DISS. ETH NO. 30102

FORWARD MODELLING AND THEORETICAL PREDICTIONS FOR COSMOLOGICAL GALAXY SURVEYS

A thesis submitted to attain the degree of

DOCTOR OF SCIENCES
(Dr. sc. ETH Zurich)

presented by

BEATRICE MOSER

Master of Science ETH in Physics
ETH Zurich

born on 30th of October 1994

accepted on the recommendation of
Prof. Dr. Alexandre Refregier
Dr. Tomasz Kacprzak
Prof. Dr. Daniel Grün

2024

So Long, and Thanks for All the Fish

— DOUGLAS ADAMS

ABSTRACT

Cosmological galaxy surveys measure the positions and fluxes of millions of galaxies, tracing the underlying distribution of matter in the Universe. The formation and evolution of cosmic structures are successfully described through the interplay between gravity and the accelerated expansion of the Universe in the ΛCDM model. The physical nature of the two dominant components of the Universe in this model, dark matter and dark energy, is not yet understood and some discrepancies have recently arisen between constraints from different observables. Large-scale structure surveys can shed light on the dark components, through the combination of galaxy clustering, cosmic shear and galaxy-galaxy lensing. The volume of cosmological data has increased dramatically in the past two decades, and will continue to do so in the near future with the advent of Stage IV cosmological surveys. The quantity and quality of data requires a continuous effort to make theoretical predictions more robust and easy to extend, while also carefully controlling observational systematics.

In this thesis, we explore how novel techniques can be developed to realize the potential of next generation cosmological galaxy surveys. We start by introducing a framework for the numerical solution of the Einstein-Boltzmann equations, describing the time evolution of linear order perturbations of cosmological fields. The key innovation lies in the automatic translation of symbolic equations into optimized code for their solution. This enables the timely implementation and testing of new models, with minimal use of approximations. We test the framework in three cases: for a dark energy model with constant equation of state, for massive neutrinos and for the radiation streaming approximation.

We then turn to observational systematics affecting cosmological galaxy surveys. One major source of systematic uncertainty is the determination of the redshift distributions of galaxy samples from a limited number of broad-band fluxes, known as photometric redshift estimation. Inaccuracies in the estimated redshift distribution can bias the resulting cosmological constraints from weak lensing or photometric galaxy clustering.

To address this problem, we use a forward modelling approach, part of the Monte Carlo Control Loops (MCCL) framework. The redshift distribution of galaxies is a result of their observed properties, which we link to an underlying galaxy population model and simulate at the image level including realistic observational and instrumental effects. We derive posteriors of parameters from the empirical galaxy population model via simulation-based inference. The resulting image simulations can be processed in the same way as the data, thus accounting for sample selection, and yield realistic redshift distributions. We apply this methodology to deep field images from the Hyper Suprime-Cam Subaru Strategic Program (HSC) and photometric redshifts from COSMOS2020 to constrain the galaxy population model at high redshift and test our method in the regime of upcoming surveys.

We then perform a blind tomographic comparison between our forward modelling pipeline and Self-Organizing Map p(z) (SOMPZ), a direct redshift calibration technique, applied on the same target sample from Dark Energy Survey (DES) Year 3 data. The cross-comparison and combination of different methodologies will become an essential tool to achieve the level of accuracy required for Stage IV surveys.

This work shows how advanced numerical methods and forward modelling offer good prospects for the analysis of next generation cosmological galaxy surveys.

