%$ confidence region.

We finally determine the ITEM prior by optimizing the volume $\overline{V}$ from Eq. 7.14. In Table 7.4 we summarize the resulting prior, with the corresponding biases, coverage, and mean errors. We observe an equilibrium of the biases of the final ITEM prior relative to the original one. This means that the ITEM prior can return similar biases for both nuisance realizations. The reduction of the prior width ensures a decrease in the total uncertainty over the target cosmological parameters in the last three columns. We can also visualize this in the projected posteriors of Ω_m and S_8 from Fig. 7.9. There is an uncertainty reduction in the 2D marginalized posterior, specifically a decrease of the higher values of Ω_m and S_8.

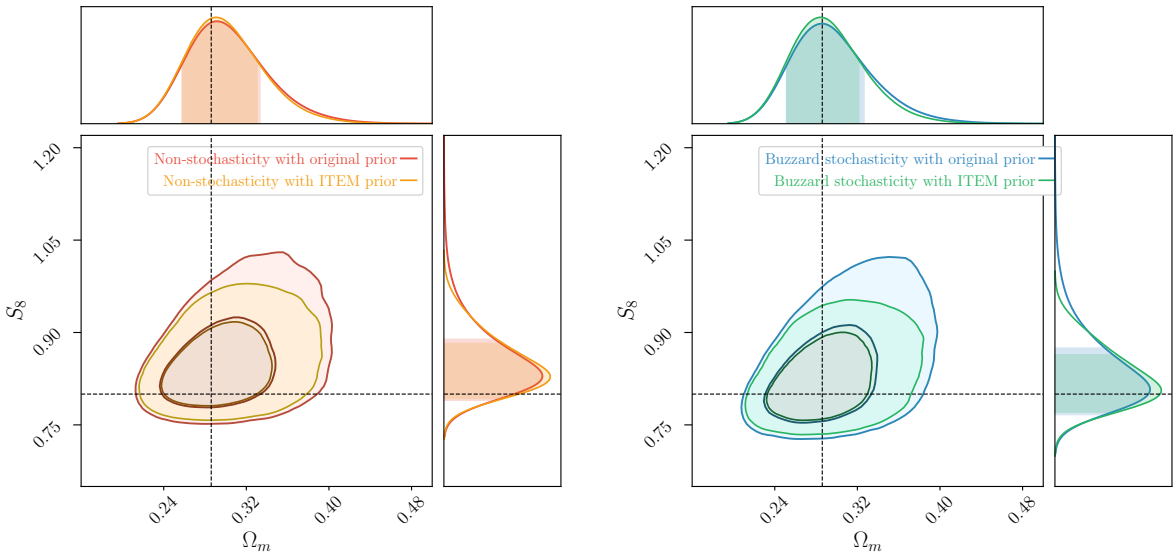


Figure 7.9: The resulting 1D and 2D marginalized posterior distributions of the two configuration data vectors from α *Model*, but analyzed using r *Model* as reference.

To further illustrate this method's scopes, we use the obtained ITEM priors (on both α and r models) on real DESY1 posterior data.

In Fig. 7.10 and Table 7.5 we reproduce Fig. 10 and Table II from [1] incorporating the constraints obtained in this work with the ITEM priors.

The constraints on the matter density did not only converge in almost an identical 1D posterior distribution ($\Omega_m = 0.25^{+0.03}_{-0.02}$), but also led to a more consistent result with the case in which there is only a linear bias (top left panel from Fig. 7.10). The constraints on the amplitude of late-time structure formation kept the observed inconsistency ($\sigma_8 = 0.94^{+0.06}_{-0.04}$ and $\sigma_8 = 0.84^{+0.04}_{-0.05}$ for the r and α models respectively) due to a reduction of the difference on the central value estimate, but also a major concentration.

These stronger priors reduce the cosmological uncertainty. In addition, the ITEM constraints of the yellow and green contours in Fig. 7.10 are closer to the black region. However, it is not feasible yet to discriminate between bias models with these results because their behavior is symmetric w.r.t. the simple linear bias. That arises because we used the same data vectors to derive the ITEM priors. Again, external information, e.g., additional nuisance realizations included in the ITEM framework, will allow us to assess between shot-noise parametrizations.

The changes in the stochasticity parameters also deserve space for interpretation. For the r model we find that for a prior that constraints the likelihood analysis to $r \in \mathcal{U}[0.75, 1.00]$, the mean estimate does not change. There is only an rise of the lower bound ($r = 0.77^{+0.10}_{-0.13}$ for the original analysis to $r = 0.77^{+0.10}_{-0.02}$ for the ITEM prior), which implies that the deviation of r from unity is still at the 2σ level.

For the α model we find a determinant preference toward lower stochasticity values after applying the ITEM priors. For α_0 we considered two nuisance configurations ($[1.00, 1.26]$) which lead to an ITEM prior $\alpha_0 \in \mathcal{U}[0.3, 1.5]$ that decrease the excess of volume for high-

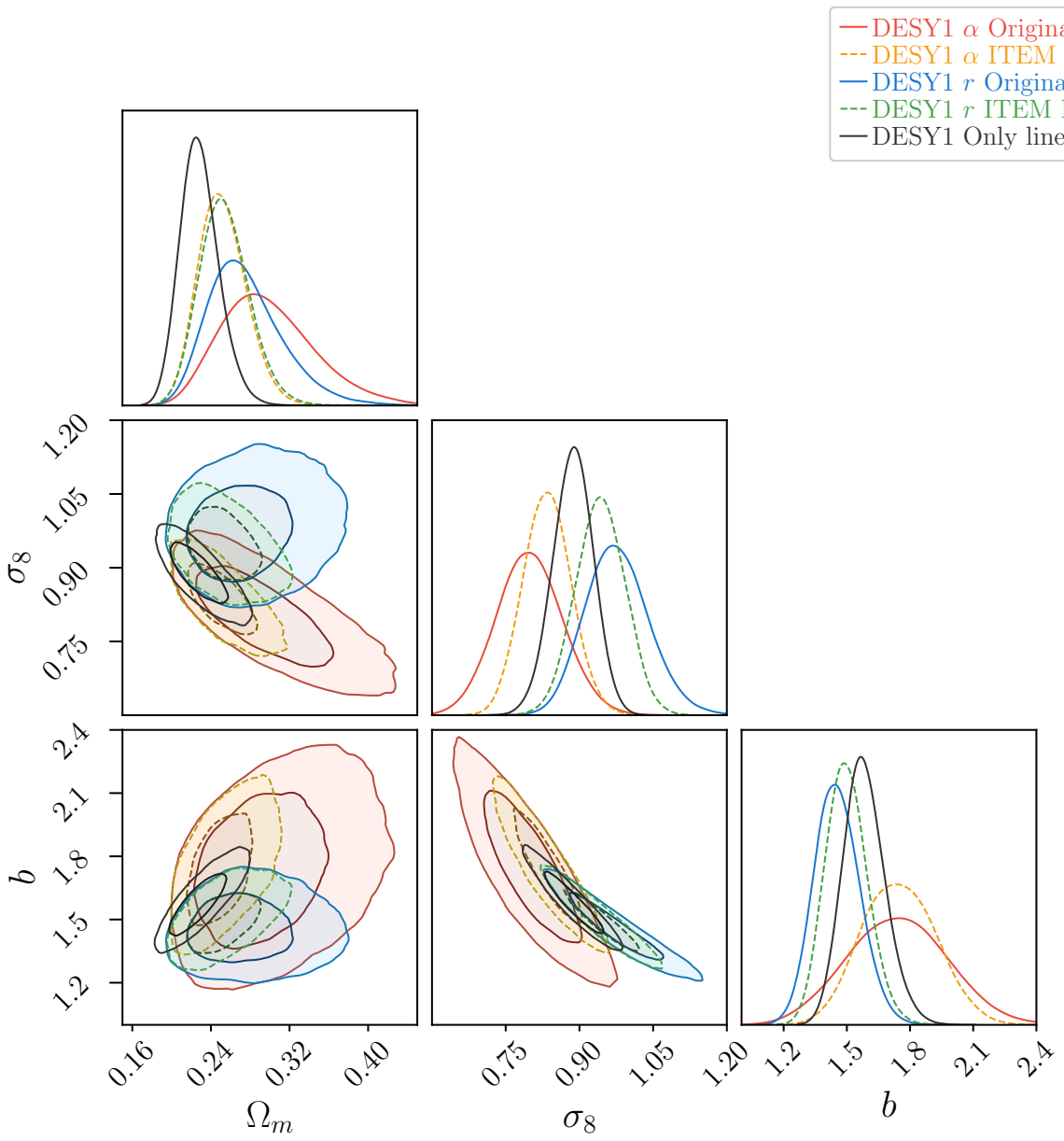


Figure 7.10: The 1D and 2D forecasr constraints for the final cosmological chains obtained from the DSS analysis of DESY1 data using the linearly biased tracers with Poissonian noise in black, the r model in blue and α models with red color. The smaller contours illustrate how the cosmological inference would change if we apply an importance sampling using the ITEM priors for the stochasticity parameters. With yellow, we show the constraint of the using the ITEM prior with $0.3\sigma_{2D}$ threshold over the α *Model*, while with green, we show the constraint using the ITEM prior with $0.5\sigma_{2D}$ threshold over the r *Model*. The original constraints are plotted in Fig. 10 from [1].

stochasticity values ($\alpha_0 = 1.5^{+0.4}_{-0.6}$ for the original analysis to $\alpha_0 = 0.9^{+0.4}_{-0.4}$ for the ITEM prior). The ITEM prior $\alpha_1 \in \mathcal{U}[-0.1, 1.1]$ obtained from the original nuisance realizations

([0.00, 0.29]) decrease the excess of volume for matter-dependent stochasticity ($\alpha_1 = 1.7^{+1.1}_{-0.9}$ for the original analysis to $\alpha_1 = 1.1^{+0.1}_{-0.5}$ for the ITEM prior). For α_0 the estimate is agrees with the case of non-stochasticity ($\alpha_0 = 1.0$ and $\alpha_1 = 0.0$), while for α_1, there is a deviation at the 2σ level, which is consistent with a super-Poissonian scatter in galaxy count at fixed matter density and with the results obtained for the r model.

Table 7.5: DESY1 likelihood 1D constraints with the original and ITEM priors respectively for both shot-noise models.

Model	S_8	Ω_m	σ_8	b	r	α_0	α_1
r Original	$0.90^{+0.11}_{-0.08}$	$0.26^{+0.04}_{-0.03}$	$0.97^{+0.07}_{-0.06}$	$1.45^{+0.10}_{-0.11}$	$0.77^{+0.10}_{-0.13}$	–	–
r ITEM Prior	$0.88^{+0.04}_{-0.06}$	$0.25^{+0.03}_{-0.02}$	$0.94^{+0.06}_{-0.04}$	$1.49^{+0.10}_{-0.09}$	$0.77^{+0.10}_{-0.02}$	–	–
α Original	$0.78^{+0.05}_{-0.04}$	$0.28^{+0.05}_{-0.04}$	$0.80^{+0.06}_{-0.07}$	$1.75^{+0.22}_{-0.26}$	–	$1.5^{+0.4}_{-0.6}$	$1.7^{+1.1}_{-0.9}$
α ITEM Prior	$0.76^{+0.03}_{-0.03}$	$0.25^{+0.03}_{-0.02}$	$0.84^{+0.04}_{-0.05}$	$1.75^{+0.17}_{-0.19}$	–	$0.9^{+0.4}_{-0.4}$	$1.1^{+0.1}_{-0.5}$

7.5. Discussions and conclusion

This work presents the ITEM prior pipeline, and its first application in the shot-noise models from the density split statistics. We show that it is possible to overcome projection effects using a series of simulated likelihood analyses.

It is common to set a flat prior to parameters without exploring if there are projection effects due to non-linear relationships between the variables. This situation is even more complicated when higher dimensions are considered because connections can be more complex and less evident. Bayesian analysis allows us to perform simulation likelihood tests to propose a prior that, when facing the real data analysis, has some prior motivation.

With the increment of cosmological data and, therefore, systematic effects, we propose a method to make statistically significant statements about the parameters of interest when considering different nuisance realizations. The pipeline sets an equilibrium between the underfitting and overestimation of the nuisance parameters in a model. This is done by requiring that the biases on the target parameters are lower than a certain fraction of their errors and, in the presence of noise, we still cover the original fiducial parameter values a certain amount of times in some desired confidence region.

We test the ITEM prior pipeline using two different nuisance models from the DSS framework: the extended Poissonian shot-noise model with galaxy stochasticity and a log-normal model for the joint PDF with a correlation term. We simulate two realistic DESY1 data vectors: one accounts for a case with no stochasticity between galaxy counts and the underlying matter density contrast. Another one considers specific stochasticity motivated by previous studies.

For these realizations, we were able to reduce the biases and the cosmological uncertainties by choosing from a list of candidate priors the one that optimized both random and systematic errors under the ITEM prior pipeline. In a future work (Britt et al. in prep.), we will consider other configurations of the shot-noise motivated by different halo occupation distribution (HOD) simulations and study the impact on more realistic scenarios.

We also studied the case in which the basic model is not the same as the one used to generate the data. In this case, the nuisance models are inherently biased due to a non-linear mapping between the models and a non-equivalent prior space. This second case returns more significant discrepancies between the actual parameters and the estimated quantities regarding their errors. This disparity required us to be less demanding when setting the threshold to find the ITEM prior.

One of the advantages of the ITEM priors is that they are numerically adjustable. This flexibility allows us to tune thresholds to values we consider more consistent and propose different nuisance realizations and prior configurations. We aimed to explore a different approach to assigning prior distributions to nuisance parameters in the absence of a physical or mathematical motivation to constrain them. We determined priors on two models using our proposal.

It is important to recall that we present a pipeline whose results depend on the fiducial values we assign to the parameter of interests. The use of ITEM priors with real data should also consider different realizations of those values. However, in cosmology, a few different configurations should be enough in the presence of smooth fields, like the ones we find in the space of the cosmological parameters.

There are still some queries we should address in the future research, which include:

- How accurate is our Gaussian assumption on the Sobol sampling used to determine the $1-\sigma$ confidence regions in the noisy realizations: the most precise constraints on the contours can be obtained when performing the actual MCMC chains. By the time of writing this thesis, alternative methods to save computational costs have been proposed since the beginning of this work. We hope, e.g., to derive faster results using the alternative method presented by [198] and publish them in a final article shortly (Ried et al. in prep.).
- How simulations on the halo occupation distribution can provide complementary nuisance realizations of the shot-noise according to a wider variety of plausible stochasticities that could be present in the universe. This is currently being explored in a companion article (Britt et al. in prep.).
- How do the considered fiducial values for the target parameters impact the determination of the ITEM prior. This supplement would imply an extra layer of complexity of the ITEM prior derivation that can be systematically added to the presented framework.

Given the extensive use of Bayesian statistical inference, we expect the ITEM priors to help overcome the prior volume effect when there is a required nuisance modeling, but not limiting itself only to the area of observational cosmology.

Acknowledgements

We would like to thank Elisabeth Krause, Steffen Hagstotz and Raffaella Capasso for helpful discussions.

BR was funded by the Chilean National Agency for Research and Development (ANID) - Subdirección de Capital Humano / Magíster Nacional / 2021 - ID 22210491 and the German Academic Exchange Service (DAAD, Short-Term Research Grant 2021 No. 57552337). BR also gratefully acknowledge support from the Chilean Centro de Astrofísica (ANID Project AFB 170002) and the Program for Astrophysics Visitor Exchange at Stanford (PAVES).

BR, OF and DG were funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC-2094 – 390783311.

Data Availability

The Dark Energy Survey Year 1 data (Density Split Statistics) used in this work is publicly available at https://des.ncsa.illinois.edu/releases/y1a1/density.

Chapter 8

Conclusions

Through this work, we extensively examined the cosmological and statistical background to address one of the difficulties that the era of precision cosmology is now facing. Fig. 7.1 can be understood as the starting point to our motivation: *how do we include nuisance parameters in our cosmological models without falling into assigning non-informative priors that, consequently, lead to projection effects?*

We started answering this question concerning why we need nuisance parameters in cosmology in the first place. Once the basics of cosmology are presented in Chapter 2, we focus on the cosmic large-scale structure formation and its observation in Chapter 3. In the latter, we dig through the general theory of cosmic structure until we express the necessity of including nuisance parameters to account for higher-order statistics. In particular, we focused on the galaxy bias and possible extensions that account for astrophysical effects of smaller scales. We concentrate on the two shot-noise models from the DSS framework, which had the remaining task of exploring informative priors on their galaxy tracer quantities.

We likewise decided to answer the question from a statistical perspective, addressing if it is possible to define a framework that systematically determines prior distributions to the nuisance parameters. We asked ourselves which would be the properties we expect the marginalized posteriors of the cosmological parameters to have. These include *unbiasedness, reliability, robustness* and *precision*. Consequently, it was necessary to review the basics of probabilities and statistical inference as done in Chapters 4 and 5, respectively. We thoroughly examined the theory of point and interval estimators, including the different classes and existing approaches. We also discussed their selection and which possible errors we can encounter when performing a likelihood analysis.

Once we specified the desired properties, we explored how to condense those statistical desires into a feasible pipeline that would help us to overcome the prior volume effect. One of the main challenges was concentrating our Bayesian and frequentist thoughts on a series of checkpoints for the nuisance priors. As the prior volume effect will lead us to shifted marginalized distributions, the coverage fractions will not be consistent with the associated confidence level. Therefore, our requirements included reducing the biases and demanding consistent coverages w.r.t. some fiducial configuration. This strategy could allow us to trust our Bayesian analyses further because we will recover the input cosmology with some statistical certainty in, e.g., an ensemble of universes.

The statistical training acquired with these discussions yielded the ITEM priors methodology presented in Chapter 7. We elaborated a systematic framework to determine nuisance prior distributions under an alternative method to concentrate prior information of realizations of the nuisance. As mentioned, we are still studying if our alternative way of deriving confidence contours is robust with respect to the MCMC sampling standard procedure. Despite that technicality, we forecasted some possible scenarios with ITEM priors that decrease the uncertainties on the cosmological parameters. We expect to publish the final results shortly.

Finally, we invite the scientific community to explore this method and study its implications in other fields. We hope this derivation enhances statistical inferences and decreases systematic errors in critical tensions like the ones we currently face in many areas of physics.

Chapter 9

Glossary

- **Astronomic distances**:

  An arcminute (') is a unit of angular measurement equal to 1/60 of one degree. An arcsecond (") is 1/60 of an arcminute.

  A astronomical unit (AU) is the mean distance between the Sun and the Earth. It is equivalent to $\sim 1.49 \cdot 10^8$ km.

  A light-year (ly) represents the distance light will travel in a straight line for a year. It is equivalent to $\sim 9.46 \cdot 10^{12}$ km.

  A parsec (pc) is the distance in which one astronomical unit subtends an angle of one arcsecond. It is equivalent to 3.26 light-years.

- **Fourier formalism**:

  Let $f(\boldsymbol{x})$ be a function such that $\boldsymbol{x} \in \mathbb{R}^3$. The Fourier transform of this function will be:

  $$\tilde{f}(\boldsymbol{k}) = \int f(\boldsymbol{x}) e^{-i\boldsymbol{x}\cdot\boldsymbol{k}} d^3x,$$

  with inverse transform:

  $$f(\boldsymbol{x}) = \frac{1}{(2\pi)^3} \int \tilde{f}(\boldsymbol{k}) e^{i\boldsymbol{x}\cdot\boldsymbol{k}} d^3k.$$

- **General Relativity notation**:
  - Symbols like x_μ represent a contra-variant 4-vector.
  - Greek letters indices: 0, 1, 2, 3
  - Latin indices: 1, 2, 3
  - Repeated indices implies sum (Einstein's sum convention).

  - The kronecker symbol δ_α^β is the identity matrix.

- **Constants**:
  - Velocity of light: c
  - Newtonian constant of gravitation: G

- **Natural unit system**:

  Throughout this work we the natural unit system:

  $$c = G = \hbar = k_B = 1$$

  unless stated otherwise.

- **Dirac delta function**:

  The Dirac delta function is defined as zero everywhere except at the some point x', where it is infinite:

  $$\delta_D(x - x') = \begin{cases} +\infty, & x = x' \\ 0, & x \neq x' \end{cases}.$$

  One of the main identities is that:

  $$\int_{-\infty}^{\infty} \delta(x - x')dx = 1.$$

- The expectation value $\langle \cdot \rangle$ will be an ensemble average over all possible states of a system.
- The natural logarithm will be given by the $\log(\cdot)$ symbol.