SOMMARIO

Le osservazioni di galassie per la cosmologia misurano le posizioni e i flussi di milioni di galassie, che ricalcano la distribuzione della materia nell'Universo. La formazione e l'evoluzione delle strutture cosmiche sono descritte con successo dall'interazione tra la gravità e l'espansione accelerata dell'Universo nel modello ΛCDM. La natura fisica delle due componenti dominanti dell'Universo in questo modello, la materia oscura e l'energia oscura, non è ancora stata compresa e di recente sono emerse alcune differenze tra i parametri misurati da diverse osservabili. Le indagini della struttura a grande scala possono far luce sulle componenti oscure, attraverso la combinazione del clustering delle galassie, la correlazione tra lenti gravitazionali deboli e il lensing galassia-galassia. Il volume dei dati cosmologici è aumentato in modo drammatico negli ultimi due decenni e continuerà a farlo nel prossimo futuro con l'avvento delle survey cosmologiche di fase IV. La quantità e qualità dei dati richiedono uno sforzo continuo per rendere le previsioni teoriche più affidabili e facili da estendere, tenendo sotto controllo allo stesso tempo gli errori sistematici osservativi.

In questa tesi esploriamo nuove tecniche per realizzare il potenziale delle osservazioni cosmologiche di galassie di prossima generazione. Iniziamo introducendo un framework per la soluzione numerica delle equazioni di Einstein-Boltzmann, che descrivono l'evoluzione temporale delle perturbazioni dei campi cosmologici al primo ordine. L'innovazione chiave consiste nella traduzione automatica delle equazioni simboliche in codice ottimizzato per la loro soluzione. Ciò consente di implementare e testare tempestivamente nuovi modelli, con un uso minimo di approssimazioni. Abbiamo testato il framework in tre casi: per un modello di energia oscura con equazione di stato costante, per i neutrini massivi e per l'approssimazione di flusso della

radiazione.

Passiamo poi al trattamento delle sistematiche osservative che interessano le osservazioni cosmologiche di galassie. Una delle principali fonti di incertezza sistematica è la determinazione delle distribuzioni di redshift dei campioni di galassie a partire da un numero limitato di flussi misurati in filtri a banda larga, nota come stima del redshift fotometrico. Errori nella ricostruzione della distribuzione del redshift possono falsare i vincoli cosmologici risultanti dal lensing debole o dal clustering fotometrico delle galassie.

Per affrontare questo problema, utilizziamo un approccio detto forward modelling, parte del framework Monte Carlo Control Loops (MCCL). La distribuzione del redshift delle galassie è il risultato delle loro proprietà osservate, che possiamo collegare a un modello di popolazione di galassie e simulare a livello di immagini includendo effetti osservativi e strumentali realistici. Ricaviamo la distribuzione dei parametri del modello empirico di popolazione di galassie mediante inferenza basata sulle simulazioni. Una volta calibrate, le immagini simulate possono essere analizzate allo stesso modo dei dati reali, tenendo conto della selezione del campione, e riproducono distribuzioni di redshift realistiche. Applichiamo questa metodologia alle immagini di campo profondo dell'Hyper Suprime-Cam Subaru Strategic Program (HSC) e ai redshift fotometrici di COSMOS2020 per vincolare il modello di popolazione delle galassie ad alto redshift e testare il nostro metodo nel regime delle survey future.

Eseguiamo poi un confronto tomografico in cieco tra la nostra pipeline che usa il forward modelling e Self-Organizing Map p(z) (SOMPZ), una tecnica di calibrazione diretta del redshift, applicata allo stesso campione di galassie misurate da tre anni di osservazioni della Dark Energy Survey (DES). Il confronto e la combinazione di diverse metodologie diventerà uno strumento essenziale per raggiungere il livello di accuratezza richiesto per le survey di fase IV.

Questo lavoro mostra come i metodi numerici avanzati e il forward modelling offrano buone prospettive per l'analisi delle survey cosmologiche di galassie di prossima generazione.

CONTENTS

CHAPTER 1

INTRODUCTION

> *The story so far:*
> *In the beginning the Universe was created.*
> *This has made a lot of people very angry and been*
> *widely regarded as a bad move.*
>
> — DOUGLAS ADAMS, The Restaurant at the End of the Universe

Cosmology is the study of the Universe as a whole, including its origin, contents and evolution through time. Modern cosmology has been developed in the last century. The theoretical foundation is Einstein's theory of general relativity, describing the geometry of spacetime and how the energy and momentum of the contents of the Universe shape it through gravity. The ΛCDM model was enstablished as the concordance model of cosmology roughly three decades ago, when experiments started confirming the theoretical predictions and constraining the parameters of the model (e.g. [1]). In this model, the Universe contains three main components: baryonic matter, which is the ordinary matter we are accustomed to, cold dark matter, which is a more abundant matter component only interacting gravitationally and dark energy, the source of energy driving the accelerated expansion of the Universe and described in ΛCDM by the cosmological constant Λ. Both the latter components are of yet unknown nature and are deduced through indirect observations. A brief overview of the cosmological background and of the perturbations of the various components of the Universe is presented in the next sections of this chapter.

Observational successes of the ΛCDM model include the detection of the black body spectrum and anisotropies of the cosmic microwave background (CMB) [2], the accelerated expansion of the Universe observed from the Hubble diagram of Supernovae Type Ia [3, 4] among other probes (see e.g. [5]), the structure of the cosmic web [6, 7] and the abundance of elements in the Universe [8]. As measurements get more precise, some cracks have started appearing in the model. First the H_0 tension [9], a disagreement between early-time and late-time cosmological probes concerning the rate of expansion of the Universe and then the S_8 tension [10], a difference in the measurements of the clumpiness of matter.

A popular way to investigate the tensions is that of increasing the constraining power on the parameters of the model, by either using summary statistics that capture more information from the data, novel probes or cross-correlations between different probes and new datasets. Instead, in this thesis we explore some of the steps that occur between the collection of data and cosmological constraints in a cosmological galaxy survey. This is necessary to test the robustness of the constraints and whether the cosmological tensions point to new physics, rather than limitations of our measurements. First, we look at the theoretical predictions that we compare the data to, which need to be accurate and include beyond ΛCDM model extensions. Then, we extend a novel technique to treat systematics that could affect large-scale structure measurements.

At the core of many cosmological theory predictions lies the Einstein-Boltzmann system of equations, that describes the evolution of perturbations in the Universe at linear order [11]. In Chapter 2, we contribute to the development of the `PyCosmo` Python-based symbolic framework [12], which allows to solve the Einstein-Boltzmann equations numerically and can be easily extended beyond the ΛCDM model. We demonstrate the extensibility of the framework by implementing two simple model extensions: a dark energy component with a constant equation of state and massive neutrinos. Allowing for a dark energy component that is not a cosmological constant is a very common theory extension and allows for a test of the dark sector [13, 14]. Neutrinos are known to be massive from oscillation experiments [15], so that accounting for neutrino mass in cosmological predictions has become important at the current level of precision. Furthermore, cosmological observations are a promising ground for distinguishing between the neutrino mass orderings and eventually detecting the neutrino mass [16]. We also extend the `PyCosmo` package to include the radiation streaming approximation [17], providing a considerable speed-up of the code with little accuracy loss. The main observable we focus on in this chapter is the linear matter power spectrum, which appears in most large-scale structure theoretical predictions.

We transition from theory to observations of large-scale structure in Chapters 3 and 4. One of the main systematics impacting large-scale structure observables and cosmic shear measurements in particular is the retrieval of redshift distributions of samples of galaxies from noisy photometry. Current cosmological surveys, such as the Dark Energy Survey (DES) [18], the Kilo Degree Survey (KiDS) [19] and Hyper Suprime Cam Subaru Strategic Survey (HSC) [20] measure the flux from galaxies in less than ten broad-band filters between the optical and the near-infrared. Upcoming Stage IV surveys, such as the Large Survey of Space and Time (LSST) [21], Euclid [22, 23] and Nancy-Grace Roman Space Telescope [24], are pursuing a similar approach with increased statistics and will thus need an improved control of systematics. Measuring photometric redshifts is challenging, prone to known colour-redshift degeneracies and errors in retrieved redshift distribution lead to biases in the resulting cosmological constraints [25, 26].