Bibliography

[1] Gruen, D., Friedrich, O., Krause, E., DeRose, J., Cawthon, R., Davis, C., Elvin-Poole, J., Rykoff, E. S., Wechsler, R. H., Alarcon, A., Bernstein, G. M., Blazek, J., Chang, C., Clampitt, J., Crocce, M., De Vicente, J., Gatti, M., Gill, M. S. S., Hartley, W. G., Hilbert, S., Hoyle, B., Jain, B., Jarvis, M., Lahav, O., MacCrann, N., McClintock, T., Prat, J., Rollins, R. P., Ross, A. J., Rozo, E., Samuroff, S., Sánchez, C., Sheldon, E., Troxel, M. A., Zuntz, J., Abbott, T. M. C., Abdalla, F. B., Allam, S., Annis, J., Bechtol, K., Benoit-Lévy, A., Bertin, E., Bridle, S. L., Brooks, D., Buckley-Geer, E., Carnero Rosell, A., Carrasco Kind, M., Carretero, J., Cunha, C. E., D'Andrea, C. B., da Costa, L. N., Desai, S., Diehl, H. T., Dietrich, J. P., Doel, P., Drlica-Wagner, A., Fernandez, E., Flaugher, B., Fosalba, P., Frieman, J., García-Bellido, J., Gaztanaga, E., Giannantonio, T., Gruendl, R. A., Gschwend, J., Gutierrez, G., Honscheid, K., James, D. J., Jeltema, T., Kuehn, K., Kuropatkin, N., Lima, M., March, M., Marshall, J. L., Martini, P., Melchior, P., Menanteau, F., Miquel, R., Mohr, J. J., Plazas, A. A., Roodman, A., Sanchez, E., Scarpine, V., Schubnell, M., Sevilla-Noarbe, I., Smith, M., Smith, R. C., Soares-Santos, M., Sobreira, F., Swanson, M. E. C., Tarle, G., Thomas, D., Vikram, V., Walker, A. R., Weller, J., Zhang, Y., and DES Collaboration, "Density split statistics: Cosmological constraints from counts and lensing in cells in DES Y1 and SDSS data," Physical Review D, vol. 98, no. 2, p. 023507, 2018, doi:10.1103/PhysRevD.98.023507.

[2] Friedrich, O., Gruen, D., DeRose, J., Kirk, D., Krause, E., McClintock, T., Rykoff, E. S., Seitz, S., Wechsler, R. H., Bernstein, G. M., Blazek, J., Chang, C., Hilbert, S., Jain, B., Kovacs, A., Lahav, O., Abdalla, F. B., Allam, S., Annis, J., Bechtol, K., Benoit-Lévy, A., Bertin, E., Brooks, D., Carnero Rosell, A., Carrasco Kind, M., Carretero, J., Cunha, C. E., D'Andrea, C. B., da Costa, L. N., Davis, C., Desai, S., Diehl, H. T., Dietrich, J. P., Drlica-Wagner, A., Eifler, T. F., Fosalba, P., Frieman, J., García-Bellido, J., Gaztanaga, E., Gerdes, D. W., Giannantonio, T., Gruendl, R. A., Gschwend, J., Gutierrez, G., Honscheid, K., James, D. J., Jarvis, M., Kuehn, K., Kuropatkin, N., Lima, M., March, M., Marshall, J. L., Melchior, P., Menanteau, F., Miquel, R., Mohr, J. J., Nord, B., Plazas, A. A., Sanchez, E., Scarpine, V., Schindler, R., Schubnell, M., Sevilla-Noarbe, I., Sheldon, E., Smith, M., Soares-Santos, M., Sobreira, F., Suchyta, E., Swanson, M. E. C., Tarle, G., Thomas, D., Troxel, M. A., Vikram, V., Weller, J., and DES Collaboration, "Density split statistics: Joint model of counts and lensing in cells," prd, vol. 98, 2018, doi:10.1103/PhysRevD.98.023508.

[3] Penzias, A. A. and Wilson, R. W., "A Measurement of Excess Antenna Temperature at 4080 Mc/s.," Apj, vol. 142, pp. 419–421, 1965, doi:10.1086/148307.

[4] Dicke, R. H., Peebles, P. J. E., Roll, P. G., and Wilkinson, D. T., "Cosmic Black-Body Radiation.," Apj, vol. 142, pp. 414–419, 1965, doi:10.1086/148306.

[5] Lemaître, A. G., "A homogeneous universe of constant mass and increasing radius accounting for the radial velocity of extra-galactic nebulae," Monthly Notices of the Royal Astronomical Society, vol. 91, no. 5, pp. 483–490, 1931, doi:10.1093/mnras/91.5.483.

[6] Hubble, E., "A relation between distance and radial velocity among extra-galactic nebulae," Proceedings of the National Academy of Sciences, vol. 15, no. 3, pp. 168–173, 1929, doi:10.1073/pnas.15.3.168.

[7] Weinberg, D. H., Mortonson, M. J., Eisenstein, D. J., Hirata, C., Riess, A. G., and Rozo, E., "Observational probes of cosmic acceleration," Physics Reports, vol. 530, no. 2, pp. 87–255, 2013, doi:10.1016/j.physrep.2013.05.001.

[8] Riess, A. G., Filippenko, A. V., Challis, P., Clocchiatti, A., Diercks, A., Garnavich, P. M., Gilliland, R. L., Hogan, C. J., Jha, S., Kirshner, R. P., Leibundgut, B., Phillips, M. M., Reiss, D., Schmidt, B. P., Schommer, R. A., Smith, R. C., Spyromilio, J., Stubbs, C., Suntzeff, N. B., and Tonry, J., "Observational evidence from supernovae for an accelerating universe and a cosmological constant," The Astronomical Journal, vol. 116, pp. 1009–1038, 1998, doi:10.1086/300499.

[9] Perlmutter, S., Aldering, G., Goldhaber, G., Knop, R. A., Nugent, P., Castro, P. G., Deustua, S., Fabbro, S., Goobar, A., Groom, D. E., Hook, I. M., Kim, A. G., Kim, M. Y., Lee, J. C., Nunes, N. J., Pain, R., Pennypacker, C. R., Quimby, R., Lidman, C., Ellis, R. S., Irwin, M., McMahon, R. G., Ruiz-Lapuente, P., Walton, N., Schaefer, B., Boyle, B. J., Filippenko, A. V., Matheson, T., Fruchter, A. S., Panagia, N., Newberg, H. J. M., Couch, W. J., and Project, T. S. C., "Measurements of ω and λ from 42 high-redshift supernovae," The Astrophysical Journal, vol. 517, pp. 565–586, 1999, doi:10.1086/307221.

[10] Dodelson, S., Modern Cosmology. Amsterdam: Academic Press, 2003.

[11] Moresco, M., Amati, L., Amendola, L., Birrer, S., Blakeslee, J. P., Cantiello, M., Cimatti, A., Darling, J., Della Valle, M., Fishbach, M., Grillo, C., Hamaus, N., Holz, D., Izzo, L., Jimenez, R., Lusso, E., Meneghetti, M., Piedipalumbo, E., Pisani, A., Pourtsidou, A., Pozzetti, L., Quartin, M., Risaliti, G., Rosati, P., and Verde, L., "Unveiling the universe with emerging cosmological probes," 2022, doi:10.48550/ARXIV.2201.07241.

[12] Peebles, P., Principles of Physical Cosmology. Princeton Series in Physics, Princeton University Press, 2020, https://books.google.cl/books?id=H2lczQEACAAJ.

[13] Djorgovski, S. G., Mahabal, A., Drake, A., Graham, M., and Donalek, C., "Sky surveys," en Planets, Stars and Stellar Systems, pp. 223–281, Springer Netherlands, 2013, doi:10.1007/978-94-007-5618-2_5.

[14] Valentino, E. D., Mena, O., Pan, S., Visinelli, L., Yang, W., Melchiorri, A., Mota, D. F., Riess, A. G., and Silk, J., "In the realm of the hubble tension, a review of solutions," Classical and Quantum Gravity, vol. 38, p. 153001, 2021, doi:10.1088/1361-6382/ac086d.

[15] Abbott, T. M. C., Aguena, M., Alarcon, A., Allam, S., Alves, O., Amon, A., Andrade-Oliveira, F., Annis, J., Avila, S., Bacon, D., Baxter, E., Bechtol, K., Becker, M. R., Bernstein, G. M., Bhargava, S., Birrer, S., Blazek, J., Brandao-Souza, A., Bridle, S. L., Brooks, D., Buckley-Geer, E., Burke, D. L., Camacho, H., Campos, A., Rosell, A. C., Kind, M. C., Carretero, J., Castander, F. J., Cawthon, R., Chang, C., Chen, A., Chen,

R., Choi, A., Conselice, C., Cordero, J., Costanzi, M., Crocce, M., da Costa, L. N., da Silva Pereira, M. E., Davis, C., Davis, T. M., Vicente, J. D., DeRose, J., Desai, S., Valentino, E. D., Diehl, H. T., Dietrich, J. P., Dodelson, S., Doel, P., Doux, C., Drlica-Wagner, A., Eckert, K., Eifler, T. F., Elsner, F., Elvin-Poole, J., Everett, S., Evrard, A. E., Fang, X., Farahi, A., Fernandez, E., Ferrero, I., Ferté, A., Fosalba, P., Friedrich, O., Frieman, J., García-Bellido, J., Gatti, M., Gaztanaga, E., Gerdes, D. W., Giannantonio, T., Giannini, G., Gruen, D., Gruendl, R. A., Gschwend, J., Gutierrez, G., Harrison, I., Hartley, W. G., Herner, K., Hinton, S. R., Hollowood, D. L., Honscheid, K., Hoyle, B., Huff, E. M., Huterer, D., Jain, B., James, D. J., Jarvis, M., Jeffrey, N., Jeltema, T., Kovacs, A., Krause, E., Kron, R., Kuehn, K., Kuropatkin, N., Lahav, O., Leget, P. F., Lemos, P., Liddle, A. R., Lidman, C., Lima, M., Lin, H., MacCrann, N., Maia, M. A. G., Marshall, J. L., Martini, P., McCullough, J., Melchior, P., Mena-Fernández, J., Menanteau, F., Miquel, R., Mohr, J. J., Morgan, R., Muir, J., Myles, J., Nadathur, S., Navarro-Alsina, A., Nichol, R. C., Ogando, R. L. C., Omori, Y., Palmese, A., Pandey, S., Park, Y., Paz-Chinchón, F., Petravick, D., Pieres, A., Malagón, A. A. P., Porredon, A., Prat, J., Raveri, M., Rodriguez-Monroy, M., Rollins, R. P., Romer, A. K., Roodman, A., Rosenfeld, R., Ross, A. J., Rykoff, E. S., Samuroff, S., Sánchez, C., Sanchez, E., Sanchez, J., Cid, D. S., Scarpine, V., Schubnell, M., Scolnic, D., Secco, L. F., Serrano, S., Sevilla-Noarbe, I., Sheldon, E., Shin, T., Smith, M., Soares-Santos, M., Suchyta, E., Swanson, M. E. C., Tabbutt, M., Tarle, G., Thomas, D., To, C., Troja, A., Troxel, M. A., Tucker, D. L., Tutusaus, I., Varga, T. N., Walker, A. R., Weaverdyck, N., Weller, J., Yanny, B., Yin, B., Zhang, Y., Zuntz, J., and DES Collaboration, "Dark Energy Survey Year 3 results: Cosmological constraints from galaxy clustering and weak lensing," 2021.

[16] Codis, S., "Nuisance parameters for large galaxy surveys," 2016, doi:10.48550/ARXIV.1612.06873.

[17] Abbott, T. M. C., Abdalla, F. B., Alarcon, A., Aleksić, J., Allam, S., Allen, S., Amara, A., Annis, J., Asorey, J., Avila, S., Bacon, D., Balbinot, E., Banerji, M., Banik, N., Barkhouse, W., Baumer, M., Baxter, E., Bechtol, K., Becker, M. R., Benoit-Lévy, A., Benson, B. A., Bernstein, G. M., Bertin, E., Blazek, J., Bridle, S. L., Brooks, D., Brout, D., Buckley-Geer, E., Burke, D. L., Busha, M. T., Campos, A., Capozzi, D., Carnero Rosell, A., Carrasco Kind, M., Carretero, J., Castander, F. J., Cawthon, R., Chang, C., Chen, N., Childress, M., Choi, A., Conselice, C., Crittenden, R., Crocce, M., Cunha, C. E., D'Andrea, C. B., da Costa, L. N., Das, R., Davis, T. M., Davis, C., De Vicente, J., DePoy, D. L., DeRose, J., Desai, S., Diehl, H. T., Dietrich, J. P., Dodelson, S., Doel, P., Drlica-Wagner, A., Eifler, T. F., Elliott, A. E., Elsner, F., Elvin-Poole, J., Estrada, J., Evrard, A. E., Fang, Y., Fernandez, E., Ferté, A., Finley, D. A., Flaugher, B., Fosalba, P., Friedrich, O., Frieman, J., García-Bellido, J., Garcia-Fernandez, M., Gatti, M., Gaztanaga, E., Gerdes, D. W., Giannantonio, T., Gill, M. S. S., Glazebrook, K., Goldstein, D. A., Gruen, D., Gruendl, R. A., Gschwend, J., Gutierrez, G., Hamilton, S., Hartley, W. G., Hinton, S. R., Honscheid, K., Hoyle, B., Huterer, D., Jain, B., James, D. J., Jarvis, M., Jeltema, T., Johnson, M. D., Johnson, M. W. G., Kacprzak, T., Kent, S., Kim, A. G., King, A., Kirk, D., Kokron, N., Kovacs, A., Krause, E., Krawiec, C., Kremin, A., Kuehn, K., Kuhlmann, S., Kuropatkin, N., Lacasa, F., Lahav, O., Li, T. S., Liddle, A. R., Lidman, C., Lima, M., Lin, H., MacCrann, N., Maia, M. A. G., Makler, M., Manera, M., March, M., Marshall, J. L., Martini, P., McMahon, R. G.,

Melchior, P., Menanteau, F., Miquel, R., Miranda, V., Mudd, D., Muir, J., Möller, A., Neilsen, E., Nichol, R. C., Nord, B., Nugent, P., Ogando, R. L. C., Palmese, A., Peacock, J., Peiris, H. V., Peoples, J., Percival, W. J., Petravick, D., Plazas, A. A., Porredon, A., Prat, J., Pujol, A., Rau, M. M., Refregier, A., Ricker, P. M., Roe, N., Rollins, R. P., Romer, A. K., Roodman, A., Rosenfeld, R., Ross, A. J., Rozo, E., Rykoff, E. S., Sako, M., Salvador, A. I., Samuroff, S., Sánchez, C., Sanchez, E., Santiago, B., Scarpine, V., Schindler, R., Scolnic, D., Secco, L. F., Serrano, S., Sevilla-Noarbe, I., Sheldon, E., Smith, R. C., Smith, M., Smith, J., Soares-Santos, M., Sobreira, F., Suchyta, E., Tarle, G., Thomas, D., Troxel, M. A., Tucker, D. L., Tucker, B. E., Uddin, S. A., Varga, T. N., Vielzeuf, P., Vikram, V., Vivas, A. K., Walker, A. R., Wang, M., Wechsler, R. H., Weller, J., Wester, W., Wolf, R. C., Yanny, B., Yuan, F., Zenteno, A., Zhang, B., Zhang, Y., and Zuntz, J., "Dark energy survey year 1 results: Cosmological constraints from galaxy clustering and weak lensing," Phys. Rev. D, vol. 98, p. 043526, 2018, doi:10.1103/PhysRevD.98.043526.

[18] Baumann, D., "Lecture notes: Cosmology," 2021.

[19] Amendola, L., "Lecture notes: Cosmology," 2022.

[20] Ryden, B., Introduction to Cosmology. Cambridge University Press, 2017, https://books.google.cl/books?id=tCqejgEACAAJ.

[21] Peacock, J., Cosmological Physics. Cambridge Astrophysics, Cambridge University Press, 1999, https://books.google.cl/books?id=t8O-yylU0j0C.

[22] Einstein, A., "Die Grundlage der allgemeinen Relativitätstheorie," Annalen der Physik, vol. 354, no. 7, pp. 769–822, 1916, doi:10.1002/andp.19163540702.

[23] Einstein, A., "Zur elektrodynamik bewegter körper," Annalen der Physik, vol. 322, no. 10, pp. 891–921, 1905, doi:https://doi.org/10.1002/andp.19053221004.

[24] Friedmann, A., "Uber die Krummung des Raumes," Zeitschrift fur Physik, vol. 10, pp. 377–386, 1922, doi:10.1007/BF01332580.

[25] Lemaître, G., "L'Univers en expansion," Annales de la Socieacute;teacute; Scientifique de Bruxelles, vol. 53, p. 51, 1933.

[26] Robertson, H. P., "Kinematics and World-Structure," Apj, vol. 82, p. 284, 1935, doi:10.1086/143681.

[27] Robertson, H. P., "Kinematics and World-Structure II.," Apj, vol. 83, p. 187, 1936, doi:10.1086/143716.

[28] Robertson, H. P., "Kinematics and World-Structure III.," Apj, vol. 83, p. 257, 1936, doi:10.1086/143726.

[29] Walker, A. G., "On Milne's Theory of World-Structure," Proceedings of the London Mathematical Society, vol. 42, pp. 90–127, 1937, doi:10.1112/plms/s2-42.1.90.

[30] Einstein, A., "Cosmological considerations on the general theory of relativity," Philosophical Problems in Science (Zagadnienia Filozoficzne w Nauce), p. 183–204, 1917, https://zfn.edu.pl/index.php/zfn/article/view/416.

[31] de Sitter, W., "Further remarks on the solutions of the field-equations of the Einstein's theory of gravitation," Koninklijke Nederlandse Akademie van Wetenschappen Proceedings Series B Physical Sciences, vol. 20, pp. 1309–1312, 1918.

[32] Zwicky, F., "Die Rotverschiebung von extragalaktischen Nebeln," Helvetica Physica Acta, vol. 6, pp. 110–127, 1933.

[33] Rubin, V. C., "The rotation of spiral galaxies," Science, vol. 220, no. 4604, pp. 1339–1344, 1983, doi:10.1126/science.220.4604.1339.

[34] Rubin, V. C., "Dark matter in spiral galaxies," Scientific American, vol. 248, no. 6, pp. 96–109, 1983, http://www.jstor.org/stable/24968923 (visitado el 2022-05-03).

[35] Silk, J. *et al.*, Particle Dark Matter: Observations, Models and Searches. Cambridge: Cambridge Univ. Press, 2010, doi:10.1017/CBO9780511770739.

[36] Mukhanov, V., Physical Foundations of Cosmology. Oxford: Cambridge University Press, 2005, doi:10.1017/CBO9780511790553.

[37] Akrami, P. C. Y., Ashdown, M., Aumont, J., Baccigalupi, C., Ballardini, M., Banday, A. J., Barreiro, R. B., Bartolo, N., Basak, S., Benabed, K., Bersanelli, M., Bielewicz, P., Bond, J. R., Borrill, J., Bouchet, F. R., Boulanger, F., Bucher, M., Burigana, C., Calabrese, E., Cardoso, J.-F., Carron, J., Casaponsa, B., Challinor, A., Colombo, L. P. L., Combet, C., Crill, B. P., Cuttaia, F., de Bernardis, P., de Rosa, A., de Zotti, G., Delabrouille, J., Delouis, J.-M., Valentino, E. D., Dickinson, C., Diego, J. M., Donzelli, S., Doré, O., Ducout, A., Dupac, X., Efstathiou, G., Elsner, F., Enßlin, T. A., Eriksen, H. K., Falgarone, E., Fernandez-Cobos, R., Finelli, F., Forastieri, F., Frailis, M., Fraisse, A. A., Franceschi, E., Frolov, A., Galeotta, S., Galli, S., Ganga, K., Génova-Santos, R. T., Gerbino, M., Ghosh, T., González-Nuevo, J., Górski, K. M., Gratton, S., Gruppuso, A., Gudmundsson, J. E., Handley, W., Hansen, F. K., Helou, G., Herranz, D., Hildebrandt, S. R., Huang, Z., Jaffe, A. H., Karakci, A., Keihänen, E., Keskitalo, R., Kiiveri, K., Kim, J., Kisner, T. S., Krachmalnicoff, N., Kunz, M., Kurki-Suonio, H., Lagache, G., Lamarre, J.-M., Lasenby, A., Lattanzi, M., Lawrence, C. R., Jeune, M. L., Levrier, F., Liguori, M., Lilje, P. B., Lindholm, V., López-Caniego, M., Lubin, P. M., Ma, Y.-Z., Macías-Pérez, J. F., Maggio, G., Maino, D., Mandolesi, N., Mangilli, A., Marcos-Caballero, A., Maris, M., Martin, P. G., Martínez-González, E., Matarrese, S., Mauri, N., McEwen, J. D., Meinhold, P. R., Melchiorri, A., Mennella, A., Migliaccio, M., Miville-Deschênes, M.-A., Molinari, D., Moneti, A., Montier, L., Morgante, G., Natoli, P., Oppizzi, F., Pagano, L., Paoletti, D., Partridge, B., Peel, M., Pettorino, V., Piacentini, F., Polenta, G., Puget, J.-L., Rachen, J. P., Reinecke, M., Remazeilles, M., Renzi, A., Rocha, G., Roudier, G., Rubiño-Martín, J. A., Ruiz-Granados, B., Salvati, L., Sandri, M., Savelainen, M., Scott, D., Seljebotn, D. S., Sirignano, C., Spencer, L. D., Suur-Uski, A.-S., Tauber, J. A., Tavagnacco, D., Tenti, M., Thommesen, H., Toffolatti, L., Tomasi, M., Trombetti, T., Valiviita, J., Tent, B. V., Vielva, P., Villa, F., Vittorio, N., Wandelt, B. D., Wehus, I. K., Zacchei, A., and Zonca, A., "Planck 2018 results iv. diffuse component separation," Astronomy & Astrophysics, vol. 641, p. A4, 2020, doi:10.1051/0004-6361/201833881.