We apply and extend the Monte Carlo Control Loops (MCCL) [27] forward modelling framework for photometric redshift calibration. The method relies on a parametric galaxy population model which is constrained by observations via simulation-based inference. The full image simulations also include observational and instrumental effects, so that the real and simulated data can be processed the same way and realistic redshift distributions can be extracted from the simulations.

In Chapter 3, we constrain the galaxy population model parameters using HSC deep data [28] and accurate photometric redshifts from COSMOS2020 [29]. This yields simulations that reproduce the photometric properties of the data well and redshift distributions that we compare with COSMOS2020 at different magnitude cuts. In Chapter 4, we perform a comparison between two different redshift calibration methodologies applied to the same target sample from DES Year 3 [30]. We compare the MCCL forward modelling approach with SOMPZ, a technique which applies a reweighting of reference redshifts through Self Organizing Maps (SOMs) and is the fiducial redshift calibration technique of the DES collaboration.

We summarize the conclusions of this thesis in Chapter 5. Chapters 2 and 3 appear in a similar form in the publications [31, 32]. The references to publications are stated at the beginning of the chapters.

1.1 The homogeneous Universe

The theoretical framework of physical cosmology is Einstein's theory of general relativity (GR), which postulates a deep interconnection between the structure of spacetime and its energy content. In this introduction, we give a brief overview of the ingredients of the standard model of cosmology, following mainly [33]. More extensive treatments can be found in [33–39].

One key observation that greatly simplifies our description of the Universe is the fact that, on sufficiently large scales, there seem to be no preferred position and spatial direction: there is no special place in the Universe. More mathematically, the Universe is statistically homogeneous and isotropic. This base assumption, known as the *cosmological principle*, imposes a unique and simple form for the geometry of spacetime on large scales. The relation between the coordinate-independent line element ds and the observer dependent coordinate system is given by the spacetime metric. In the case of an homogeneous and isotropic Universe, this reads

$$ds^2 = g_{\mu\nu} dx^\mu dx^\nu = -c^2 dt^2 + a^2(t) \left[d\chi^2 + S_k^2(\chi) d\Omega^2 \right] \tag{1.1}$$

where we use the Einstein summation convention with $\mu, \nu \in 0, ...3$. In Equation 1.1 c is the speed of light, $a(t)$ is the scale factor relating comoving and physical spatial coordinates, $d\Omega^2 = d\theta^2 + \sin^2\theta d\phi^2$ is the solid angle in polar coordinates and k relates to the spatial curvature of the Universe. $S_k(\chi)$ parametrizes the three possible spatial geometries:

$$S_k(\chi) = R_0 \begin{cases} \frac{1}{\sqrt{|k|}} \sinh\left(\frac{\sqrt{|k|}\chi}{R_0}\right), & k = -1 \text{ (open universe)} \\ \frac{\chi}{R_0}, & k = 0 \text{ (flat universe)} \\ \frac{1}{\sqrt{k}} \sin\left(\frac{\sqrt{k}\chi}{R_0}\right), & k = 1 \text{ (closed universe)} \end{cases} \tag{1.2}$$

where R_0 is the curvature scaling. If we use the rescaling symmetry of the metric to set the value of the scale factor today (t=t_0) such that $a(t_0) = 1$, then $R_0 = \frac{c}{H_0}$ is the physical curvature scale today. This metric is known as the Friedmann-Lemaître-Robertson-Walker (FLRW) metric [40–43].

In general relativity free falling particles follow geodesics, the shortest path between two points on a curved spacetime surface. The trajectories $x^\mu(\lambda)$ are determined by the geodesic equation

$$\frac{d^2x^\mu}{d\lambda^2} = -\Gamma^\mu_{\alpha\beta}\frac{dx^\alpha}{d\lambda}\frac{dx^\beta}{d\lambda} \tag{1.3}$$

where the Christoffel symbols $\Gamma^\mu_{\alpha\beta}$ are defined as

$$\Gamma^\mu_{\alpha\beta} = \frac{g^{\mu\nu}}{2}\left[\frac{\partial g_{\alpha\nu}}{\partial x^\beta} + \frac{\partial g_{\beta\nu}}{\partial x^\alpha} - \frac{\partial g_{\alpha\beta}}{\partial x^\nu}\right]. \tag{1.4}$$