[38] Peebles, P. J. E., "Primeval helium abundance and the primeval fireball," Phys. Rev. Lett., vol. 16, pp. 410–413, 1966, doi:10.1103/PhysRevLett.16.410.

[39] Abazajian, K., Adelman-McCarthy, J., Agüeros, M., Allam, S., Anderson, S., Annis, J., Bahcall, N., Baldry, I., Bastian, S., Berlind, A., Bernardi, M., Blanton, M., Blythe, N., Bochanski, J., Boroski, W., Brewington, H., Briggs, J., Brinkmann, J., Brunner, R., and Zucker, D., "The first data release of the sloan digital sky survey," The Astronomical Journal, v.126, 2081-2086 (2003), vol. 126, 2003, doi:10.1086/378165.

[40] Peacock, J. A., Cole, S., Norberg, P., Baugh, C. M., Bland-Hawthorn, J., Bridges, T., Cannon, R. D., Colless, M., Collins, C., Couch, W., Dalton, G., Deeley, K., De Propris, R., Driver, S. P., Efstathiou, G., Ellis, R. S., Frenk, C. S., Glazebrook, K., Jackson, C., Lahav, O., Lewis, I., Lumsden, S., Maddox, S., Percival, W. J., Peterson, B. A., Price, I., Sutherland, W., and Taylor, K., "A measurement of the cosmological mass density from clustering in the 2dF Galaxy Redshift Survey," Nature, vol. 410, no. 6825, pp. 169–173, 2001.

[41] Freedman, R., Geller, R., and Kaufmann, W., Universe. W. H. Freeman, 2010, https://books.google.cl/books?id=XfR8QgAACAAJ.

[42] Jones, D. O., Rodney, S. A., Riess, A. G., Mobasher, B., Dahlen, T., McCully, C., Frederiksen, T. F., Casertano, S., Hjorth, J., Keeton, C. R., Koekemoer, A., Strolger, L.-G., Wiklind, T. G., Challis, P., Graur, O., Hayden, B., Patel, B., Weiner, B. J., Filippenko, A. V., Garnavich, P., Jha, S. W., Kirshner, R. P., Ferguson, H. C., Grogin, N. A., and Kocevski, D., "The discovery of the most distant known type ia supernova at redshift 1.914," The Astrophysical Journal, vol. 768, p. 166, 2013, doi:10.1088/0004-637x/768/2/166.

[43] Hogg, D. W., "Distance measures in cosmology," 1999, doi:10.48550/ARXIV.ASTRO-PH/9905116.

[44] Aghanim, N., Akrami, Y., Ashdown, M., Aumont, J., Baccigalupi, C., Ballardini, M., Banday, A. J., Barreiro, R. B., Bartolo, N., and et al., "Planck 2018 results," Astronomy and Astrophysics, vol. 641, p. A6, 2020, doi:10.1051/0004-6361/201833910.

[45] Springel, V., White, S. D. M., Jenkins, A., Frenk, C. S., Yoshida, N., Gao, L., Navarro, J., Thacker, R., Croton, D., Helly, J., Peacock, J. A., Cole, S., Thomas, P., Couchman, H., Evrard, A., Colberg, J., and Pearce, F., "Simulations of the formation, evolution and clustering of galaxies and quasars," Nature, vol. 435, no. 7042, pp. 629–636, 2005, doi:10.1038/nature03597.

[46] Leclercq, F., Pisani, A., and Wandelt, B., "Cosmology: From theory to data, from data to theory," 2014, doi:10.3254/978-1-61499-476-3-189.

[47] Bernardeau, F., Colombi, S., Gaztañaga, E., and Scoccimarro, R., "Large-scale structure of the universe and cosmological perturbation theory," Physics Reports, vol. 367, pp. 1–248, 2002, doi:10.1016/s0370-1573(02)00135-7.

[48] Friedrich, O., "Statistical properties of the cosmic density field beyond 2-point statistics," 2018, http://nbn-resolving.de/urn:nbn:de:bvb:19-234011.

[49] Hamilton, A. J. S., "Formulae for growth factors in expanding universes containing matter and a cosmological constant," mnras, vol. 322, no. 2, pp. 419–425, 2001, doi:10.1046/j.1365-8711.2001.04137.x.

[50] Akrami, P. C. Y., Arroja, F. M. A., Ashdown, M., Aumont, J., Baccigalupi, C., Ballardini, M., Banday, A. J., Barreiro, R. B., Bartolo, N., Basak, S., Battye, R. A., Benabed, K., Bernard, J.-P., Bersanelli, M., Bielewicz, P., Bock, J. J., Bond, J. R., Borrill, J., Bouchet, F. R., Boulanger, F., Bucher, M., Burigana, C., Butler, R. C., Calabrese, E., Cardoso, J.-F., Carron, J., Casaponsa, B., Challinor, A., Chiang, H. C., Colombo, L. P. L., Combet, C., Contreras, D., Crill, B. P., Cuttaia, F., Bernardis, P., de Zotti, G., Delabrouille, J., Delouis, J. M., D'esert, F. X., Valentino, E. D., Dickinson, C., Diego, J. M., Donzelli, S., Dor'e, O., Douspis, M., Ducout, A., Dupac, X., Efstathiou, G. P.,

Elsner, F., Ensslin, T. A., Eriksen, H. K., Falgarone, E., Fantaye, Y. T., Fergusson, J. R., Fernández-Cobos, R., Finelli, F., Forastieri, F., Frailis, M., Franceschi, E., Frolov, A. V., Galeotta, S., Galli, S., Ganga, K., G'enova-Santos, R. T., Gerbino, M., Ghosh, T., Gonz'alez-Nuevo, J., G'orski, K. M., Gratton, S., Gruppuso, A., Gudmundsson, J. E., Hamann, J., Handley, W. J., Hansen, F. K., Helou, G., Herranz, D., Hivon, E., Huang, Z., Jaffe, A. H., Jones, W. C., Karakci, A., Keihanen, E., Keskitalo, R., Kiiveri, K., Kim, J., Kisner, T., Knox, L., Krachmalnicoff, N., Kunz, M., Kurki-Suonio, H., Lagache, G., Lamarre, J.-M., Langer, M., Lasenby, A. N., Lattanzi, M., Lawrence, C. R., Jeune, M. L., Leahy, J., Lesgourgues, J., Levrier, F., Lewis, A. M., Liguori, M., Lilje, P. B., Lilley, M., Lindholm, V., L'opez-Caniego, M., Lubin, P., Ma, Y.-Z., Mac'ias-P'erez, J.-F., Maggio, G., Maino, D., Mandolesi, N., Mangilli, A., Marcos-Caballero, A., Maris, M., Martin, P. G., Mart'inez-Gonz'alez, E., Matarrese, S., Mauri, N., McEwen, J. D., Meerburg, P. D., Meinhold, P., Melchiorri, A., Mennella, A., Migliaccio, M., Millea, M., Mitra, S., Miville-Deschênes, M.-A., Molinari, D., Moneti, A., Montier, L., Morgante, G., Moss, A., Mottet, S., Munchmeyer, M., Natoli, P., Nørgaard-Nielsen, H. U., Oxborrow, C. A., Pagano, L., Paoletti, D., Partridge, B., Patanchon, G., Pearson, T. J., Peel, M., Peiris, H. V., Perrotta, F., Pettorino, V., Piacentini, F., Polastri, L., Polenta, G., Puget, J.-L., Rachen, J. P., Reinecke, M., Remazeilles, M., Renzi, A., Rocha, G. M., Rosset, C. D., Roudier, G. M., Rubino-Mart'in, J. A., Ruiz-Granados, B., Salvati, L., Sandri, M., Savelainen, M., Scott, D., Shellard, E. P. S., Shiraishi, M., Sirignano, C., Sirri, G., Spencer, L. D., Sunyaev, R. A., Suur-Uski, A.-S., Tauber, J., Tavagnacco, D., Tenti, M., Terenzi, L., Toffolatti, L., Tomasi, M., Trombetti, T., Valiviita, J., van Tent, B., Vibert, L., Vielva, P., Villa, F., Vittorio, N., Wandelt, B. D., Wehus, I. K., White, M., White, S. D. M., Zacchei, A., and Zonca, A., "Planck 2018 results. i. overview and the cosmological legacy of planck," Astronomy and Astrophysics, vol. 641, pp. 1–56, 2018.

[51] Bernardeau, F. and Kofman, L., "Properties of the Cosmological Density Distribution Function," Apj, vol. 443, p. 479, 1995, doi:10.1086/175542.

[52] Bartelmann, M. and Schneider, P., "Weak gravitational lensing," Physics Reports, vol. 340, pp. 291–472, 2001, doi:10.1016/s0370-1573(00)00082-x.

[53] Kilbinger, M., Fu, L., Heymans, C., Simpson, F., Benjamin, J., Erben, T., Harnois-Dé raps, J., Hoekstra, H., Hildebrandt, H., Kitching, T. D., Mellier, Y., Miller, L., Waerbeke, L. V., Benabed, K., Bonnett, C., Coupon, J., Hudson, M. J., Kuijken, K., Rowe, B., Schrabback, T., Semboloni, E., Vafaei, S., and Velander, M., "CFHTLenS: combined probe cosmological model comparison using 2d weak gravitational lensing," Monthly Notices of the Royal Astronomical Society, vol. 430, pp. 2200–2220, 2013, doi:10.1093/mnras/stt041.

[54] Heymans, C., Grocutt, E., Heavens, A., Kilbinger, M., Kitching, T. D., Simpson, F., Benjamin, J., Erben, T., Hildebrandt, H., Hoekstra, H., Mellier, Y., Miller, L., Waerbeke, L. V., Brown, M. L., Coupon, J., Fu, L., Harnois-Dé raps, J., Hudson, M. J., Kuijken, K., Rowe, B., Schrabback, T., Semboloni, E., Vafaei, S., and Velander, M., "CFHTLenS tomographic weak lensing cosmological parameter constraints: Mitigating the impact of intrinsic galaxy alignments," Monthly Notices of the Royal Astronomical Society, vol. 432, pp. 2433–2453, 2013, doi:10.1093/mnras/stt601.

[55] Abbott, T., Abdalla, F. B., Allam, S., Amara, A., Annis, J., Armstrong, R., Bacon, D.,

Banerji, M., Bauer, A. H., Baxter, E., Becker, M. R., Benoit-Lévy, A., Bernstein, R. A., Bernstein, G. M., Bertin, E., Blazek, J., Bonnett, C., Bridle, S. L., Brooks, D., Bruderer, C., Buckley-Geer, E., Burke, D. L., Busha, M. T., Capozzi, D., Carnero Rosell, A., Carrasco Kind, M., Carretero, J., Castander, F. J., Chang, C., Clampitt, J., Crocce, M., Cunha, C. E., D'Andrea, C. B., da Costa, L. N., Das, R., DePoy, D. L., Desai, S., Diehl, H. T., Dietrich, J. P., Dodelson, S., Doel, P., Drlica-Wagner, A., Efstathiou, G., Eifler, T. F., Erickson, B., Estrada, J., Evrard, A. E., Fausti Neto, A., Fernandez, E., Finley, D. A., Flaugher, B., Fosalba, P., Friedrich, O., Frieman, J., Gangkofner, C., Garcia-Bellido, J., Gaztanaga, E., Gerdes, D. W., Gruen, D., Gruendl, R. A., Gutierrez, G., Hartley, W., Hirsch, M., Honscheid, K., Huff, E. M., Jain, B., James, D. J., Jarvis, M., Kacprzak, T., Kent, S., Kirk, D., Krause, E., Kravtsov, A., Kuehn, K., Kuropatkin, N., Kwan, J., Lahav, O., Leistedt, B., Li, T. S., Lima, M., Lin, H., MacCrann, N., March, M., Marshall, J. L., Martini, P., McMahon, R. G., Melchior, P., Miller, C. J., Miquel, R., Mohr, J. J., Neilsen, E., Nichol, R. C., Nicola, A., Nord, B., Ogando, R., Palmese, A., Peiris, H. V., Plazas, A. A., Refregier, A., Roe, N., Romer, A. K., Roodman, A., Rowe, B., Rykoff, E. S., Sabiu, C., Sadeh, I., Sako, M., Samuroff, S., Sanchez, E., Sánchez, C., Seo, H., Sevilla-Noarbe, I., Sheldon, E., Smith, R. C., Soares-Santos, M., Sobreira, F., Suchyta, E., Swanson, M. E. C., Tarle, G., Thaler, J., Thomas, D., Troxel, M. A., Vikram, V., Walker, A. R., Wechsler, R. H., Weller, J., Zhang, Y., and Zuntz, J., "Cosmology from cosmic shear with dark energy survey science verification data," Phys. Rev. D, vol. 94, p. 022001, 2016, doi:10.1103/PhysRevD.94.022001.

[56] van Uitert, E., Joachimi, B., Joudaki, S., Amon, A., Heymans, C., Köhlinger, F., Asgari, M., Blake, C., Choi, A., Erben, T., Farrow, D. J., Harnois-Déraps, J., Hildebrandt, H., Hoekstra, H., Kitching, T. D., Klaes, D., Kuijken, K., Merten, J., Miller, L., Nakajima, R., Schneider, P., Valentijn, E., and Viola, M., "KiDS+GAMA: cosmology constraints from a joint analysis of cosmic shear, galaxy-galaxy lensing, and angular clustering," Monthly Notices of the Royal Astronomical Society, vol. 476, no. 4, pp. 4662–4689, 2018, doi:10.1093/mnras/sty551.

[57] Albrecht, A., Bernstein, G., Cahn, R., Freedman, W. L., Hewitt, J., Hu, W., Huth, J., Kamionkowski, M., Kolb, E. W., Knox, L., Mather, J. C., Staggs, S., and Suntzeff, N. B., "Report of the dark energy task force," 2006, doi:10.48550/ARXIV.ASTRO-PH/0609591.

[58] Collaboration:, D. E. S., Abbott, T., Abdalla, F. B., Aleksić, J., Allam, S., Amara, A., Bacon, D., Balbinot, E., Banerji, M., Bechtol, K., Benoit-Lévy, A., Bernstein, G. M., Bertin, E., Blazek, J., Bonnett, C., Bridle, S., Brooks, D., Brunner, R. J., Buckley-Geer, E., Burke, D. L., Caminha, G. B., Capozzi, D., Carlsen, J., Carnero-Rosell, A., Carollo, M., Carrasco-Kind, M., Carretero, J., Castander, F. J., Clerkin, L., Collett, T., Conselice, C., Crocce, M., Cunha, C. E., D'Andrea, C. B., da Costa, L. N., Davis, T. M., Desai, S., Diehl, H. T., Dietrich, J. P., Dodelson, S., Doel, P., Drlica-Wagner, A., Estrada, J., Etherington, J., Evrard, A. E., Fabbri, J., Finley, D. A., Flaugher, B., Foley, R. J., Fosalba, P., Frieman, J., García-Bellido, J., Gaztanaga, E., Gerdes, D. W., Giannantonio, T., Goldstein, D. A., Gruen, D., Gruendl, R. A., Guarnieri, P., Gutierrez, G., Hartley, W., Honscheid, K., Jain, B., James, D. J., Jeltema, T., Jouvel, S., Kessler, R., King, A., Kirk, D., Kron, R., Kuehn, K., Kuropatkin, N., Lahav, O., Li, T. S., Lima, M., Lin, H., Maia, M. A. G., Makler, M., Manera, M., Maraston, C., Marshall, J. L., Martini, P., McMahon, R. G., Melchior, P., Merson, A., Miller, C. J., Miquel,

R., Mohr, J. J., Morice-Atkinson, X., Naidoo, K., Neilsen, E., Nichol, R. C., Nord, B., Ogando, R., Ostrovski, F., Palmese, A., Papadopoulos, A., Peiris, H. V., Peoples, J., Percival, W. J., Plazas, A. A., Reed, S. L., Refregier, A., Romer, A. K., Roodman, A., Ross, A., Rozo, E., Rykoff, E. S., Sadeh, I., Sako, M., Sánchez, C., Sanchez, E., Santiago, B., Scarpine, V., Schubnell, M., Sevilla-Noarbe, I., Sheldon, E., Smith, M., Smith, R. C., Soares-Santos, M., Sobreira, F., Soumagnac, M., Suchyta, E., Sullivan, M., Swanson, M., Tarle, G., Thaler, J., Thomas, D., Thomas, R. C., Tucker, D., Vieira, J. D., Vikram, V., Walker, A. R., Wechsler, R. H., Weller, J., Wester, W., Whiteway, L., Wilcox, H., Yanny, B., Zhang, Y., and Zuntz, J., "The Dark Energy Survey: more than dark energy – an overview," Monthly Notices of the Royal Astronomical Society, vol. 460, pp. 1270–1299, 2016, doi:10.1093/mnras/stw641.

[59] Sevilla-Noarbe, I., Bechtol, K., Carrasco Kind, M., Carnero Rosell, A., Becker, M. R., Drlica-Wagner, A., Gruendl, R. A., Rykoff, E. S., Sheldon, E., Yanny, B., and et al., "Dark Energy Survey Year 3 results: Photometric data set for cosmology," The Astrophysical Journal Supplement Series, vol. 254, p. 24, 2021, doi:10.3847/1538-4365/abeb66.

[60] de Jong, R. S., Bellido-Tirado, O., Chiappini, C., Depagne, E., Haynes, R., Johl, D., Schnurr, O., Schwope, A., Walcher, J., Dionies, F., and et al., "4most: 4-metre multi-object spectroscopic telescope," Ground-based and Airborne Instrumentation for Astronomy IV, 2012, doi:10.1117/12.926239.

[61] Aihara, H., Arimoto, N., Armstrong, R., Arnouts, S., Bahcall, N. A., Bickerton, S., Bosch, J., Bundy, K., Capak, P. L., Chan, J. H. H., and et al., "The hyper suprime-cam ssp survey: Overview and survey design," Publications of the Astronomical Society of Japan, vol. 70, 2017, doi:10.1093/pasj/psx066.

[62] Dawson, K. S., Kneib, J.-P., Percival, W. J., Alam, S., Albareti, F. D., Anderson, S. F., Armengaud, E., Aubourg, É ., Bailey, S., Bautista, J. E., Berlind, A. A., Bershady, M. A., Beutler, F., Bizyaev, D., Blanton, M. R., Blomqvist, M., Bolton, A. S., Bovy, J., Brandt, W. N., Brinkmann, J., Brownstein, J. R., Burtin, E., Busca, N. G., Cai, Z., Chuang, C.-H., Clerc, N., Comparat, J., Cope, F., Croft, R. A. C., Cruz-Gonzalez, I., da Costa, L. N., Cousinou, M.-C., Darling, J., de la Macorra, A., de la Torre, S., Delubac, T., du Mas des Bourboux, H., Dwelly, T., Ealet, A., Eisenstein, D. J., Eracleous, M., Escoffier, S., Fan, X., Finoguenov, A., Font-Ribera, A., Frinchaboy, P., Gaulme, P., Georgakakis, A., Green, P., Guo, H., Guy, J., Ho, S., Holder, D., Huehnerhoff, J., Hutchinson, T., Jing, Y., Jullo, E., Kamble, V., Kinemuchi, K., Kirkby, D., Kitaura, F.-S., Klaene, M. A., Laher, R. R., Lang, D., Laurent, P., Goff, J.-M. L., Li, C., Liang, Y., Lima, M., Lin, Q., Lin, W., Lin, Y.-T., Long, D. C., Lundgren, B., MacDonald, N., Maia, M. A. G., Malanushenko, E., Malanushenko, V., Mariappan, V., McBride, C. K., McGreer, I. D., Ménard, B., Merloni, A., Meza, A., Montero-Dorta, A. D., Muna, D., Myers, A. D., Nandra, K., Naugle, T., Newman, J. A., Noterdaeme, P., Nugent, P., Ogando, R., Olmstead, M. D., Oravetz, A., Oravetz, D. J., Padmanabhan, N., Palanque-Delabrouille, N., Pan, K., Parejko, J. K., Pâris, I., Peacock, J. A., Petitjean, P., Pieri, M. M., Pisani, A., Prada, F., Prakash, A., Raichoor, A., Reid, B., Rich, J., Ridl, J., Rodriguez-Torres, S., Rosell, A. C., Ross, A. J., Rossi, G., Ruan, J., Salvato, M., Sayres, C., Schneider, D. P., Schlegel, D. J., Seljak, U., Seo, H.-J., Sesar, B., Shandera, S., Shu, Y., Slosar, A., Sobreira, F., Streblyanska, A., Suzuki, N., Taylor, D., Tao, C., Tinker, J. L., Tojeiro, R., Vargas-Magaña, M., Wang, Y., Weaver, B. A., Weinberg, D. H.,

White, M., Wood-Vasey, W. M., Yeche, C., Zhai, Z., Zhao, C., bo Zhao, G., Zheng, Z., Zhu, G. B., and Zou, H., "The SDSS-IV extended baryon oscillation spectroscopic survey: overview and early data," The Astronomical Journal, vol. 151, p. 44, 2016, doi:10.3847/0004-6256/151/2/44.