The Einstein field equations describe the dynamics of the Universe by relating the spacetime curvature encoded in the Einstein tensor

$$G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R \tag{1.5}$$

where the Ricci tensor and scalar are defined as

$$R_{\mu\nu} = \frac{\partial\Gamma^\rho_{\mu\nu}}{\partial x^\rho} - \frac{\partial\Gamma^\rho_{\mu\rho}}{\partial x^\nu} + \Gamma^\rho_{\sigma\rho}\Gamma^\sigma_{\mu\nu} - \Gamma^\rho_{\nu\sigma}\Gamma^\sigma_{\mu\rho}, \quad R = g^{\mu\nu}R_{\mu\nu} \tag{1.6}$$

with the contents of the Universe represented by the energy momentum tensor $T_{\mu\nu}$:

$$G_{\mu\nu} + \Lambda g_{\mu\nu} = \frac{8\pi G}{c^4} T_{\mu\nu} \tag{1.7}$$

where G is Newton's gravitational constant and Λ is the cosmological constant, introduced by Einstein to allow for a static Universe solution of the equations. If we assume that we can describe the components of the Universe as perfect fluids, then the energy-momentum tensor of each component can be written as

$$T_{\mu\nu} = \left(\rho + \frac{P}{c^2}\right) U_\mu U_\nu + P g_{\mu\nu} \tag{1.8}$$

where ρ is the density, P the pressure and $U^\mu = c\frac{dx^\mu}{\sqrt{-ds^2}}$ the four velocity of the perfect fluid. The Universe contains both relativistic components, like photons and neutrinos, and non-relativistic components such as baryonic matter. Furthermore, since the 1930s [44] various observations have pointed to the existence of an additional form of matter, dark matter, which only interacts gravitationally and makes up the majority of the matter in the Universe.

By inserting the FLRW metric in the Einstein field equations for an ideal fluid, we obtain the Friedmann equations

$$H^2(t) = \left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho - \frac{kc^2}{a^2 R_0^2} \tag{1.9}$$

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}\left(\rho + \frac{3P}{c^2}\right) \tag{1.10}$$

describing the time evolution of the scale factor $a(t)$. The function $H(t)$ is called the Hubble parameter and plays an important role in cosmology. The time evolution of pressure and density is derived from the conservation of the stress-energy tensor

$$\nabla_\mu T^{\mu\nu} = 0 \tag{1.11}$$

together with the equation of state

$$P = w\rho c^2 \tag{1.12}$$

with $w = 0$ for dark matter and baryons and $1/3$ for radiation components. The combination of Equation 1.11 and 1.12 gives us

$$\frac{d\rho}{dt} + 3\frac{\dot{a}}{a}\rho(1+w) = 0 \quad \Rightarrow \quad \rho(a) \propto a^{-3(1+w)} \tag{1.13}$$

and thus

$$\rho_m(a) \propto a^{-3} \qquad\qquad \rho_r(a) \propto a^{-4} \tag{1.14}$$

where $\rho_m(a)$ is the density of matter, either dark or baryonic, and $\rho_r(a)$ that of radiation.

Edwin Hubble in 1929 observed that there is a positive linear relationship between the recession velocity of a galaxy and its distance from the observer [45]. This means that the value of the Hubble parameter today $H_0 \equiv H(t = t_0)$ is positive and consequently that the Universe is expanding. Tracing back the expansion history leads to the prediction that the Universe was initially in a hot dense state called Big Bang and then cooled down during the expansion. During this evolution, different components dominated the energy density of the Universe: radiation was dominant at early times whereas matter started dominating at intermediate time-scales. In the 1990s, the observational evidence from supernovae Type Ia [3, 4] and other probes (see e.g. [5]) showed that the Universe is currently undergoing a phase of accelerated expansion and lead to the addition of a further component to the model: dark energy. This mysterious energy source drives the accelerated expansion of the Universe $\ddot{a} > 0$ and must therefore have an equation of state parameter $w < -1/3$. The simplest case is when $w = -1$ so that ρ_Λ is constant and $T_{\mu\nu}$ proportional to $g_{\mu\nu}$ which corresponds to the cosmological constant Λ introduced in Equation 1.7.

We can reformulate the Friedmann equations 1.9 and 1.10, by defining the critical density of the Universe

$$\rho_{\text{crit}}(a) = \frac{3H(a)^2}{8\pi G} \tag{1.15}$$

and the density parameters Ω_i

$$\Omega_i = \frac{\rho_i(a)}{\rho_{\text{crit}}(a)}. \tag{1.16}$$

We can use the quantities just defined to write

$$\frac{H(a)^2}{H_0^2} = \Omega_{r,0}a^{-4} + \Omega_{m,0}a^{-3} + \Omega_\Lambda + \Omega_k a^{-2} \tag{1.17}$$

where $\Omega_{r,0}$ includes all radiation components and $\Omega_{m,0}$ all matter components.

$$\Omega_k = -\frac{kc^2}{H_0^2 R_0^2} \tag{1.18}$$

is the curvature parameter, observationally shown to be very close to zero so that

$$\Omega_0 = \Omega_{r,0} + \Omega_\Lambda + \Omega_{m,0} = 1. \tag{1.19}$$

From here on we assume a flat cosmology.

The wavelength of light emitted by a distant object in an expanding Universe is stretched along its path

$$\frac{\lambda_{\text{obs}}}{\lambda_{\text{emit}}} = \frac{a(t_{\text{obs}})}{a(t_{\text{emit}})} \tag{1.20}$$

where λ_{emit} and t_{emit} are the emitted wavelength and time of emission and λ_{obs} and t_{obs} the observed one. The cosmological redshift is then defined as

$$z \equiv \frac{\lambda_{\text{obs}} - \lambda_{\text{emit}}}{\lambda_{\text{emit}}} = \frac{a(t_{\text{obs}})}{a(t_{\text{emit}})} - 1 \tag{1.21}$$

and since we observe the Universe today at $t_{\text{obs}} = t_0$

$$1 + z = \frac{1}{a}. \tag{1.22}$$

By measuring the spectrum of a distant galaxy, we can obtain the scale factor value at the time the light was emitted by comparing the wavelength of emission and absorption lines in the galactic spectrum with that of spectral lines of different elements measured on Earth.

Besides cosmological redshift, there are several other important measures of distance in cosmology. We have already mentioned the comoving distance χ, which defines a spatial coordinate system that does not change over time. Let us look at a photon travelling in space. Due to isotropy, we can always consider its motion as radial only such that

$$ds^2 = 0 = -c^2 dt^2 + a^2(t) d\chi^2. \tag{1.23}$$

We thus obtain

$$d\chi = c\frac{dt}{a(t)} = c d\eta \tag{1.24}$$

where η is called the conformal time. We call the maximum distance that a photon can travel since the Big Bang comoving horizon

$$\chi_H = c\int_0^{t_0} \frac{dt}{a(t)}. \tag{1.25}$$

The comoving horizon χ_H corresponds to the size of the observable Universe.

Distances that are directly related to astronomical observations are the angular diameter distance and the luminosity distance. Consider an object of size D and luminosity L, which we call a standard ruler, and of which we measure the angular size $\Delta\theta$ and the flux F. Then, from the angular part of the metric in Equation 1.1, we obtain the angular diameter distance d_A

$$d_A(z) = \frac{D}{\Delta\theta} = a S_k(\chi). \tag{1.26}$$

The luminosity distance d_L is defined through

$$F = \frac{L}{4\pi d_L^2} \tag{1.27}$$

and, since the solid angle subtended by the object depends on d_A^{-2} and the photon's energy density evolves as a^{-4}, we have

$$d_L = \frac{S_k(\chi)}{a}. \tag{1.28}$$

1.2 The inhomogeneous Universe

Despite the *cosmological principle*, the Universe is only homogeneous and isotropic on sufficiently large scales. When we look at smaller areas of the Universe, inhomogeneities and anisotropies arise: we find filaments, voids and collapsed objects like galaxy clusters, galaxies, stars and planets. These structures are formed by the collapse of initially small inhomogeneities. We thus treat the inhomogeneities as perturbations on top of an otherwise homogeneous and isotropic background.