[63] Collaboration, D., Aghamousa, A., Aguilar, J., Ahlen, S., Alam, S., Allen, L. E., Prieto, C. A., Annis, J., Bailey, S., Balland, C., Ballester, O., Baltay, C., Beaufore, L., Bebek, C., Beers, T. C., Bell, E. F., Bernal, J. L., Besuner, R., Beutler, F., Blake, C., Bleuler, H., Blomqvist, M., Blum, R., Bolton, A. S., Briceno, C., Brooks, D., Brownstein, J. R., Buckley-Geer, E., Burden, A., Burtin, E., Busca, N. G., Cahn, R. N., Cai, Y.-C., Cardiel-Sas, L., Carlberg, R. G., Carton, P.-H., Casas, R., Castander, F. J., Cervantes-Cota, J. L., Claybaugh, T. M., Close, M., Coker, C. T., Cole, S., Comparat, J., Cooper, A. P., Cousinou, M. C., Crocce, M., Cuby, J.-G., Cunningham, D. P., Davis, T. M., Dawson, K. S., de la Macorra, A., Vicente, J. D., Delubac, T., Derwent, M., Dey, A., Dhungana, G., Ding, Z., Doel, P., Duan, Y. T., Ealet, A., Edelstein, J., Eftekharzadeh, S., Eisenstein, D. J., Elliott, A., Escoffier, S., Evatt, M., Fagrelius, P., Fan, X., Fanning, K., Farahi, A., Farihi, J., Favole, G., Feng, Y., Fernandez, E., Findlay, J. R., Finkbeiner, D. P., Fitzpatrick, M. J., Flaugher, B., Flender, S., Font-Ribera, A., Forero-Romero, J. E., Fosalba, P., Frenk, C. S., Fumagalli, M., Gaensicke, B. T., Gallo, G., Garcia-Bellido, J., Gaztanaga, E., Fusillo, N. P. G., Gerard, T., Gershkovich, I., Giannantonio, T., Gillet, D., de Rivera, G. G., Gonzalez-Perez, V., Gott, S., Graur, O., Gutierrez, G., Guy, J., Habib, S., Heetderks, H., Heetderks, I., Heitmann, K., Hellwing, W. A., Herrera, D. A., Ho, S., Holland, S., Honscheid, K., Huff, E., Hutchinson, T. A., Huterer, D., Hwang, H. S., Laguna, J. M. I., Ishikawa, Y., Jacobs, D., Jeffrey, N., Jelinsky, P., Jennings, E., Jiang, L., Jimenez, J., Johnson, J., Joyce, R., Jullo, E., Juneau, S., Kama, S., Karcher, A., Karkar, S., Kehoe, R., Kennamer, N., Kent, S., Kilbinger, M., Kim, A. G., Kirkby, D., Kisner, T., Kitanidis, E., Kneib, J.-P., Koposov, S., Kovacs, E., Koyama, K., Kremin, A., Kron, R., Kronig, L., Kueter-Young, A., Lacey, C. G., Lafever, R., Lahav, O., Lambert, A., Lampton, M., Landriau, M., Lang, D., Lauer, T. R., Goff, J.-M. L., Guillou, L. L., Suu, A. L. V., Lee, J. H., Lee, S.-J., Leitner, D., Lesser, M., Levi, M. E., L'Huillier, B., Li, B., Liang, M., Lin, H., Linder, E., Loebman, S. R., Lukić, Z., Ma, J., MacCrann, N., Magneville, C., Makarem, L., Manera, M., Manser, C. J., Marshall, R., Martini, P., Massey, R., Matheson, T., McCauley, J., McDonald, P., McGreer, I. D., Meisner, A., Metcalfe, N., Miller, T. N., Miquel, R., Moustakas, J., Myers, A., Naik, M., Newman, J. A., Nichol, R. C., Nicola, A., da Costa, L. N., Nie, J., Niz, G., Norberg, P., Nord, B., Norman, D., Nugent, P., O'Brien, T., Oh, M., Olsen, K. A. G., Padilla, C., Padmanabhan, H., Padmanabhan, N., Palanque-Delabrouille, N., Palmese, A., Pappalardo, D., Pâris, I., Park, C., Patej, A., Peacock, J. A., Peiris, H. V., Peng, X., Percival, W. J., Perruchot, S., Pieri, M. M., Pogge, R., Pollack, J. E., Poppett, C., Prada, F., Prakash, A., Probst, R. G., Rabinowitz, D., Raichoor, A., Ree, C. H., Refregier, A., Regal, X., Reid, B., Reil, K., Rezaie, M., Rockosi, C. M., Roe, N., Ronayette, S., Roodman, A., Ross, A. J., Ross, N. P., Rossi, G., Rozo, E., Ruhlmann-Kleider, V., Rykoff, E. S., Sabiu, C., Samushia, L., Sanchez, E., Sanchez, J., Schlegel, D. J., Schneider, M., Schubnell, M., Secroun, A., Seljak, U., Seo, H.-J., Serrano, S., Shafieloo, A., Shan, H., Sharples, R., Sholl, M. J., Shourt, W. V., Silber, J. H., Silva, D. R., Sirk, M. M., Slosar, A., Smith, A., Smoot, G. F., Som, D., Song, Y.-S., Sprayberry, D., Staten, R., Stefanik, A., Tarle, G., Tie, S. S., Tinker, J. L., Tojeiro, R., Valdes, F., Valenzuela, O., Valluri, M., Vargas-Magana, M., Verde, L., Walker, A. R., Wang, J.,

Wang, Y., Weaver, B. A., Weaverdyck, C., Wechsler, R. H., Weinberg, D. H., White, M., Yang, Q., Yeche, C., Zhang, T., Zhao, G.-B., Zheng, Y., Zhou, X., Zhou, Z., Zhu, Y., Zou, H., and Zu, Y., "The DESI Experiment Part I: Science, targeting, and survey design," 2016.

[64] LSST Science Collaboration, Abell, P. A., Allison, J., Anderson, S. F., Andrew, J. R., Angel, J. R. P., Armus, L., Arnett, D., Asztalos, S. J., Axelrod, T. S., and et al., "LSST Science Book, Version 2.0," ArXiv e-prints, 2009.

[65] Ivezić, v., Kahn, S. M., Tyson, J. A., Abel, B., Acosta, E., Allsman, R., Alonso, D., AlSayyad, Y., Anderson, S. F., Andrew, J., and et al., "Lsst: From science drivers to reference design and anticipated data products," The Astrophysical Journal, vol. 873, p. 111, 2019, doi:10.3847/1538-4357/ab042c.

[66] Laureijs, R., Amiaux, J., Arduini, S., Augueres, J. L., Brinchmann, J., Cole, R., Cropper, M., Dabin, C., Duvet, L., Ealet, A., Garilli, B., Gondoin, P., Guzzo, L., Hoar, J., Hoekstra, H., Holmes, R., Kitching, T., Maciaszek, T., Mellier, Y., Pasian, F., Percival, W., Rhodes, J., Criado, G. S., Sauvage, M., Scaramella, R., Valenziano, L., Warren, S., Bender, R., Castander, F., Cimatti, A., Fèvre, O. L., Kurki-Suonio, H., Levi, M., Lilje, P., Meylan, G., Nichol, R., Pedersen, K., Popa, V., Lopez, R. R., Rix, H. W., Rottgering, H., Zeilinger, W., Grupp, F., Hudelot, P., Massey, R., Meneghetti, M., Miller, L., Paltani, S., Paulin-Henriksson, S., Pires, S., Saxton, C., Schrabback, T., Seidel, G., Walsh, J., Aghanim, N., Amendola, L., Bartlett, J., Baccigalupi, C., Beaulieu, J. P., Benabed, K., Cuby, J. G., Elbaz, D., Fosalba, P., Gavazzi, G., Helmi, A., Hook, I., Irwin, M., Kneib, J. P., Kunz, M., Mannucci, F., Moscardini, L., Tao, C., Teyssier, R., Weller, J., Zamorani, G., Osorio, M. R. Z., Boulade, O., Foumond, J. J., Di Giorgio, A., Guttridge, P., James, A., Kemp, M., Martignac, J., Spencer, A., Walton, D., Blümchen, T., Bonoli, C., Bortoletto, F., Cerna, C., Corcione, L., Fabron, C., Jahnke, K., Ligori, S., Madrid, F., Martin, L., Morgante, G., Pamplona, T., Prieto, E., Riva, M., Toledo, R., Trifoglio, M., Zerbi, F., Abdalla, F., Douspis, M., Grenet, C., Borgani, S., Bouwens, R., Courbin, F., Delouis, J. M., Dubath, P., Fontana, A., Frailis, M., Grazian, A., Koppenhöfer, J., Mansutti, O., Melchior, M., Mignoli, M., Mohr, J., Neissner, C., Noddle, K., Poncet, M., Scodeggio, M., Serrano, S., Shane, N., Starck, J. L., Surace, C., Taylor, A., Verdoes-Kleijn, G., Vuerli, C., Williams, O. R., Zacchei, A., Altieri, B., Sanz, I. E., Kohley, R., Oosterbroek, T., Astier, P., Bacon, D., Bardelli, S., Baugh, C., Bellagamba, F., Benoist, C., Bianchi, D., Biviano, A., Branchini, E., Carbone, C., Cardone, V., Clements, D., Colombi, S., Conselice, C., Cresci, G., Deacon, N., Dunlop, J., Fedeli, C., Fontanot, F., Franzetti, P., Giocoli, C., Garcia-Bellido, J., Gow, J., Heavens, A., Hewett, P., Heymans, C., Holland, A., Huang, Z., Ilbert, O., Joachimi, B., Jennins, E., Kerins, E., Kiessling, A., Kirk, D., Kotak, R., Krause, O., Lahav, O., van Leeuwen, F., Lesgourgues, J., Lombardi, M., Magliocchetti, M., Maguire, K., Majerotto, E., Maoli, R., Marulli, F., Maurogordato, S., McCracken, H., McLure, R., Melchiorri, A., Merson, A., Moresco, M., Nonino, M., Norberg, P., Peacock, J., Pello, R., Penny, M., Pettorino, V., Di Porto, C., Pozzetti, L., Quercellini, C., Radovich, M., Rassat, A., Roche, N., Ronayette, S., Rossetti, E., Sartoris, B., Schneider, P., Semboloni, E., Serjeant, S., Simpson, F., Skordis, C., Smadja, G., Smartt, S., Spano, P., Spiro, S., Sullivan, M., Tilquin, A., Trotta, R., Verde, L., Wang, Y., Williger, G., Zhao, G., Zoubian, J., and Zucca, E., "Euclid definition study report," 2011, doi:10.48550/ARXIV.1110.3193.

[67] Akeson, R., Armus, L., Bachelet, E., Bailey, V., Bartusek, L., Bellini, A., Benford, D.,

Bennett, D., Bhattacharya, A., Bohlin, R., Boyer, M., Bozza, V., Bryden, G., Calchi Novati, S., Carpenter, K., Casertano, S., Choi, A., Content, D., Dayal, P., Dressler, A., Doré, O., Fall, S. M., Fan, X., Fang, X., Filippenko, A., Finkelstein, S., Foley, R., Furlanetto, S., Kalirai, J., Gaudi, B. S., Gilbert, K., Girard, J., Grady, K., Greene, J., Guhathakurta, P., Heinrich, C., Hemmati, S., Hendel, D., Henderson, C., Henning, T., Hirata, C., Ho, S., Huff, E., Hutter, A., Jansen, R., Jha, S., Johnson, S., Jones, D., Kasdin, J., Kelly, P., Kirshner, R., Koekemoer, A., Kruk, J., Lewis, N., Macintosh, B., Madau, P., Malhotra, S., Mandel, K., Massara, E., Masters, D., McEnery, J., McQuinn, K., Melchior, P., Melton, M., Mennesson, B., Peeples, M., Penny, M., Perlmutter, S., Pisani, A., Plazas, A., Poleski, R., Postman, M., Ranc, C., Rauscher, B., Rest, A., Roberge, A., Robertson, B., Rodney, S., Rhoads, J., Rhodes, J., Ryan, Russell, J., Sahu, K., Sand, D., Scolnic, D., Seth, A., Shvartzvald, Y., Siellez, K., Smith, A., Spergel, D., Stassun, K., Street, R., Strolger, L.-G., Szalay, A., Trauger, J., Troxel, M. A., Turnbull, M., van der Marel, R., von der Linden, A., Wang, Y., Weinberg, D., Williams, B., Windhorst, R., Wollack, E., Wu, H.-Y., Yee, J., and Zimmerman, N., "The Wide Field Infrared Survey Telescope: 100 Hubbles for the 2020s," arXiv e-prints, p. arXiv:1902.05569, 2019.

[68] Abbott, T., Abdalla, F., Alarcon, A., Aleksić, J., Allam, S., Allen, S., Amara, A., Annis, J., Asorey, J., Avila, S., Bacon, D., Balbinot, E., Banerji, M., Banik, N., Barkhouse, W., Baumer, M., Baxter, E., Bechtol, K., Becker, M., Benoit-Lévy, A., Benson, B., Bernstein, G., Bertin, E., Blazek, J., Bridle, S., Brooks, D., Brout, D., Buckley-Geer, E., Burke, D., Busha, M., Campos, A., Capozzi, D., Rosell, A. C., Kind, M. C., Carretero, J., Castander, F., Cawthon, R., Chang, C., Chen, N., Childress, M., Choi, A., Conselice, C., Crittenden, R., Crocce, M., Cunha, C., D'Andrea, C., da Costa, L., Das, R., Davis, T., Davis, C., Vicente, J. D., DePoy, D., DeRose, J., Desai, S., Diehl, H., Dietrich, J., Dodelson, S., Doel, P., Drlica-Wagner, A., Eifler, T., Elliott, A., Elsner, F., Elvin-Poole, J., Estrada, J., Evrard, A., Fang, Y., Fernandez, E., Ferté, A., Finley, D., Flaugher, B., Fosalba, P., Friedrich, O., Frieman, J., García-Bellido, J., Garcia-Fernandez, M., Gatti, M., Gaztanaga, E., Gerdes, D., Giannantonio, T., Gill, M., Glazebrook, K., Goldstein, D., Gruen, D., Gruendl, R., Gschwend, J., Gutierrez, G., Hamilton, S., Hartley, W., Hinton, S., Honscheid, K., Hoyle, B., Huterer, D., Jain, B., James, D., Jarvis, M., Jeltema, T., Johnson, M., Johnson, M., Kacprzak, T., Kent, S., Kim, A., King, A., Kirk, D., Kokron, N., Kovacs, A., Krause, E., Krawiec, C., Kremin, A., Kuehn, K., Kuhlmann, S., Kuropatkin, N., Lacasa, F., Lahav, O., Li, T., Liddle, A., Lidman, C., Lima, M., Lin, H., MacCrann, N., Maia, M., Makler, M., Manera, M., March, M., Marshall, J., Martini, P., McMahon, R., Melchior, P., Menanteau, F., Miquel, R., Miranda, V., Mudd, D., Muir, J., Möller, A., Neilsen, E., Nichol, R., Nord, B., Nugent, P., Ogando, R., Palmese, A., Peacock, J., Peiris, H., Peoples, J., Percival, W., Petravick, D., Plazas, A., Porredon, A., Prat, J., Pujol, A., Rau, M., Refregier, A., Ricker, P., Roe, N., Rollins, R., Romer, A., Roodman, A., Rosenfeld, R., Ross, A., Rozo, E., Rykoff, E., Sako, M., Salvador, A., Samuroff, S., Sánchez, C., Sanchez, E., Santiago, B., Scarpine, V., Schindler, R., Scolnic, D., Secco, L., Serrano, S., Sevilla-Noarbe, I., Sheldon, E., Smith, R., Smith, M., Smith, J., Soares-Santos, M., Sobreira, F., Suchyta, E., Tarle, G., Thomas, D., Troxel, M., Tucker, D., Tucker, B., Uddin, S., Varga, T., Vielzeuf, P., Vikram, V., Vivas, A., Walker, A., Wang, M., Wechsler, R., Weller, J., Wester, W., Wolf, R., Yanny, B., Yuan, F., Zenteno, A., Zhang, B., Zhang, Y., and and, J. Z., "Dark energy survey year 1 results: Cosmological constraints from galaxy clustering and weak

lensing," Physical Review D, vol. 98, 2018, doi:10.1103/physrevd.98.043526.

[69] Gruen, D., Friedrich, O., Amara, A., Bacon, D., Bonnett, C., Hartley, W., Jain, B., Jarvis, M., Kacprzak, T., Krause, E., Mana, A., Rozo, E., Rykoff, E., Seitz, S., Sheldon, E., Troxel, M., Vikram, V., Abbott, T., Abdalla, F., Allam, S., Armstrong, R., Banerji, M., Bauer, A., Becker, M., Benoit-Levy, A., Bernstein, G., Bernstein, R., Bertin, E., Bridle, S., Brooks, D., Buckley-Geer, E., Burke, D., Capozzi, D., Rosell, A., Kind, M., Carretero, J., Crocce, M., Cunha, C., D'Andrea, C., Costa, L., DePoy, D., Desai, S., Diehl, H., Dietrich, J., Doel, P., Eifler, T., Neto, A., Fernandez, E., Flaugher, B., Fosalba, P., Frieman, J., Gerdes, D., Gruendl, R., Gutierrez, G., Honscheid, K., James, D., Kuehn, K., Kuropatkin, N., Lahav, O., Li, T., Lima, M., Maia, M., March, M., Martini, P., Melchior, P., Miller, C., Miquel, R., Mohr, J., Nord, B., Ogando, R., Plazas, A., Reil, K., Romer, A., Roodman, A., Sako, M., Sanchez, E., Scarpine, V., Schubnell, M., Sevilla-Noarbe, I., Smith, R., Soares-Santos, M., Sobreira, F., Suchyta, E., Swanson, M., Tarle, G., Thaler, J., Thomas, D., Walker, A., Wechsler, R., Weller, J., Zhang, Y., and Zuntz, J., "Weak lensing by galaxy troughs in des science verification data," MNRAS, vol. 455, no. 3, pp. 3367–3380, 2016, doi:10.1093/mnras/stv2506.

[70] DeRose, J., Wechsler, R. H., Becker, M. R., Busha, M. T., Rykoff, E. S., MacCrann, N., Erickson, B., Evrard, A. E., Kravtsov, A., Gruen, D., Allam, S., Avila, S., Bridle, S., Brooks, D., Buckley-Geer, E., Rosell, A. C., Kind, M. C., Carretero, J., Castander, F. J., Cawthon, R., Crocce, M., da Costa, L. N., Davis, C., De Vicente, J., Dietrich, J. P., Doel, P., Drlica-Wagner, A., Fosalba, P., Frieman, J., Garcia-Bellido, J., Gutierrez, G., Hartley, W. G., Hollowood, D. L., Hoyle, B., James, D. J., Krause, E., Kuehn, K., Kuropatkin, N., Lima, M., Maia, M. A. G., Menanteau, F., Miller, C. J., Miquel, R., Ogando, R. L. C., Malagón, A. P., Romer, A. K., Sanchez, E., Schindler, R., Serrano, S., Sevilla-Noarbe, I., Smith, M., Suchyta, E., Swanson, M. E. C., Tarle, G., and Vikram, V., "The buzzard flock: Dark energy survey synthetic sky catalogs," 2019, doi:10.48550/ARXIV.1901.02401.

[71] Ptolemy, Ptolemy's Almagest (translated and annotated by G.J. Toomer). Duckworth London, 1984.

[72] González-Morán, A. L., Chávez, R., Terlevich, E., Terlevich, R., Fernández-Arenas, D., Bresolin, F., Plionis, M., Melnick, J., Basilakos, S., and Telles, E., "Independent cosmological constraints from high-z hii galaxies: new results from vlt-kmos data," Monthly Notices of the Royal Astronomical Society, vol. 505, no. 1, pp. 1441–1457, 2021, doi:10.1093/mnras/stab1385.

[73] Wei, J.-J. and Melia, F., "Cosmology-independent estimate of the hubble constant and spatial curvature using time-delay lenses and quasars," The Astrophysical Journal, vol. 897, p. 127, 2020, doi:10.3847/1538-4357/ab959b.

[74] Marc, V., Rasanen, S., and Durrer, R., "Model-independent cosmological constraints from the cmb," Journal of Cosmology and Astroparticle Physics, vol. 2010, 2010, doi:10.1088/1475-7516/2010/08/023.

[75] Feigelson, E. D. and Babu, G. J., Modern Statistical Methods for Astronomy: With R Applications. Cambridge University Press, 2012, doi:10.1017/CBO9781139015653.

[76] Kolmogorov, A. N., Foundations of the Theory of Probability. Chelsea Pub Co, 2 ed., 1960.

[77] Kotz, S., Balakrishnan, N., Read, C. B., and Vidakovic, B., Encyclopedia of Statistical Sciences. pub-WILEY, second ed., 2004, doi:https://doi.org/10.1002/0471667196.ess0783.pub2.

[78] Taylor, A., "Course notes for astronomical statistics," 2004.

[79] VanderPlas, J., "Frequentism and Bayesianism: A Python-driven primer," 2014, doi:10.48550/ARXIV.1411.5018.

[80] Crow, E. L. and Shimizu, K., Lognormal Distributions. Statistics: A Series of Textbooks and Monographs, Boca Raton, FL: CRC Press, 1987.

[81] Trotta, R., "Bayes in the sky: Bayesian inference and model selection in cosmology," Contemporary Physics, vol. 49, p. 71–104, 2008, doi:10.1080/00107510802066753.

[82] Bayarri, M. J. and Berger, J. O., "The interplay of Bayesian and frequentist analysis," Statistical Science, vol. 19, no. 1, pp. 58–80, 2004, http://www.jstor.org/stable/4144373 (visitado el 2022-04-24).

[83] Wang, Y. and Blei, D. M., "Frequentist consistency of variational Bayes," Journal of the American Statistical Association, vol. 114, pp. 1147–1161, 2018, doi:10.1080/01621459.2018.1473776.

[84] Lucy, L. B., "Frequentist tests for Bayesian models," Astronomy and Astrophysics, vol. 588, p. A19, 2016, doi:10.1051/0004-6361/201527709.

[85] Silva, I. R., "On the correspondence between frequentist and Bayesian tests," Communications in Statistics - Theory and Methods, vol. 47, no. 14, pp. 3477–3487, 2017, doi:10.1080/03610926.2017.1359296.

[86] Niu, S.-C., "Course notes for opre-6301, introduction to operations research," 2004.

[87] Fisher, R. A. and Russell, E. J., "On the mathematical foundations of theoretical statistics," Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character, vol. 222, no. 594-604, pp. 309–368, 1922, doi:10.1098/rsta.1922.0009.

[88] Aswani, A., "Lecture notes in ieor 165 – engineering statistics, quality control, and forecasting," 2021.

[89] Lambert, J., Photometria Sive De Mensura Et Gradibus Luminis, Colorum Et Umbrae. Augustae Vindelicorum, 1760, http://slubdd.de/katalog?TN_libero_mab2)500288888.

[90] Lagrange, J.-L., "Mémoire sur l'utilité de la méthode de prendre le milieu entre les résultats de plusieurs observations," Miscellanea Taurinensia, pp. 173–234, 1770-1773, https://gallica.bnf.fr/ark:/12148/bpt6k215570z/f174.

[91] Edwards, A. W. F., "The history of likelihood," International Statistical Review / Revue Internationale de Statistique, vol. 42, no. 1, pp. 9–15, 1974, http://www.jstor.org/stable/1402681.

[92] Fisher, R. A., "On an absolute criterion for fitting frequency curves," Statistical Science, vol. 12, no. 1, pp. 39–41, 1912, http://www.jstor.org/stable/2246266.

[93] Fisher, R. A., "On the 'probable error' of a coefficient of correlation deduced from a small sample," Metron, vol. 1, pp. 1–32, 1921, https://ci.nii.ac.jp/naid/10012392243/en/.

[94] Yan, L., "Lecture notes in cs109 – probability for computer scientists," 2019.

[95] Lehmann, E. and Lehmann, A., Testing Statistical Hypotheses. Probability and Statistics Series, Wiley, 1986, https://books.google.de/books?id=jexQAAAAMAAJ.

[96] Johnson, D., "Course notes for elec531, statistical signal processing," 2011.

[97] Birnbaum, A., "Median unbiased estimators," Bulletin of Mathematical Statistics, vol. 11, pp. 25–34, 1964, doi:10.5109/13011.

[98] Neyman, J., "Outline of a theory of statistical estimation based on the classical theory of probability," Philosophical Transactions of the Royal Society of London. Series A, Mathematical and Physical Sciences, vol. 236, no. 767, pp. 333–380, 1937, doi:10.1098/rsta.1937.0005.

[99] Dodge, Y., The Concise Encyclopedia of Statistics. Springer New York, 2008, doi:10.1007/978-0-387-32833-1.

[100] Dawson, A., Research Methods and Statistics in Physical Education. ETP, 2018, https://books.google.de/books?id=juQxwAEACAAJ.

[101] Feldman, G. J. and Cousins, R. D., "Unified approach to the classical statistical analysis of small signals," Physical Review D, vol. 57, pp. 3873–3889, 1998, doi:10.1103/physrevd.57.3873.

[102] Secco, L., Samuroff, S., Krause, E., Jain, B., Blazek, J., Raveri, M., Campos, A., Amon, A., Chen, A., Doux, C., Choi, A., Gruen, D., Bernstein, G., Chang, C., DeRose, J., Myles, J., Ferté, A., Lemos, P., Huterer, D., Prat, J., Troxel, M., MacCrann, N., Liddle, A., Kacprzak, T., Fang, X., Sánchez, C., Pandey, S., Dodelson, S., Chintalapati, P., Hoffmann, K., Alarcon, A., Alves, O., Andrade-Oliveira, F., Baxter, E., Bechtol, K., Becker, M., Brandao-Souza, A., Camacho, H., Rosell, A. C., Kind, M. C., Cawthon, R., Cordero, J., Crocce, M., Davis, C., Valentino, E. D., Drlica-Wagner, A., Eckert, K., Eifler, T., Elidaiana, M., Elsner, F., Elvin-Poole, J., Everett, S., Fosalba, P., Friedrich, O., Gatti, M., Giannini, G., Gruendl, R., Harrison, I., Hartley, W., Herner, K., Huang, H., Huff, E., Jarvis, M., Jeffrey, N., Kuropatkin, N., Leget, P.-F., Muir, J., Mccullough, J., Alsina, A. N., Omori, Y., Park, Y., Porredon, A., Rollins, R., Roodman, A., Rosenfeld, R., Ross, A., Rykoff, E., Sanchez, J., Sevilla-Noarbe, I., Sheldon, E., Shin, T., Troja, A., Tutusaus, I., Varga, T., Weaverdyck, N., Wechsler, R., Yanny, B., Yin, B., Zhang, Y., Zuntz, J., Abbott, T., Aguena, M., Allam, S., Annis, J., Bacon, D., Bertin, E., Bhargava, S., Bridle, S., Brooks, D., Buckley-Geer, E., Burke, D., Carretero, J., Costanzi, M., da Costa, L., Vicente, J. D., Diehl, H., Dietrich, J., Doel, P., Ferrero, I., Flaugher, B., Frieman, J., García-Bellido, J., Gaztanaga, E., Gerdes, D., Giannantonio, T., Gschwend, J., Gutierrez, G., Hinton, S., Hollowood, D., Honscheid, K., Hoyle, B., James, D., Jeltema, T., Kuehn, K., Lahav, O., Lima, M., Lin, H., Maia, M., Marshall, J., Martini, P., Melchior, P., Menanteau, F., Miquel, R., Mohr, J., Morgan, R., Ogando, R., Palmese, A., Paz-Chinchón, F., Petravick, D., Pieres, A., Malagón, A. P., Rodriguez-Monroy, M., Romer, A., Sanchez, E., Scarpine, V., Schubnell, M., Scolnic, D., Serrano, S., Smith, M., Soares-Santos, M., Suchyta, E., Swanson, M., Tarle, G., Thomas, D., and and, C. T., "Dark Energy Survey Year 3 results: Cosmology from cosmic shear and robustness to modeling uncertainty," Physical Review D, vol. 105, 2022, doi:10.1103/physrevd.105.023515.

[103] Krause, E., Fang, X., Pandey, S., Secco, L. F., Alves, O., Huang, H., Blazek, J., Prat,

J., Zuntz, J., Eifler, T. F., MacCrann, N., DeRose, J., Crocce, M., Porredon, A., Jain, B., Troxel, M. A., Dodelson, S., Huterer, D., Liddle, A. R., Leonard, C. D., Amon, A., Chen, A., Elvin-Poole, J., Ferté, A., Muir, J., Park, Y., Samuroff, S., Brandao-Souza, A., Weaverdyck, N., Zacharegkas, G., Rosenfeld, R., Campos, A., Chintalapati, P., Choi, A., Di Valentino, E., Doux, C., Herner, K., Lemos, P., Mena-Fernández, J., Omori, Y., Paterno, M., Rodriguez-Monroy, M., Rogozenski, P., Rollins, R. P., Troja, A., Tutusaus, I., Wechsler, R. H., Abbott, T. M. C., Aguena, M., Allam, S., Andrade-Oliveira, F., Annis, J., Bacon, D., Baxter, E., Bechtol, K., Bernstein, G. M., Brooks, D., Buckley-Geer, E., Burke, D. L., Rosell, A. C., Kind, M. C., Carretero, J., Castander, F. J., Cawthon, R., Chang, C., Costanzi, M., da Costa, L. N., Pereira, M. E. S., De Vicente, J., Desai, S., Diehl, H. T., Doel, P., Everett, S., Evrard, A. E., Ferrero, I., Flaugher, B., Fosalba, P., Frieman, J., García-Bellido, J., Gaztanaga, E., Gerdes, D. W., Giannantonio, T., Gruen, D., Gruendl, R. A., Gschwend, J., Gutierrez, G., Hartley, W. G., Hinton, S. R., Hollowood, D. L., Honscheid, K., Hoyle, B., Huff, E. M., James, D. J., Kuehn, K., Kuropatkin, N., Lahav, O., Lima, M., Maia, M. A. G., Marshall, J. L., Martini, P., Melchior, P., Menanteau, F., Miquel, R., Mohr, J. J., Morgan, R., Myles, J., Palmese, A., Paz-Chinchón, F., Petravick, D., Pieres, A., Malagón, A. A. P., Sanchez, E., Scarpine, V., Schubnell, M., Serrano, S., Sevilla-Noarbe, I., Smith, M., Soares-Santos, M., Suchyta, E., Tarle, G., Thomas, D., To, C., Varga, T. N., and Weller, J., "Dark energy survey year 3 results: Multi-probe modeling strategy and validation," 2021, doi:10.48550/ARXIV.2105.13548.

[104] Biller, S. D. and Oser, S. M., "Another look at confidence intervals: Proposal for a more relevant and transparent approach," Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, vol. 774, pp. 103–119, 2015, doi:10.1016/j.nima.2014.11.081.

[105] Casella, G. and George, E. I., "Explaining the Gibbs sampler," The American Statistician, vol. 46, no. 3, pp. 167–174, 1992, http://www.jstor.org/stable/2685208 (visitado el 2022-04-20).

[106] Hastings, W. K., "Monte Carlo sampling methods using Markov chains and their applications," Biometrika, vol. 57, pp. 97–109, 1970, doi:10.1093/biomet/57.1.97.

[107] Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., and Teller, E., "Equation of state calculations by fast computing machines," The Journal of Chemical Physics, vol. 21, no. 6, pp. 1087–1092, 1953, doi:10.1063/1.1699114.

[108] Casella, G., Robert, C. P., and Wells, M. T., "Generalized accept-reject sampling schemes," en Institute of Mathematical Statistics Lecture Notes - Monograph Series, pp. 342–347, Institute of Mathematical Statistics, 2004, doi:10.1214/lnms/1196285403.

[109] Kahn, H. and Harris, T. E., "Estimation of particle transmission by random sampling," National Bureau of Standards applied mathematics series, vol. 12, pp. 27–30, 1951.

[110] Andrieu, C., de Freitas, N., Doucet, A., and Jordan, M. I., "An introduction to MCMC for machine learning," Machine Learning, vol. 50, no. 1/2, pp. 5–43, 2003, doi:10.1023/a:1020281327116.

[111] Newman, M. E. J. and Barkema, G. T., Monte Carlo methods in statistical physics. Oxford: Clarendon Press, 1999.

[112] Huterer, D., "Lecture notes in structure formation in the universe," 2021.

[113] Kitagawa, G. and Sato, S., Sequential Monte Carlo methods in practice. Springer, 2001.

[114] Geweke, J., "Bayesian inference in econometric models using monte carlo integration," Econometrica, vol. 57, no. 6, pp. 1317–1339, 1989, http://www.jstor.org/stable/1913710 (visitado el 2022-04-19).

[115] Devroye, L., Non-Uniform Random Variate Generation. Springer-Verlag, 1986, http://cg.scs.carleton.ca/~luc/rnbookindex.html.

[116] Marriott, F. H. C. and Institute., I. S., A dictionary of statistical terms / by F.H.C. Marriott. Longman Scientific & Technical Harlow, Essex, 5th ed. ed., 1989.

[117] Taylor, J. R., An Introduction to Error Analysis: The Study of Uncertainties in Physical Measurements. University Science Books, 2 sub ed., 1996.

[118] Spergel, D. N., Verde, L., Peiris, H. V., Komatsu, E., Nolta, M. R., Bennett, C. L., Halpern, M., Hinshaw, G., Jarosik, N., Kogut, A., and et al., "First-year wilkinson microwave anisotropy probe (wmap) observations: Determination of cosmological parameters," The Astrophysical Journal Supplement Series, vol. 148, p. 175–194, 2003, doi:10.1086/377226.

[119] Spergel, D. N., Bean, R., Dore, O., Nolta, M. R., Bennett, C. L., Dunkley, J., Hinshaw, G., Jarosik, N., Komatsu, E., Page, L., and et al., "Three-year wilkinson microwave anisotropy probe (wmap) observations: Implications for cosmology," The Astrophysical Journal Supplement Series, vol. 170, p. 377–408, 2007, doi:10.1086/513700.

[120] Hinshaw, G., Weiland, J. L., Hill, R. S., Odegard, N., Larson, D., Bennett, C. L., Dunkley, J., Gold, B., Greason, M. R., Jarosik, N., and et al., "Five-year wilkinson microwave anisotropy probe observations: Data processing, sky maps, and basic results," The Astrophysical Journal Supplement Series, vol. 180, p. 225–245, 2009, doi:10.1088/0067-0049/180/2/225.

[121] Jarosik, N., Bennett, C. L., Dunkley, J., Gold, B., Greason, M. R., Halpern, M., Hill, R. S., Hinshaw, G., Kogut, A., Komatsu, E., and et al., "Seven-year wilkinson microwave anisotropy probe (wmap) observations: Sky maps, systematic errors, and basic results," The Astrophysical Journal Supplement Series, vol. 192, p. 14, 2011, doi:10.1088/0067-0049/192/2/14.

[122] Ade, P. A. R., Aghanim, N., Alves, M. I. R., Armitage-Caplan, C., Arnaud, M., Ashdown, M., Atrio-Barandela, F., Aumont, J., Aussel, H., and et al., "Planck2013 results. i. overview of products and scientific results," Astronomy and Astrophysics, vol. 571, p. A1, 2014, doi:10.1051/0004-6361/201321529.

[123] Ade, P. A. R., Aghanim, N., Ahmed, Z., Aikin, R. W., Alexander, K. D., Arnaud, M., Aumont, J., Baccigalupi, C., Banday, A. J., Barkats, D., Barreiro, R. B., Bartlett, J. G., Bartolo, N., Battaner, E., Benabed, K., Benoît, A., Benoit-Lévy, A., Benton, S. J., Bernard, J.-P., Bersanelli, M., Bielewicz, P., Bischoff, C. A., Bock, J. J., Bonaldi, A., Bonavera, L., Bond, J. R., Borrill, J., Bouchet, F. R., Boulanger, F., Brevik, J. A., Bucher, M., Buder, I., Bullock, E., Burigana, C., Butler, R. C., Buza, V., Calabrese, E., Cardoso, J.-F., Catalano, A., Challinor, A., Chary, R.-R., Chiang, H. C., Christensen, P. R., Colombo, L. P. L., Combet, C., Connors, J., Couchot, F., Coulais, A., Crill, B. P., Curto, A., Cuttaia, F., Danese, L., Davies, R. D., Davis, R. J., de Bernardis, P., de Rosa, A., de Zotti, G., Delabrouille, J., Delouis, J.-M., Désert, F.-X., Dickinson, C., Diego,

J. M., Dole, H., Donzelli, S., Doré, O., Douspis, M., Dowell, C. D., Duband, L., Ducout, A., Dunkley, J., Dupac, X., Dvorkin, C., Efstathiou, G., Elsner, F., Enßlin, T. A., Eriksen, H. K., Falgarone, E., Filippini, J. P., Finelli, F., Fliescher, S., Forni, O., Frailis, M., Fraisse, A. A., Franceschi, E., Frejsel, A., Galeotta, S., Galli, S., Ganga, K., Ghosh, T., Giard, M., Gjerløw, E., Golwala, S. R., González-Nuevo, J., Górski, K. M., Gratton, S., Gregorio, A., Gruppuso, A., Gudmundsson, J. E., Halpern, M., Hansen, F. K., Hanson, D., Harrison, D. L., Hasselfield, M., Helou, G., Henrot-Versillé, S., Herranz, D., Hildebrandt, S. R., Hilton, G. C., Hivon, E., Hobson, M., Holmes, W. A., Hovest, W., Hristov, V. V., Huffenberger, K. M., Hui, H., Hurier, G., Irwin, K. D., Jaffe, A. H., Jaffe, T. R., Jewell, J., Jones, W. C., Juvela, M., Karakci, A., Karkare, K. S., Kaufman, J. P., Keating, B. G., Kefeli, S., Keihänen, E., Kernasovskiy, S. A., Keskitalo, R., Kisner, T. S., Kneissl, R., Knoche, J., Knox, L., Kovac, J. M., Krachmalnicoff, N., Kunz, M., Kuo, C. L., Kurki-Suonio, H., Lagache, G., Lähteenmäki, A., Lamarre, J.-M., Lasenby, A., Lattanzi, M., Lawrence, C. R., Leitch, E. M., Leonardi, R., Levrier, F., Lewis, A., Liguori, M., Lilje, P. B., Linden-Vørnle, M., López-Caniego, M., Lubin, P. M., Lueker, M., Macías-Pérez, J. F., Maffei, B., Maino, D., Mandolesi, N., Mangilli, A., Maris, M., Martin, P. G., Martínez-González, E., Masi, S., Mason, P., Matarrese, S., Megerian, K. G., Meinhold, P. R., Melchiorri, A., Mendes, L., Mennella, A., Migliaccio, M., Mitra, S., Miville-Deschênes, M.-A., Moneti, A., Montier, L., Morgante, G., Mortlock, D., Moss, A., Munshi, D., Murphy, J. A., Naselsky, P., Nati, F., Natoli, P., Netterfield, C. B., Nguyen, H. T., Nørgaard-Nielsen, H. U., Noviello, F., Novikov, D., Novikov, I., O'Brient, R., Ogburn, R. W., Orlando, A., Pagano, L., Pajot, F., Paladini, R., Paoletti, D., Partridge, B., Pasian, F., Patanchon, G., Pearson, T. J., Perdereau, O., Perotto, L., Pettorino, V., Piacentini, F., Piat, M., Pietrobon, D., Plaszczynski, S., Pointecouteau, E., Polenta, G., Ponthieu, N., Pratt, G. W., Prunet, S., Pryke, C., Puget, J.-L., Rachen, J. P., Reach, W. T., Rebolo, R., Reinecke, M., Remazeilles, M., Renault, C., Renzi, A., Richter, S., Ristorcelli, I., Rocha, G., Rossetti, M., Roudier, G., Rowan-Robinson, M., Rubiño Martín, J. A., Rusholme, B., Sandri, M., Santos, D., Savelainen, M., Savini, G., Schwarz, R., Scott, D., Seiffert, M. D., Sheehy, C. D., Spencer, L. D., Staniszewski, Z. K., Stolyarov, V., Sudiwala, R., Sunyaev, R., Sutton, D., Suur-Uski, A.-S., Sygnet, J.-F., Tauber, J. A., Teply, G. P., Terenzi, L., Thompson, K. L., Toffolatti, L., Tolan, J. E., Tomasi, M., Tristram, M., Tucci, M., Turner, A. D., Valenziano, L., Valiviita, J., Van Tent, B., Vibert, L., Vielva, P., Vieregg, A. G., Villa, F., Wade, L. A., Wandelt, B. D., Watson, R., Weber, A. C., Wehus, I. K., White, M., White, S. D. M., Willmert, J., Wong, C. L., Yoon, K. W., Yvon, D., Zacchei, A., and Zonca, A., "Joint analysis of BICEP2/Keck array and Planck data," Phys. Rev. Lett., vol. 114, p. 101301, 2015, doi:10.1103/PhysRevLett.114.101301.