1.2.1 Linear perturbation theory

We start by adding a perturbation to the FLRW metric $\bar{g}_{\mu\nu}$ so that

$$g_{\mu\nu} = \bar{g}_{\mu\nu} + \delta g_{\mu\nu} \tag{1.29}$$

where $\delta g_{\mu\nu}$ is symmetric and $|\delta g_{\mu\nu}| \ll \bar{g}_{\mu\nu}$. The perturbations in the metric can be separated in a scalar, a vector and a tensor part (SVT decomposition) [46]. This decomposition is important, as the three sets of equations do not mix at linear order. Here we only focus on the scalar perturbations, which are relevant for structure formation in our Universe. Both vector and tensor perturbations produce no density fluctuations and are thus not important for structure formation [47], even though tensor perturbations correspond to primordial gravitational waves potentially detectable in the cosmic microwave background. We follow closely the linear perturbation treatment from [11].

The form of the perturbed metric depends on the choice of gauge, which determines the time slicing of spacetime and the spatial coordinates defined on this time slicing. In the following, we work in Newtonian conformal gauge, which offers simple physical interpretations and in which the perturbed metric is diagonal [11]. Another common gauge is the synchronous gauge.

The line element in Newtonian conformal gauge assuming a flat Universe reads [35]

$$ds^2 = a^2(\eta)\left[-(1+2\Psi(\eta,\vec{x}))d\eta^2 + \left(1+2\Phi(\eta,\vec{x})\right)dx^i dx_i\right] \tag{1.30}$$

where $\Psi(\eta,\vec{x})$ and $\Phi(\eta,\vec{x})$ are the two Newtonian gravitational potentials and we use conformal time η and x^i as the spatial components of x^μ in Cartesian coordinates. In the case of dark and baryonic matter, we can again write the perturbed stress-energy tensor using the perfect fluid assumption

$$T^0{}_0 = -(\bar{\rho}+\delta\rho) \tag{1.31}$$

$$T^i{}_0 = -(\bar{\rho}+\frac{\bar{P}}{c^2})v^i \tag{1.32}$$

$$T^i{}_j = (\bar{P}+\delta P)\delta^i{}_j + \Pi^i{}_j \tag{1.33}$$

where $\delta\rho$ and δP are the density and pressure perturbations, $v^i = \frac{dx^i}{d\eta}$ is the bulk velocity of the fluid and is also a first order perturbation and $\Pi^i{}_j$ is the anisotropic stress ($\Pi^i{}_i = 0$). We define the density contrast of a species i

$$\delta_i(\eta,\vec{x}) = \frac{\rho_i(\eta,\vec{x}) - \bar{\rho}_i(\eta,\vec{x})}{\bar{\rho}_i(\eta,\vec{x})}. \tag{1.34}$$

Photons and massless neutrinos can only be appropriately described through their phase space distribution function. In general, the stress-energy tensor $T_{\mu\nu}$ can be written in terms of the distribution function of a particle species $f(x^i, P_j, t)$ as

$$T_{\mu\nu} = \int dP_1 dP_2 dP_3 (-g)^{-\frac{1}{2}} \frac{P_\mu P_\nu}{P^0} f(\eta, x^i, P_j) \tag{1.35}$$

where P_j is the conjugate momentum (the spatial part of the 4-momentum with lower indices) related to the proper momentum by