[124] Freedman, W. L., Madore, B. F., Gibson, B. K., Ferrarese, L., Kelson, D. D., Sakai, S., Mould, J. R., Kennicutt, Jr., R. C., Ford, H. C., Graham, J. A., and et al., "Final results from the hubble space telescope: Key project to measure the hubble constant," The Astrophysical Journal, vol. 553, p. 47–72, 2001, doi:10.1086/320638.

[125] Bonamente, M., Joy, M. K., LaRoque, S. J., Carlstrom, J. E., Reese, E. D., and Dawson, K. S., "Determination of the cosmic distance scale from Sunyaev-Zel'dovich effect and Chandra X-ray measurements of high-redshift galaxy clusters," The Astrophysical Journal, vol. 647, p. 25–54, 2006, doi:10.1086/505291.

[126] Riess, A. G., Macri, L., Casertano, S., Lampeitl, H., Ferguson, H. C., Filippenko, A. V.,

Jha, S. W., Li, W., and Chornock, R., "A 3 % solution: Determination of the Hubble constant with the Hubble Space Telescope and Wide Field Camera 3," The Astrophysical Journal, vol. 730, p. 119, 2011, doi:10.1088/0004-637x/730/2/119.

[127] Freedman, W. L., Madore, B. F., Scowcroft, V., Burns, C., Monson, A., Persson, S. E., Seibert, M., and Rigby, J., "Carnegie Hubble program: A mid-infrared calibration of the Hubble constant," The Astrophysical Journal, vol. 758, p. 24, 2012, doi:10.1088/0004-637x/758/1/24.

[128] Sorce, J. G., Courtois, H. M., Tully, R. B., Seibert, M., Scowcroft, V., Freedman, W. L., Madore, B. F., Persson, S. E., Monson, A., and Rigby, J., "Calibration of the mid-infrared Tully-Fisher relation," The Astrophysical Journal, vol. 765, p. 94, 2013, doi:10.1088/0004-637x/765/2/94.

[129] Efstathiou, G., "H_0 revisited," Monthly Notices of the Royal Astronomical Society, vol. 440, pp. 1138–1152, 2014, doi:10.1093/mnras/stu278.

[130] Zhang, B. R., Childress, M. J., Davis, T. M., Karpenka, N. V., Lidman, C., Schmidt, B. P., and Smith, M., "A blinded determination of h_0 from low-redshift Type Ia supernovae, calibrated by Cepheid variables," Monthly Notices of the Royal Astronomical Society, vol. 471, p. 2254–2285, 2017, doi:10.1093/mnras/stx1600.

[131] Riess, A. G., Casertano, S., Yuan, W., Macri, L., Anderson, J., MacKenty, J. W., Bowers, J. B., Clubb, K. I., Filippenko, A. V., Jones, D. O., and et al., "New parallaxes of galactic Cepheids from spatially scanning the Hubble Space Telescope: Implications for the Hubble constant," The Astrophysical Journal, vol. 855, p. 136, 2018, doi:10.3847/1538-4357/aaadb7.

[132] Riess, A. G., Casertano, S., Yuan, W., Macri, L. M., and Scolnic, D., "Large Magellanic Cloud Cepheid standards provide a 1 % foundation for the determination of the Hubble constant and stronger evidence for physics beyond ΛCDM," The Astrophysical Journal, vol. 876, p. 85, 2019, doi:10.3847/1538-4357/ab1422.

[133] Riess, A. G., Casertano, S., Yuan, W., Bowers, J. B., Macri, L., Zinn, J. C., and Scolnic, D., "Cosmic distances calibrated to 1 % precision with Gaia EDR3 parallaxes and Hubble Space Telescope photometry of 75 Milky Way Cepheids confirm tension with ΛCDM," The Astrophysical Journal Letters, vol. 908, p. L6, 2020, doi:10.3847/2041-8213/abdbaf.

[134] Riess, A. G., Yuan, W., Macri, L. M., Scolnic, D., Brout, D., Casertano, S., Jones, D. O., Murakami, Y., Breuval, L., Brink, T. G., Filippenko, A. V., Hoffmann, S., Jha, S. W., Kenworthy, W. D., Mackenty, J., Stahl, B. E., and Zheng, W., "A comprehensive measurement of the local value of the Hubble constant with 1 km/s/Mpc uncertainty from the Hubble Space Telescope and the SH0ES team," 2021.

[135] Heymans, C., Tröster, T., Asgari, M., Blake, C., Hildebrandt, H., Joachimi, B., Kuijken, K., Lin, C.-A., Sánchez, A. G., van den Busch, J. L., and et al., "KiDS-1000 cosmology: Multi-probe weak gravitational lensing and spectroscopic galaxy clustering constraints," Astronomy & Astrophysics, vol. 646, p. A140, 2021, doi:10.1051/0004-6361/202039063.

[136] Abdalla, E., Abellán, G. F., Aboubrahim, A., Agnello, A., Akarsu, O., Akrami, Y., Alestas, G., Aloni, D., Amendola, L., Anchordoqui, L. A., Anderson, R. I., Arendse, N., Asgari, M., Ballardini, M., Barger, V., Basilakos, S., Batista, R. C., Battistelli, E. S., Battye, R., Benetti, M., Benisty, D., Berlin, A., de Bernardis, P., Berti, E., Bi-

denko, B., Birrer, S., Blakeslee, J. P., Boddy, K. K., Bom, C. R., Bonilla, A., Borghi, N., Bouchet, F. R., Braglia, M., Buchert, T., Buckley-Geer, E., Calabrese, E., Caldwell, R. R., Camarena, D., Capozziello, S., Casertano, S., Chen, G. C. F., Chluba, J., Chen, A., Chen, H.-Y., Chudaykin, A., Cicoli, M., Copi, C. J., Courbin, F., Cyr-Racine, F.-Y., Czerny, B., Dainotti, M., D'Amico, G., Davis, A.-C., Pérez, J. d. C., de Haro, J., Delabrouille, J., Denton, P. B., Dhawan, S., Dienes, K. R., Di Valentino, E., Du, P., Eckert, D., Escamilla-Rivera, C., Ferté, A., Finelli, F., Fosalba, P., Freedman, W. L., Frusciante, N., Gaztañaga, E., Giarè, W., Giusarma, E., Gómez-Valent, A., Handley, W., Harrison, I., Hart, L., Hazra, D. K., Heavens, A., Heinesen, A., Hildebrandt, H., Hill, J. C., Hogg, N. B., Holz, D. E., Hooper, D. C., Hosseininejad, N., Huterer, D., Ishak, M., Ivanov, M. M., Jaffe, A. H., Jang, I. S., Jedamzik, K., Jimenez, R., Joseph, M., Joudaki, S., Kamionkowski, M., Karwal, T., Kazantzidis, L., Keeley, R. E., Klasen, M., Komatsu, E., Koopmans, L. V. E., Kumar, S., Lamagna, L., Lazkoz, R., Lee, C.-C., Lesgourgues, J., Said, J. L., Lewis, T. R., L'Huillier, B., Lucca, M., Maartens, R., Macri, L. M., Marfatia, D., Marra, V., Martins, C. J. A. P., Masi, S., Matarrese, S., Mazumdar, A., Melchiorri, A., Mena, O., Mersini-Houghton, L., Mertens, J., Milakovic, D., Minami, Y., Miranda, V., Moreno-Pulido, C., Moresco, M., Mota, D. F., Mottola, E., Mozzon, S., Muir, J., Mukherjee, A., Mukherjee, S., Naselsky, P., Nath, P., Nesseris, S., Niedermann, F., Notari, A., Nunes, R. C., Colgain, E. O., Owens, K. A., Ozulker, E., Pace, F., Paliathanasis, A., Palmese, A., Pan, S., Paoletti, D., Bergliaffa, S. E. P., Perivolaropoulos, L., Pesce, D. W., Pettorino, V., Philcox, O. H. E., Pogosian, L., Poulin, V., Poulot, G., Raveri, M., Reid, M. J., Renzi, F., Riess, A. G., Sabla, V. I., Salucci, P., Salzano, V., Saridakis, E. N., Sathyaprakash, B. S., Schmaltz, M., Schöneberg, N., Scolnic, D., Sen, A. A., Sehgal, N., Shafieloo, A., Sheikh-Jabbari, M. M., Silk, J., Silvestri, A., Skara, F., Sloth, M. S., Soares-Santos, M., Peracaula, J. S., Songsheng, Y.-Y., Soriano, J. F., Staicova, D., Starkman, G. D., Szapudi, I., Teixeira, E. M., Thomas, B., Treu, T., Trott, E., van de Bruck, C., Vazquez, J. A., Verde, L., Visinelli, L., Wang, D., Wang, J.-M., Wang, S.-J., Watkins, R., Watson, S., Webb, J. K., Weiner, N., Weltman, A., Witte, S. J., Wojtak, R., Yadav, A. K., Yang, W., Zhao, G.-B., and Zumalacárregui, M., "Cosmology intertwined: A review of the particle physics, astrophysics, and cosmology associated with the cosmological tensions and anomalies," 2022, doi:10.48550/ARXIV.2203.06142.

[137] Dodge, Y., Cox, D., Institute, I. S., and Commenges, D., The Oxford Dictionary of Statistical Terms. Oxford University Press, 2006, https://books.google.de/books?id=_OnjBgpuhWcC.

[138] McLachlan, G. J., "Mahalanobis distance," Resonance, vol. 4, pp. 20–26, 1999, doi:10.1007/bf02834632.

[139] Mahalanobis, P. C., "On the generalized distance in statistics," Proceedings of the National Institute of Sciences (Calcutta), vol. 2, pp. 49–55, 1936.

[140] Etherington, T. R., "Mahalanobis distances and ecological niche modelling: correcting a chi-squared probability error," PeerJ, vol. 7, p. e6678, 2019.

[141] Ruggles, R. and Brodie, H., "An empirical approach to economic intelligence in World War II," Journal of the American Statistical Association, vol. 42, no. 237, pp. 72–91, 1947, doi:10.1080/01621459.1947.10501915.

[142] Johnson, R. W., "Estimating the size of a population," Teaching Statistics, vol. 16,

no. 2, pp. 50–52, 1994, doi:https://doi.org/10.1111/j.1467-9639.1994.tb00688.x.

[143] Laureijs, R., Amiaux, J., Arduini, S., Augueres, J. L., Brinchmann, J., Cole, R., Cropper, M., Dabin, C., Duvet, L., Ealet, A., Garilli, B., Gondoin, P., Guzzo, L., Hoar, J., Hoekstra, H., Holmes, R., Kitching, T., Maciaszek, T., Mellier, Y., Pasian, F., Percival, W., Rhodes, J., Criado, G. S., Sauvage, M., Scaramella, R., Valenziano, L., Warren, S., Bender, R., Castander, F., Cimatti, A., Fèvre, O. L., Kurki-Suonio, H., Levi, M., Lilje, P., Meylan, G., Nichol, R., Pedersen, K., Popa, V., Lopez, R. R., Rix, H. W., Rottgering, H., Zeilinger, W., Grupp, F., Hudelot, P., Massey, R., Meneghetti, M., Miller, L., Paltani, S., Paulin-Henriksson, S., Pires, S., Saxton, C., Schrabback, T., Seidel, G., Walsh, J., Aghanim, N., Amendola, L., Bartlett, J., Baccigalupi, C., Beaulieu, J. P., Benabed, K., Cuby, J. G., Elbaz, D., Fosalba, P., Gavazzi, G., Helmi, A., Hook, I., Irwin, M., Kneib, J. P., Kunz, M., Mannucci, F., Moscardini, L., Tao, C., Teyssier, R., Weller, J., Zamorani, G., Osorio, M. R. Z., Boulade, O., Foumond, J. J., Giorgio, A. D., Guttridge, P., James, A., Kemp, M., Martignac, J., Spencer, A., Walton, D., Blümchen, T., Bonoli, C., Bortoletto, F., Cerna, C., Corcione, L., Fabron, C., Jahnke, K., Ligori, S., Madrid, F., Martin, L., Morgante, G., Pamplona, T., Prieto, E., Riva, M., Toledo, R., Trifoglio, M., Zerbi, F., Abdalla, F., Douspis, M., Grenet, C., Borgani, S., Bouwens, R., Courbin, F., Delouis, J. M., Dubath, P., Fontana, A., Frailis, M., Grazian, A., Koppenhöfer, J., Mansutti, O., Melchior, M., Mignoli, M., Mohr, J., Neissner, C., Noddle, K., Poncet, M., Scodeggio, M., Serrano, S., Shane, N., Starck, J. L., Surace, C., Taylor, A., Verdoes-Kleijn, G., Vuerli, C., Williams, O. R., Zacchei, A., Altieri, B., Sanz, I. E., Kohley, R., Oosterbroek, T., Astier, P., Bacon, D., Bardelli, S., Baugh, C., Bellagamba, F., Benoist, C., Bianchi, D., Biviano, A., Branchini, E., Carbone, C., Cardone, V., Clements, D., Colombi, S., Conselice, C., Cresci, G., Deacon, N., Dunlop, J., Fedeli, C., Fontanot, F., Franzetti, P., Giocoli, C., Garcia-Bellido, J., Gow, J., Heavens, A., Hewett, P., Heymans, C., Holland, A., Huang, Z., Ilbert, O., Joachimi, B., Jennins, E., Kerins, E., Kiessling, A., Kirk, D., Kotak, R., Krause, O., Lahav, O., van Leeuwen, F., Lesgourgues, J., Lombardi, M., Magliocchetti, M., Maguire, K., Majerotto, E., Maoli, R., Marulli, F., Maurogordato, S., McCracken, H., McLure, R., Melchiorri, A., Merson, A., Moresco, M., Nonino, M., Norberg, P., Peacock, J., Pello, R., Penny, M., Pettorino, V., Porto, C. D., Pozzetti, L., Quercellini, C., Radovich, M., Rassat, A., Roche, N., Ronayette, S., Rossetti, E., Sartoris, B., Schneider, P., Semboloni, E., Serjeant, S., Simpson, F., Skordis, C., Smadja, G., Smartt, S., Spano, P., Spiro, S., Sullivan, M., Tilquin, A., Trotta, R., Verde, L., Wang, Y., Williger, G., Zhao, G., Zoubian, J., and Zucca, E., "Euclid definition study report," 2011.

[144] Troxel, M. and Ishak, M., "The intrinsic alignment of galaxies and its impact on weak gravitational lensing in an era of precision cosmology," Physics Reports, vol. 558, p. 1–59, 2015, doi:10.1016/j.physrep.2014.11.001.

[145] Secco, L. F., Samuroff, S., Krause, E., Jain, B., Blazek, J., Raveri, M., Campos, A., Amon, A., Chen, A., Doux, C., Choi, A., Gruen, D., Bernstein, G. M., Chang, C., DeRose, J., Myles, J., Ferté, A., Lemos, P., Huterer, D., Prat, J., Troxel, M. A., MacCrann, N., Liddle, A. R., Kacprzak, T., Fang, X., Sánchez, C., Pandey, S., Dodelson, S., Chintalapati, P., Hoffmann, K., Alarcon, A., Alves, O., Andrade-Oliveira, F., Baxter, E. J., Bechtol, K., Becker, M. R., Brandao-Souza, A., Camacho, H., Rosell, A. C., Kind, M. C., Cawthon, R., Cordero, J. P., Crocce, M., Davis, C., Valentino, E. D., Drlica-Wagner, A., Eckert, K., Eifler, T. F., Elidaiana, M., Elsner, F., Elvin-Poole, J.,

Everett, S., Fosalba, P., Friedrich, O., Gatti, M., Giannini, G., Gruendl, R. A., Harrison, I., Hartley, W. G., Herner, K., Huang, H., Huff, E. M., Jarvis, M., Jeffrey, N., Kuropatkin, N., Leget, P. F., Muir, J., Mccullough, J., Alsina, A. N., Omori, Y., Park, Y., Porredon, A., Rollins, R., Roodman, A., Rosenfeld, R., Ross, A. J., Rykoff, E. S., Sanchez, J., Sevilla-Noarbe, I., Sheldon, E. S., Shin, T., Tutusaus, I., Varga, T. N., Weaverdyck, N., Wechsler, R. H., Yanny, B., Yin, B., Zhang, Y., Zuntz, J., Abbott, T. M. C., Aguena, M., Allam, S., Annis, J., Bacon, D., Bertin, E., Bhargava, S., Bridle, S. L., Brooks, D., Buckley-Geer, E., Burke, D. L., Carretero, J., Costanzi, M., da Costa, L. N., Vicente, J. D., Diehl, H. T., Dietrich, J. P., Doel, P., Ferrero, I., Flaugher, B., Frieman, J., García-Bellido, J., Gaztanaga, E., Gerdes, D. W., Giannantonio, T., Gschwend, J., Gutierrez, G., Hinton, S. R., Hollowood, D. L., Honscheid, K., Hoyle, B., James, D. J., Jeltema, T., Kuehn, K., Lahav, O., Lima, M., Lin, H., Maia, M. A. G., Marshall, J. L., Martini, P., Melchior, P., Menanteau, F., Miquel, R., Mohr, J. J., Morgan, R., Ogando, R. L. C., Palmese, A., Paz-Chinchón, F., Petravick, D., Pieres, A., Malagón, A. A. P., Rodriguez-Monroy, M., Romer, A. K., Sanchez, E., Scarpine, V., Schubnell, M., Scolnic, D., Serrano, S., Smith, M., Soares-Santos, M., Suchyta, E., Swanson, M. E. C., Tarle, G., Thomas, D., and To, C., "Dark energy survey year 3 results: Cosmology from cosmic shear and robustness to modeling uncertainty," 2021.

[146] Asgari, M., Lin, C.-A., Joachimi, B., Giblin, B., Heymans, C., Hildebrandt, H., Kannawadi, A., Stölzner, B., Tröster, T., van den Busch, J. L., and et al., "KiDS-1000 cosmology: Cosmic shear constraints and comparison between two point statistics," Astronomy & Astrophysics, vol. 645, p. A104, 2021, doi:10.1051/0004-6361/202039070.

[147] Li, X., Miyatake, H., Luo, W., More, S., Oguri, M., Hamana, T., Mandelbaum, R., Shirasaki, M., Takada, M., Armstrong, R., Kannawadi, A., Takita, S., Miyazaki, S., Nishizawa, A. J., Malagón, A. A. P., Strauss, M. A., Tanaka, M., and Yoshida, N., "The three-year shear catalog of the Subaru Hyper Suprime-Cam SSP Survey," 2021.

[148] Cui, W., Borgani, S., and Murante, G., "The effect of active galactic nuclei feedback on the halo mass function," Monthly Notices of the Royal Astronomical Society, vol. 441, p. 1769–1782, 2014, doi:10.1093/mnras/stu673.

[149] Velliscig, M., van Daalen, M. P., Schaye, J., McCarthy, I. G., Cacciato, M., Le Brun, A. M. C., and Vecchia, C. D., "The impact of galaxy formation on the total mass, mass profile and abundance of haloes," Monthly Notices of the Royal Astronomical Society, vol. 442, p. 2641–2658, 2014, doi:10.1093/mnras/stu1044.

[150] Mummery, B. O., McCarthy, I. G., Bird, S., and Schaye, J., "The separate and combined effects of baryon physics and neutrino free streaming on large-scale structure," Monthly Notices of the Royal Astronomical Society, vol. 471, p. 227–242, 2017, doi:10.1093/mnras/stx1469.

[151] Schmidt, F., "An n-th order lagrangian forward model for large-scale structure," Journal of Cosmology and Astroparticle Physics, vol. 2021, p. 033, 2021, doi:10.1088/1475-7516/2021/04/033.

[152] Dekel, A. and Lahav, O., "Stochastic nonlinear galaxy biasing," The Astrophysical Journal, vol. 520, p. 24–34, 1999, doi:10.1086/307428.

[153] Scherrer, R. J. and Weinberg, D. H., "Constraints on the effects of locally biased galaxy formation," The Astrophysical Journal, vol. 504, p. 607–611, 1998, doi:10.1086/306113.

[154] Baldauf, T., Seljak, U., Senatore, L., and Zaldarriaga, M., "Galaxy bias and non-linear structure formation in general relativity," Journal of Cosmology and Astroparticle Physics, vol. 2011, p. 031–031, 2011, doi:10.1088/1475-7516/2011/10/031.

[155] Pandey, S., Krause, E., DeRose, J., MacCrann, N., Jain, B., Crocce, M., Blazek, J., Choi, A., Huang, H., To, C., Fang, X., Elvin-Poole, J., Prat, J., Porredon, A., Secco, L. F., Rodriguez-Monroy, M., Weaverdyck, N., Park, Y., Raveri, M., Rozo, E., Rykoff, E. S., Bernstein, G. M., Sánchez, C., Jarvis, M., Troxel, M. A., Zacharegkas, G., Chang, C., Alarcon, A., Alves, O., Amon, A., Andrade-Oliveira, F., Baxter, E., Bechtol, K., Becker, M. R., Camacho, H., Campos, A., Rosell, A. C., Kind, M. C., Cawthon, R., Chen, R., Chintalapati, P., Davis, C., Valentino, E. D., Diehl, H. T., Dodelson, S., Doux, C., Drlica-Wagner, A., Eckert, K., Eifler, T. F., Elsner, F., Everett, S., Farahi, A., Ferté, A., Fosalba, P., Friedrich, O., Gatti, M., Giannini, G., Gruen, D., Gruendl, R. A., Harrison, I., Hartley, W. G., Huff, E. M., Huterer, D., Leget, P. F., McCullough, J., Muir, J., Myles, J., Navarro-Alsina, A., Omori, Y., Rollins, R. P., Roodman, A., Rosenfeld, R., Sevilla-Noarbe, I., Sheldon, E., Shin, T., Troja, A., Tutusaus, I., Varga, T. N., Wechsler, R. H., Yanny, B., Yin, B., Zhang, Y., Zuntz, J., Abbott, T. M. C., Aguena, M., Allam, S., Annis, J., Bacon, D., Bertin, E., Brooks, D., Burke, D. L., Carretero, J., Conselice, C., Costanzi, M., da Costa, L. N., Pereira, M. E. S., Vicente, J. D., Dietrich, J. P., Doel, P., Evrard, A. E., Ferrero, I., Flaugher, B., Frieman, J., García-Bellido, J., Gaztanaga, E., Gerdes, D. W., Giannantonio, T., Gschwend, J., Gutierrez, G., Hinton, S. R., Hollowood, D. L., Honscheid, K., James, D. J., Jeltema, T., Kuehn, K., Kuropatkin, N., Lahav, O., Lima, M., Lin, H., Maia, M. A. G., Marshall, J. L., Melchior, P., Menanteau, F., Miller, C. J., Miquel, R., Mohr, J. J., Morgan, R., Palmese, A., Paz-Chinchón, F., Petravick, D., Pieres, A., Malagón, A. A. P., Sanchez, E., Scarpine, V., Serrano, S., Smith, M., Soares-Santos, M., Suchyta, E., Tarle, G., Thomas, D., and Weller, J., "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," 2021.

[156] Ishikawa, S., Okumura, T., Oguri, M., and Lin, S.-C., "Halo-model analysis of the clustering of photometric luminous red galaxies at $0.10 \leq z \leq 1.05$ from the subaru hyper suprime-cam survey," 2021.

[157] Seo, H. and Eisenstein, D. J., "Improved forecasts for the baryon acoustic oscillations and cosmological distance scale," The Astrophysical Journal, vol. 665, p. 14–24, 2007, doi:10.1086/519549.

[158] Ding, Z., Seo, H.-J., Vlah, Z., Feng, Y., Schmittfull, M., and Beutler, F., "Theoretical systematics of future baryon acoustic oscillation surveys," Monthly Notices of the Royal Astronomical Society, 2018, doi:10.1093/mnras/sty1413.

[159] Rosell, A. C., Rodriguez-Monroy, M., Crocce, M., Elvin-Poole, J., Porredon, A., Ferrero, I., Mena-Fernandez, J., Cawthon, R., Vicente, J. D., Gaztanaga, E., Ross, A. J., Sanchez, E., Sevilla-Noarbe, I., Alves, O., Andrade-Oliveira, F., Asorey, J., Avila, S., Brandao-Souza, A., Camacho, H., Chan, K. C., Ferte, A., Muir, J., Riquelme, W., Rosenfeld, R., Cid, D. S., Hartley, W. G., Weaverdyck, N., Abbott, T., Aguena, M., Sahar, A., Annis, J., Bertin, E., Brooks, D., Buckley-Geer, E., Burke, D., Calcino, J., Carollo, D., Kind, M. C., Carretero, J., Castander, F., Choi, A., Conselice, C., Costanzi, M., da Costa, L., da Silva Pereira, M. E., Davis, T., Desai, S., Diehl, H. T., Doel,

P., Drlica-Wagner, A., Eckert, K., Everett, S., Evrard, A., Flaugher, B., Fosalba, P., Frieman, J., Garcia-Bellido, J., Gerdes, D., Giannantonio, T., Glazebrook, K., Gruen, D., Gruendl, R., Gschwend, J., Gutierrez, G., Hinton, S., Hollowood, D., Honscheid, K., Hoyle, B., Huterer, D., James, D., Kim, A., Krause, E., Kuehn, K., Lahav, O., Lewis, G., Lidman, C., Lima, M., Maia, M., Malik, U., Marshall, J., Menanteau, F., Miquel, R., Mohr, J., Moller, A., Morgan, R., Ogando, R., Palmese, A., Paz-Chinchon, F., Percival, W., Pieres, A., Malagon, A. P., Roodman, A., Scarpine, V., Schubnell, M., Serrano, S., Sharp, R., Sheldon, E., Smith, M., Soares-Santos, M., Suchyta, E., Swanson, M., Tarle, G., Thomas, D., To, C., Tucker, B., Tucker, D., Uddin, S., Varga, T. N., and DES Collaboration, "Dark Energy Survey Year 3 results: Galaxy sample for BAO measurement," 2021.

[160] Padmanabhan, N., Xu, X., Eisenstein, D. J., Scalzo, R., Cuesta, A. J., Mehta, K. T., and Kazin, E., "A 2 % distance to $z = 0.35$ by reconstructing baryon acoustic oscillations – I. Methods and application to the Sloan Digital Sky Survey," Monthly Notices of the Royal Astronomical Society, vol. 427, p. 2132–2145, 2012, doi:10.1111/j.1365-2966.2012.21888.x.

[161] Mandelbaum, R., Blazek, J., Chisari, N. E., Collett, T., Galbany, L., Gawiser, E., Hložek, R. A., Kim, A. G., Leonard, C. D., Lochner, M., Newman, J. A., Perrefort, D. J., Schmidt, S. J., Singh, S., and Sullivan, M., "Wide-field Multi-object Spectroscopy to Enhance Dark Energy Science from LSST," 2019.

[162] MacCrann, N., Becker, M. R., McCullough, J., Amon, A., Gruen, D., Jarvis, M., Choi, A., Troxel, M. A., Sheldon, E., Yanny, B., Herner, K., Dodelson, S., Zuntz, J., Eckert, K., Rollins, R. P., Varga, T. N., Bernstein, G. M., Gruendl, R. A., Harrison, I., Hartley, W. G., Sevilla-Noarbe, I., Pieres, A., Bridle, S. L., Myles, J., Alarcon, A., Everett, S., Sánchez, C., Huff, E. M., Tarsitano, F., Gatti, M., Secco, L. F., Abbott, T. M. C., Aguena, M., Allam, S., Annis, J., Bacon, D., Bertin, E., Brooks, D., Burke, D. L., Rosell, A. C., Kind, M. C., Carretero, J., Costanzi, M., Crocce, M., Pereira, M. E. S., Vicente, J. D., Desai, S., Diehl, H. T., Dietrich, J. P., Doel, P., Eifler, T. F., Ferrero, I., Ferté, A., Flaugher, B., Fosalba, P., Frieman, J., García-Bellido, J., Gaztanaga, E., Gerdes, D. W., Giannantonio, T., Gschwend, J., Gutierrez, G., Hinton, S. R., Hollowood, D. L., Honscheid, K., James, D. J., Lahav, O., Lima, M., Maia, M. A. G., March, M., Marshall, J. L., Martini, P., Melchior, P., Menanteau, F., Miquel, R., Mohr, J. J., Morgan, R., Muir, J., Ogando, R. L. C., Palmese, A., Paz-Chinchón, F., Plazas, A. A., Rodriguez-Monroy, M., Roodman, A., Samuroff, S., Sanchez, E., Scarpine, V., Serrano, S., Smith, M., Soares-Santos, M., Suchyta, E., Swanson, M. E. C., Tarle, G., Thomas, D., To, C., and Wilkinson, R. D., "DES Y3 results: Blending shear and redshift biases in image simulations," 2020.

[163] Georgiou, C., Hoekstra, H., Kuijken, K., Bilicki, M., Dvornik, A., Erben, T., Giblin, B., Heymans, C., Hildebrandt, H., de Jong, J. T. A., and et al., "Halo shapes constrained from a pure sample of central galaxies in KiDS-1000," Astronomy & Astrophysics, vol. 647, p. A185, 2021, doi:10.1051/0004-6361/201937405.

[164] Hoyle, B., Gruen, D., Bernstein, G. M., Rau, M. M., De Vicente, J., Hartley, W. G., Gaztanaga, E., DeRose, J., Troxel, M. A., Davis, C., and et al., "Dark Energy Survey Year 1 results: redshift distributions of the weak-lensing source galaxies," Monthly Notices of the Royal Astronomical Society, vol. 478, p. 592–610, 2018, doi:10.1093/mn

ras/sty957.

[165] Cordero, J. P., Harrison, I., Rollins, R. P., Bernstein, G. M., Bridle, S. L., Alarcon, A., Alves, O., Amon, A., Andrade-Oliveira, F., Camacho, H., Campos, A., Choi, A., DeRose, J., Dodelson, S., Eckert, K., Eifler, T. F., Everett, S., Fang, X., Friedrich, O., Gruen, D., Gruendl, R. A., Hartley, W. G., Huff, E. M., Krause, E., Kuropatkin, N., MacCrann, N., McCullough, J., Myles, J., Pandey, S., Raveri, M., Rosenfeld, R., Rykoff, E. S., Sánchez, C., Sánchez, J., Sevilla-Noarbe, I., Sheldon, E., Troxel, M., Wechsler, R., Yanny, B., Yin, B., Zhang, Y., Aguena, M., Allam, S., Bertin, E., Brooks, D., Burke, D. L., Rosell, A. C., Kind, M. C., Carretero, J., Castander, F. J., Cawthon, R., Costanzi, M., da Costa, L., da Silva Pereira, M. E., Vicente, J. D., Diehl, H. T., Dietrich, J., Doel, P., Elvin-Poole, J., Ferrero, I., Flaugher, B., Fosalba, P., Frieman, J., Garcia-Bellido, J., Gerdes, D., Gschwend, J., Gutierrez, G., Hinton, S., Hollowood, D. L., Honscheid, K., Hoyle, B., James, D., Kuehn, K., Lahav, O., Maia, M. A. G., March, M., Menanteau, F., Miquel, R., Morgan, R., Muir, J., Palmese, A., Paz-Chinchon, F., Pieres, A., Malagón, A. P., Sánchez, E., Scarpine, V., Serrano, S., Smith, M., Soares-Santos, M., Suchyta, E., Swanson, M., Tarle, G., Thomas, D., To, C., and Varga, T. N., "Dark Energy Survey Year 3 results: Marginalisation over redshift distribution uncertainties using ranking of discrete realisations," 2021.

[166] Hildebrandt, H., van den Busch, J. L., Wright, A. H., Blake, C., Joachimi, B., Kuijken, K., Tröster, T., Asgari, M., Bilicki, M., de Jong, J. T. A., and et al., "KiDS-1000 catalogue: Redshift distributions and their calibration," Astronomy & Astrophysics, vol. 647, p. A124, 2021, doi:10.1051/0004-6361/202039018.

[167] Tanaka, M., Coupon, J., Hsieh, B.-C., Mineo, S., Nishizawa, A. J., Speagle, J., Furusawa, H., Miyazaki, S., and Murayama, H., "Photometric redshifts for Hyper Suprime-Cam Subaru strategic program Data Release 1," Publications of the Astronomical Society of Japan, vol. 70, 2017, doi:10.1093/pasj/psx077.

[168] Huang, S., Leauthaud, A., Murata, R., Bosch, J., Price, P., Lupton, R., Mandelbaum, R., Lackner, C., Bickerton, S., Miyazaki, S., and et al., "Characterization and photometric performance of the Hyper Suprime-Cam software pipeline," Publications of the Astronomical Society of Japan, vol. 70, 2017, doi:10.1093/pasj/psx126.

[169] Cawthon, R., Elvin-Poole, J., Porredon, A., Crocce, M., Giannini, G., Gatti, M., Ross, A. J., Rykoff, E. S., Rosell, A. C., DeRose, J., Lee, S., Rodriguez-Monroy, M., Amon, A., Bechtol, K., Vicente, J. D., Gruen, D., Morgan, R., Sanchez, E., Sanchez, J., Sevilla-Noarbe, I., Abbott, T. M. C., Aguena, M., Allam, S., Annis, J., Avila, S., Bacon, D., Bertin, E., Brooks, D., Burke, D. L., Kind, M. C., Carretero, J., Castander, F. J., Choi, A., Costanzi, M., da Costa, L. N., Pereira, M. E. S., Dawson, K., Desai, S., Diehl, H. T., Eckert, K., Everett, S., Ferrero, I., Fosalba, P., Frieman, J., García-Bellido, J., Gaztanaga, E., Gruendl, R. A., Gschwend, J., Gutierrez, G., Hinton, S. R., Hollowood, D. L., Honscheid, K., Huterer, D., James, D. J., Kim, A. G., Kneib, J. P., Kuehn, K., Kuropatkin, N., Lahav, O., Lima, M., Lin, H., Maia, M. A. G., Melchior, P., Menanteau, F., Miquel, R., Mohr, J. J., Muir, J., Myles, J., Palmese, A., Pandey, S., Paz-Chinchón, F., Percival, W. J., Plazas, A. A., Roodman, A., Rossi, G., Scarpine, V., Serrano, S., Smith, M., Soares-Santos, M., Suchyta, E., Swanson, M. E. C., Tarle, G., To, C., Troxel, M. A., Wilkinson, R. D., and the DES Collaboration, "Dark Energy Survey Year 3 results: Calibration of lens sample redshift distributions using clustering redshifts with

BOSS/eBOSS," 2020.

[170] Bernstein, G. M., "Shape measurement biases from underfitting and ellipticity gradients," Monthly Notices of the Royal Astronomical Society, vol. 406, p. 2793–2804, 2010, doi:10.1111/j.1365-2966.2010.16883.x.

[171] Amon, A., Gruen, D., Troxel, M. A., MacCrann, N., Dodelson, S., Choi, A., Doux, C., Secco, L. F., Samuroff, S., Krause, E., Cordero, J., Myles, J., DeRose, J., Wechsler, R. H., Gatti, M., Navarro-Alsina, A., Bernstein, G. M., Jain, B., Blazek, J., Alarcon, A., Ferté, A., Raveri, M., Lemos, P., Campos, A., Prat, J., Sánchez, C., Jarvis, M., Alves, O., Andrade-Oliveira, F., Baxter, E., Bechtol, K., Becker, M. R., Bridle, S. L., Camacho, H., Campos, A., Rosell, A. C., Kind, M. C., Cawthon, R., Chang, C., Chen, R., Chintalapati, P., Crocce, M., Davis, C., Diehl, H. T., Drlica-Wagner, A., Eckert, K., Eifler, T. F., Elvin-Poole, J., Everett, S., Fang, X., Fosalba, P., Friedrich, O., Giannini, G., Gruendl, R. A., Harrison, I., Hartley, W. G., Herner, K., Huang, H., Huff, E. M., Huterer, D., Kuropatkin, N., Leget, P. F., Liddle, A. R., McCullough, J., Muir, J., Pandey, S., Park, Y., Porredon, A., Refregier, A., Rollins, R. P., Roodman, A., Rosenfeld, R., Ross, A. J., Rykoff, E. S., Sanchez, J., Sevilla-Noarbe, I., Sheldon, E., Shin, T., Troja, A., Tutusaus, I., Varga, T. N., Weaverdyck, N., Yanny, B., Yin, B., Zhang, Y., Zuntz, J., Aguena, M., Allam, S., Annis, J., Bacon, D., Bertin, E., Bhargava, S., Brooks, D., Buckley-Geer, E., Burke, D. L., Carretero, J., Costanzi, M., da Costa, L. N., Pereira, M. E. S., Vicente, J. D., Desai, S., Dietrich, J. P., Doel, P., Ferrero, I., Flaugher, B., Frieman, J., García-Bellido, J., Gaztanaga, E., Gerdes, D. W., Giannantonio, T., Gschwend, J., Gutierrez, G., Hinton, S. R., Hollowood, D. L., Honscheid, K., Hoyle, B., James, D. J., Kron, R., Kuehn, K., Lahav, O., Lima, M., Lin, H., Maia, M. A. G., Marshall, J. L., Martini, P., Melchior, P., Menanteau, F., Miquel, R., Mohr, J. J., Morgan, R., Ogando, R. L. C., Palmese, A., Paz-Chinchón, F., Petravick, D., Pieres, A., Malagón, A. A. P., Romer, A. K., Sanchez, E., Scarpine, V., Schubnell, M., Serrano, S., Smith, M., Soares-Santos, M., Suchyta, E., Tarle, G., Thomas, D., To, C., and Weller, J., "Dark Energy Survey Year 3 results: Cosmology from cosmic shear and robustness to data calibration," 2021.

[172] Sugiyama, S., Takada, M., Kobayashi, Y., Miyatake, H., Shirasaki, M., Nishimichi, T., and Park, Y., "Validating a minimal galaxy bias method for cosmological parameter inference using HSC-SDSS mock catalogs," Physical Review D, vol. 102, 2020, doi:10.1103/physrevd.102.083520.

[173] Sartoris, B., Biviano, A., Rosati, P., Mercurio, A., Grillo, C., Ettori, S., Nonino, M., Umetsu, K., Bergamini, P., Caminha, G. B., and et al., "CLASH-VLT: a full dynamical reconstruction of the mass profile of Abell S1063 from 1 kpc out to the virial radius," Astronomy & Astrophysics, vol. 637, p. A34, 2020, doi:10.1051/0004-6361/202037521.

[174] Sellentin, E. and Heavens, A. F., "Parameter inference with estimated covariance matrices," mnras, vol. 456, no. 1, pp. L132–L136, 2016, doi:10.1093/mnrasl/slv190.

[175] Joachimi, B., Lin, C. A., Asgari, M., Tröster, T., Heymans, C., Hildebrandt, H., Köhlinger, F., Sánchez, A. G., Wright, A. H., Bilicki, M., Blake, C., van den Busch, J. L., Crocce, M., Dvornik, A., Erben, T., Getman, F., Giblin, B., Hoekstra, H., Kannawadi, A., Kuijken, K., Napolitano, N. R., Schneider, P., Scoccimarro, R., Sellentin, E., Shan, H. Y., von Wietersheim-Kramsta, M., and Zuntz, J., "KiDS-1000 methodology: Modelling and inference for joint weak gravitational lensing and spectroscopic galaxy

clustering analysis," aap, vol. 646, p. A129, 2021, doi:10.1051/0004-6361/202038831.

[176] Percival, W. J., Friedrich, O., Sellentin, E., and Heavens, A., "Matching Bayesian and frequentist coverage probabilities when using an approximate data covariance matrix," 2021.

[177] van Daalen, M. P., McCarthy, I. G., and Schaye, J., "Exploring the effects of galaxy formation on matter clustering through a library of simulation power spectra," Monthly Notices of the Royal Astronomical Society, vol. 491, p. 2424–2446, 2019, doi:10.1093/mnras/stz3199.

[178] Walther, M., Armengaud, E., Ravoux, C., Palanque-Delabrouille, N., Yèche, C., and Lukić, Z., "Simulating intergalactic gas for DESI-like small scale lymanα forest observations," Journal of Cosmology and Astroparticle Physics, vol. 2021, p. 059, 2021, doi:10.1088/1475-7516/2021/04/059.

[179] Hirata, C. M. and Seljak, U. c. v., "Intrinsic alignment-lensing interference as a contaminant of cosmic shear," Phys. Rev. D, vol. 70, p. 063526, 2004, doi:10.1103/PhysRevD.70.063526.

[180] Kiessling, A., Cacciato, M., Joachimi, B., Kirk, D., Kitching, T. D., Leonard, A., Mandelbaum, R., Schaefer, B. M., Sifon, C., Brown, M. L., and et al., "Galaxy alignments: Theory, modelling & simulations," Space Science Reviews, vol. 193, p. 67–136, 2015, doi:10.1007/s11214-015-0203-6.

[181] Blazek, J. A., MacCrann, N., Troxel, M., and Fang, X., "Beyond linear galaxy alignments," Physical Review D, vol. 100, 2019, doi:10.1103/physrevd.100.103506.

[182] Szewciw, A. O., Beltz-Mohrmann, G. D., Berlind, A. A., and Sinha, M., "Toward accurate modeling of galaxy clustering on small scales: Constraining the galaxy-halo connection with optimal statistics," 2021.

[183] Voivodic, R. and Barreira, A., "Responses of halo occupation distributions: a new ingredient in the halo model & the impact on galaxy bias," Journal of Cosmology and Astroparticle Physics, vol. 2021, p. 069, 2021, doi:10.1088/1475-7516/2021/05/069.

[184] Friedrich, O., Halder, A., Boyle, A., Uhlemann, C., Britt, D., Codis, S., Gruen, D., and Hahn, C., "The pdf perspective on the tracer-matter connection: Lagrangian bias and non-poissonian shot noise," 2021.

[185] Sugiyama, S., Takada, M., Miyatake, H., Nishimichi, T., Shirasaki, M., Kobayashi, Y., More, S., Takahashi, R., Osato, K., Oguri, M., Coupon, J., Hikage, C., Hsieh, B.-C., Komiyama, Y., Leauthaud, A., Li, X., Luo, W., Lupton, R. H., Murayama, H., Nishizawa, A. J., Park, Y., Price, P. A., Simet, M., Speagle, J. S., Strauss, M. A., and Tanaka, M., "HSC Year 1 cosmology results with the minimal bias method: HSC×BOSS galaxy-galaxy weak lensing and boss galaxy clustering," 2021.

[186] Chisari, N. E., Richardson, M. L. A., Devriendt, J., Dubois, Y., Schneider, A., Le Brun, A. M. C., Beckmann, R. S., Peirani, S., Slyz, A., and Pichon, C., "The impact of baryons on the matter power spectrum from the Horizon-AGN cosmological hydrodynamical simulation," Monthly Notices of the Royal Astronomical Society, vol. 480, no. 3, pp. 3962–3977, 2018, doi:10.1093/mnras/sty2093.

[187] Siegmund, D., "Importance Sampling in the Monte Carlo study of sequential tests," The Annals of Statistics, vol. 4, no. 4, p. 673–684, 1976.

[188] Hesterberg, T. C., "Advances in Importance Sampling," 2003.

[189] Huterer, D. and Turner, M. S., "Probing dark energy: Methods and strategies," Physical Review D, vol. 64, 2001, doi:10.1103/physrevd.64.123527.

[190] Lemos, P., Raveri, M., Campos, A., Park, Y., Chang, C., Weaverdyck, N., Huterer, D., Liddle, A. R., Blazek, J., Cawthon, R., and et al., "Assessing tension metrics with dark energy survey and planck data," Monthly Notices of the Royal Astronomical Society, vol. 505, p. 6179–6194, 2021, doi:10.1093/mnras/stab1670.

[191] Raveri, M. and Hu, W., "Concordance and discordance in cosmology," Physical Review D, vol. 99, 2019, doi:10.1103/physrevd.99.043506.

[192] Anchordoqui, L. A., Di Valentino, E., Pan, S., and Yang, W., "Dissecting the h_0 and s_8 tensions with Planck + BAO + supernova type Ia in multi-parameter cosmologies," Journal of High Energy Astrophysics, vol. 32, p. 28–64, 2021, doi:10.1016/j.jheap.2021.08.001.

[193] Rozo, E., Rykoff, E. S., Abate, A., Bonnett, C., Crocce, M., Davis, C., Hoyle, B., Leistedt, B., Peiris, H. V., Wechsler, R. H., and et al., "redMaGiC: selecting luminous red galaxies from the DES Science Verification data," Monthly Notices of the Royal Astronomical Society, vol. 461, p. 1431–1450, 2016, doi:10.1093/mnras/stw1281.

[194] DeRose, J., Wechsler, R. H., Becker, M. R., Busha, M. T., Rykoff, E. S., MacCrann, N., Erickson, B., Evrard, A. E., Kravtsov, A., Gruen, D., Allam, S., Avila, S., Bridle, S., Brooks, D., Buckley-Geer, E., Rosell, A. C., Kind, M. C., Carretero, J., Castander, F. J., Cawthon, R., Crocce, M., da Costa, L. N., Davis, C., Vicente, J. D., Dietrich, J. P., Doel, P., Drlica-Wagner, A., Fosalba, P., Frieman, J., Garcia-Bellido, J., Gutierrez, G., Hartley, W. G., Hollowood, D. L., Hoyle, B., James, D. J., Krause, E., Kuehn, K., Kuropatkin, N., Lima, M., Maia, M. A. G., Menanteau, F., Miller, C. J., Miquel, R., Ogando, R. L. C., Malagón, A. P., Romer, A. K., Sanchez, E., Schindler, R., Serrano, S., Sevilla-Noarbe, I., Smith, M., Suchyta, E., Swanson, M. E. C., Tarle, G., and Vikram, V., "The buzzard flock: Dark energy survey synthetic sky catalogs," 2019.

[195] Wechsler, R. H., DeRose, J., Busha, M. T., Becker, M. R., Rykoff, E., and Evrard, A., "ADDGALS: Simulated Sky Catalogs for Wide Field Galaxy Surveys," 2021.

[196] DeRose, J., Wechsler, R. H., Becker, M. R., Rykoff, E. S., Pandey, S., MacCrann, N., Amon, A., Myles, J., Krause, E., Gruen, D., Jain, B., Troxel, M. A., Prat, J., Alarcon, A., Sánchez, C., Blazek, J., Crocce, M., Giannini, G., Gatti, M., Bernstein, G. M., Zuntz, J., Dodelson, S., Fang, X., Friedrich, O., Secco, L. F., Elvin-Poole, J., Everett, S., Choi, A., Harrison, I., Cordero, J., Rodriguez-Monroy, M., McCullough, J., Cawthon, R., Chen, A., Alves, O., Camacho, H., Campos, A., Diehl, H. T., Drlica-Wagner, A., Eifler, T. F., Fosalba, P., Huang, H., Porredon, A., Raveri, M., Rosenfeld, R., Ross, A. J., Sanchez, J., Sheldon, E., Yanny, B., Yin, B., Aguena, M., Allam, S., Andrade-Oliveira, F., Annis, J., Avila, S., Bacon, D., Bechtol, K., Bhargava, S., Brooks, D., Buckley-Geer, E., Burke, D. L., Carnero Rosell, A., Carrasco Kind, M., Chang, C., Costanzi, M., da Costa, L. N., Pereira, M. E. S., De Vicente, J., Desai, S., Dietrich, J. P., Doel, P., Eckert, K., Evrard, A. E., Ferrero, I., Ferté, A., Flaugher, B., Frieman, J., García-Bellido, J., Gaztanaga, E., Giannantonio, T., Gruendl, R. A., Gschwend, J., Gutierrez, G., Hartley, W. G., Hinton, S. R., Hollowood, D. L., Honscheid, K., Huff, E. M., Huterer, D., James, D. J., Kuehn, K., Kuropatkin, N., Lahav, O., Lima, M.,

Maia, M. A. G., Marshall, J. L., Melchior, P., Menanteau, F., Miquel, R., Mohr, J. J., Morgan, R., Palmese, A., Paz-Chinchón, F., Pieres, A., Plazas Malagón, A. A., Sanchez, E., Scarpine, V., Serrano, S., Sevilla-Noarbe, I., Smith, M., Soares-Santos, M., Suchyta, E., Tarle, G., Thomas, D., To, C., Varga, T. N., and Zhang, Y., "Dark Energy Survey Year 3 results: cosmology from combined galaxy clustering and lensing – validation on cosmological simulations," arXiv e-prints, p. arXiv:2105.13547, 2021.

[197] Sobol, I., "On the distribution of points in a cube and the approximate evaluation of integrals," USSR Computational Mathematics and Mathematical Physics, vol. 7, no. 4, pp. 86–112, 1967, doi:10.1016/0041-5553(67)90144-9.

[198] To, C.-H., Rozo, E., Krause, E., Wu, H.-Y., Wechsler, R. H., and Salcedo, A. N., "LINNA: Likelihood Inference Neural Network Accelerator," 2022.

[199] Bratley, P. and Fox, B. L., "Algorithm 659: Implementing Sobol's quasirandom sequence generator," ACM Trans. Math. Softw., vol. 14, p. 88–100, 1988, doi:10.1145/42288.214372.

Annexed A

Implementing the Sobol's quasi-random sampling

To sample the ith point of a Sobol sequence, we need to choose a polynomial of degree n_i in the space of $\mathbb{Z}_2$: $x^{n_i} + a_{1,i}x^{n_i-1} + ... + a_{n_i-1,i}x + 1$, where the coefficients $a_{1,...,n_i-1,i}$ are either 0 or 1 [199].

We define the recurrence relation:

$$m_{j,i} = 2a_{1,i}m_{j-1,i} \oplus 2^2 a_{2,i}m_{j-2,i} \oplus ... \oplus 2^{n_i-1}a_{n_i-1,i}m_{j-(n_i-1),i} \oplus 2^{n_i}m_{j-n_i,i} \oplus m_{j-n_i,i}, \quad \text{(A.1)}$$

where the initial values $m_{1,i}, ..., m_{n_i,i}$ are chosen such that each $m_{j,i}$, $1 \leq j \leq n_i$, is odd and less than 2^j and $\oplus$ is the bit-by-bit exclusive-or operator. A consequence of this is that $m_{j,i}$ will be positive integers. The direction numbers $v_{1,i}, v_{2,i}, ...$ are defined by:

$$v_{j,i} = \frac{m_{j,i}}{2^j}. \quad \text{(A.2)}$$

The jth component of the kth point in a Sobol sequence is given by:

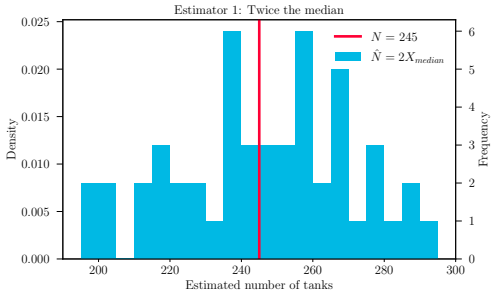
$$v_{j,k} = i_1 v_{1,k} \oplus i_2 v_{2,k} \oplus ... \quad \text{(A.3)}$$

where i_j is the jth digit from the right when i is written in binary $i = (...i_3 i_2 i_1)_2$.

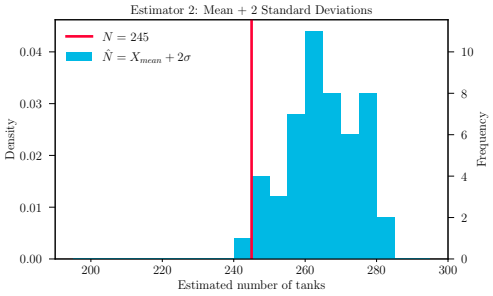
Annexed B

Extensions of the tank problem

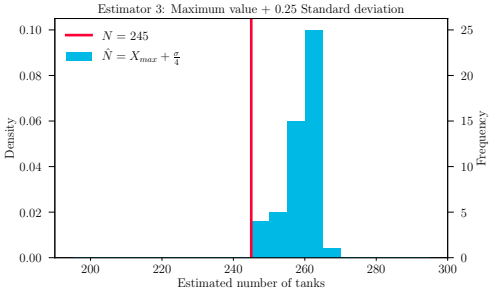
B.1. Cases with a larger sample



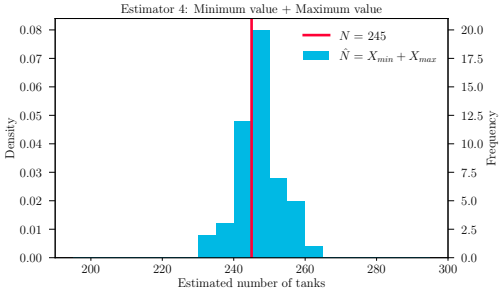
(a) Twice the median



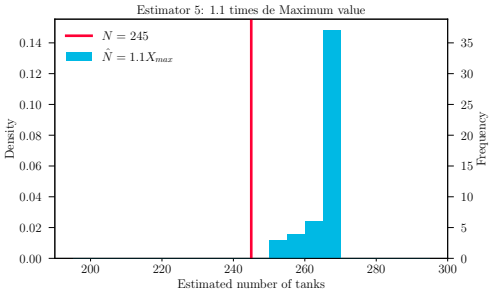
(b) Mean + 2 Standard Deviations



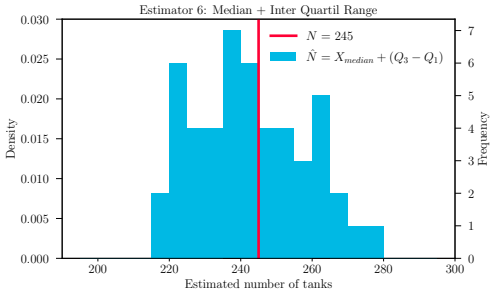
(c) Maximum value + 0.25 Standard deviation



(d) Maximum value + Minimum value



(e) 1.1 times de Maximum value



(f) Median + Inter Quartil Range

Figure B.1: Histograms of the frequentist estimators $\hat{N}$ for $N = 245$ and 50 random sampled serial numbers of 50 observations, doubling the serial numbers per observation w.r.t the case presented in Fig. 6.1.

If we increase the size of the sample from to 50 in each of the observations from the frequentist approach, the described behaviors from Sec. 6.2 hold as shown in Fig. B.1.

For the unbiased estimator (Eq. 6.6), as we increase the observed numbers, $(n+1)/n \sim 1$, and therefore, $\hat{N} \sim X_{max} - 1$. We would also have a decrease of the variance $\mathbb{V}(\hat{N}) \sim 0$, independently of N. We corroborate that performance for the case in which we double the samples from 25 to 50 in Fig. B.2.

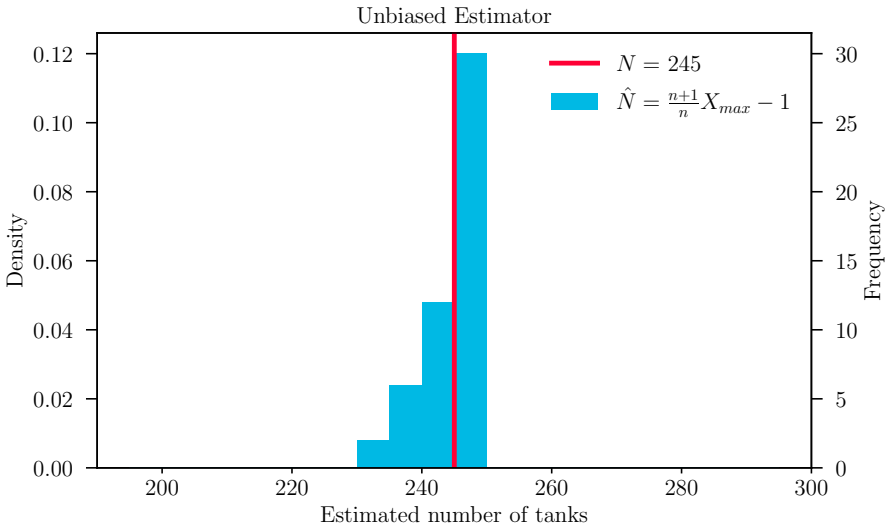


Figure B.2: Histogram of the unbiased frequentist estimator from Eq. 6.6 for the same quantities as the listed in Fig. B.1.

B.2. Other prior distributions for the Bayesian approach

It is also possible to derive the posterior for the case of a geometric prior:

$$P(N) = (1-p)^N p, \tag{B.1}$$

where p accounts for the probability of success in each experiment. This would give us the following posterior:

$$\mathbb{P}(N|x) = \frac{\binom{x}{n}(1-p)^{N-x}}{\binom{N}{n}{}_2F_1[1, 1+x-n; 1+x; 1-p]} \text{ if } N = x, x+1, x+2, ..., \text{ and 0 otherwise.} \tag{B.2}$$

The main challenge with this is that we need to assume even more information: a value for p, which is not clear and will affect the estimates on N. The same problem arises when the negative binomial prior distribution is considered:

$$P(N) = \binom{N+r-1}{r-1}(1-p)^N p^r, \tag{B.3}$$

where we would have another quantity, r which represents the number of failures until the experiment is stopped. This would give us the following posterior:

$$\mathbb{P}(N|x) = \frac{\binom{N+r-1}{r-1}\binom{x}{n}(1-p)^{N-x}}{\binom{N}{n}\binom{N+r-1}{r-1}{}_3F_2[1, 1+x-n, x+r; 1+x, 1+x; 1-p]} \text{ if } N = x, x+1, x+2, ... \tag{B.4}$$

and 0 otherwise.

The respective expected values will, therefore, depend on more information which is not evident how to summary.

Annexed C

Summary of the DSS and its Stochasticity models

The spatial distribution of galaxies and matter can be described by their respective density contrast fields

$$\delta(x) = \frac{\rho(x) - \overline{\rho}}{\overline{\rho}}, \tag{C.1}$$

where $\rho(x)$ is the density at position x and $\overline{\rho}$ is the mean density, of either matter or galaxies, indicated in the following by subscripts m and g, respectively.

The distribution of matter δ_m is not directly observed, but it is traced by galaxies δ_g distributed in the sky. On large scales, the relation between galaxy density contrast and matter density contrast can be approximated as a constant multiplicative factor b between the two,

$$\delta_g \approx b\delta_m, \tag{C.2}$$

which is called linear galaxy bias.

We start by considering a certain distribution of galaxies in a fixed range around a redshift z_f, that we will call *foreground galaxy population*. The 2D density field of the position of galaxies is obtained by applying a circular top-hat filter: at redshift z_f, we would count how many of the galaxies relies inside of it. With this information, the sky can be divided in quintiles of spatial number density. Next, we add another set of galaxies, but extended and located at a higher redshift z_b ($z_f < z_b$) that we will call *background galaxy population*.

The light emitted by the background galaxies will be subject to gravitational shear by passing through the tidal gravitional field of the foreground. The effect will be stronger in zones where the density of foreground galaxies is higher. To evaluate this, we measure the tangential shear that the background sources will suffer in each quintile of foreground density. The observed lensed galaxies will trace the distribution and number density of the source galaxies. This information will be synthesised in a data vector of both the distribution of foreground galaxies and the lensing signals.

If galaxy counts and the matter density were perfectly correlated, then a split of the sky by

galaxy density would be identical to a split by the matter density. In a realistic scenario, however, the connection of these maps would be subjected to astrophysical and systematic effects. The one we focus on is the shot-noise of galaxies, which appears at the two point statistics and as a probe of higher-order statistics. Adding this effect to the linear bias model from Eq. C.2 constitute a powerful probe of cosmology, because it accounts for the skewness of the density field and the connection of galaxies and matter.

It is often assumed that the distribution of galaxy counts N_g traces a smooth field δ_g following a Poisson shot-noise:

$$P(N_g|\delta_m) = \text{Poisson}[N_g, (1 + b\delta_m)\overline{N}_g], \tag{C.3}$$

where $\overline{N}_g$ is the mean count of galaxies in a random such region.

This means that the variance of N_g for fixed δ_g satisfies

$$\frac{\text{Var}[N_g|\delta_g]}{\langle N_g|\delta_g\rangle} = 1. \tag{C.4}$$

When assuming a deterministic relationship between δ_m and δ_g, a consequence is that the variance of N_g for fixed δ_m is also

$$\frac{\text{Var}[N_g|\delta_m]}{\langle N_g|\delta_m\rangle} = 1. \tag{C.5}$$

In the DSS, a nonlinear biasing relationship between δ_g and δ_m is considered by adding degrees of freedom to Eq. C.5 to account for the effect of stochasticity. This quantity is one of the main ingredients of the DSS method, and in this context it could arise from either a dependence on higher powers of δ_m, physical random processes in galaxy formation, or effects from halo exclusion.

C.1. α Model: Parametric model for non-Poissonianity

As a way of generalizing the galaxy-matter connection, it is possible to add two stochasticity parameters, α_0 and α_1, to obtain an equation for a generalized Poisson distribution,

$$\frac{\text{Var}[N_g|\delta_m]}{\langle N_g|\delta_m\rangle} = \alpha_0 + \alpha_1\delta_m, \tag{C.6}$$

in which, for $\alpha_0 = 1.0$ and $\alpha_1 = 0.0$, we recover Eq. C.5.

The original prior distributions for the α parameters assigned in the original work from [1, 2] are described at Table 7.1. The only actual constraint in stochasticity is that $\alpha_0 > 0$, but the boundary $0.1 < \alpha_0$ was motivated with numerical arguments. See appendix E from [2] for more details on the derivation for the other boundaries.

C.2. r Model: correlation $r \neq 1$etween galaxy density and matter density

The second approach used in DSS is the introduction of a free parameter to the baseline linear bias model: a correlation coefficient $r \neq 1$ between the random fields δ_g and δ_m (described in section IV. of [2]). This new degree of freedom captures a possible stochasticity. For $r = 1$ the random fields δ_g and δ_m would be perfectly correlated.

The covariance of δ_g and δ_m can be parametrized by the correlation coefficient:

$$r = \frac{\langle \delta_g \delta_m \rangle}{\sqrt{\langle \delta_g^2 \rangle \langle \delta_m^2 \rangle}}. \tag{C.7}$$

This can lead to a δ_m-dependence of the ratio in Eq. C.5 and in the general model for the galaxy count:

$$\mathrm{Var}[N_g] = \overline{N_g} + \overline{N_g}^2 b^2 \mathrm{Var}[\delta_m], \tag{C.8}$$

$$\mathrm{Cov}[N_g, \delta_m] = \overline{N_g} b r \mathrm{Var}[\delta_m]. \tag{C.9}$$

Theoretically, the possible values that r could take are $-1 \leq r \leq 1$. The given theory predicts a positive correlation value, so the original priors for r followed a uniform distribution $\Pi(r) \sim \mathcal{U}[0.0,1.0]$.