$$P_i = a(1 - \Phi(\eta, \mathbf{x}))p_i, \tag{1.36}$$

and g the determinant of the metric. The perturbed distribution function is the sum of the equilibrium distribution function and a perturbation

$$f(x^i, P_j, t) = f_0(\epsilon)\left[1 + \mathcal{F}(\eta, x^i, q_j)\right] \tag{1.37}$$

where $\epsilon = a(p^2 + m^2)^{\frac{1}{2}}$ is the comoving energy, $q_j = ap_j = q\hat{q}_j$ (q and $\hat{q}_j$ are the magnitude and direction of q_j) and $f_0(\epsilon)$ is either the Fermi-Dirac distribution for fermions or the Bose-Einstein distribution for bosons [11].

The Einstein field equations describe how the perturbations in the energy density of the various components of the Universe perturb the spacetime metric. In order to close the system, we also need additional equations describing how the time evolution of matter and radiation is affected by metric perturbations: the Boltzmann equations. The Boltzmann equation for the phase space distribution function $f(x^i, P_j, t)$ reads

$$\frac{df}{dt} = \frac{\partial f}{\partial t} + \frac{\partial x^i}{\partial t}\frac{\partial f}{\partial x^i} + \frac{\partial q}{\partial t}\frac{\partial f}{\partial q} + \frac{\partial \hat{q}_i}{\partial t}\frac{\partial f}{\partial \hat{q}_i} = C[f]. \tag{1.38}$$

Here $C[f]$ is the collision term describing interactions between different particle species. Combining these equations for each particle species with the Einstein field equations, we obtain the Einstein-Boltzmann system of ordinary differential equations. The system of coupled ordinary differential equations determines the evolution of the different components of the Universe at linear order and needs to be solved numerically. It is commonly expressed in Fourier space, where the derivatives become simple multiplications and where each Fourier mode evolves indipendently. The Fourier transform of $\delta(\eta, \vec{x})$ reads

$$\delta(\eta, \vec{k}) = \frac{1}{(2\pi)^{\frac{3}{2}}}\int d^3x\,\delta(\eta, \vec{x})e^{-i\vec{k}\cdot\vec{x}}. \tag{1.39}$$

We now look at the equations for different components in terms of conformal time η.

Dark matter and baryonic matter

Dark matter is described as a collisionless fluid, which only interacts gravitationally. Its evolution can easily be derived from the conservation equation of the perturbed stress-energy tensor ${T^{\mu\nu}}_{;\nu} = 0$ and results in

$$\dot{\delta} + ikv = -3\dot{\Phi} \tag{1.40}$$

$$\dot{v} + \frac{\dot{a}}{a}v = -ik\Psi \tag{1.41}$$

where dots corresponds to derivatives in η. We can also describe baryons as a non-relativistic fluid, but need to include a collision term due to Thomson scattering with the photons. The scattering mechanism conserves the number of particles but modifies their momenta. The Boltzmann equations for baryons describe the evolution of the baryon overdensity δ_b and velocity perturbation v_b

$$\dot{\delta}_b + ikv_b = -3\dot{\Phi} \tag{1.42}$$

$$\dot{v}_b + \frac{\dot{a}}{a}v_b = -ik\Psi - ikc_s^2\delta_b + \frac{4\rho_\gamma}{3\rho_b}\dot{\tau}\left[v_b + 3i\Theta_1\right] \tag{1.43}$$

where τ is the optical depth and the inverse of its comoving time derivative corresponds to the free streaming distance of the particles. Θ_1 is the first moment of the perturbations of the photons, that will be introduced further below. The time evolution of τ and c_s^2 require precise knowledge of the ionization history of electrons in the Universe, which is derived by solving the equations describing the atomic processes happening during recombination numerically [48–50].

Radiation

For the radiation components, we use the stress-energy tensor from equation 1.35. Neglecting the second order term in Equation 1.38 and using the geodesic equation to rewrite $\frac{\partial q}{\partial t}$, we obtain

$$\dot{\mathcal{F}} + i\frac{q}{\epsilon}(\vec{k}\cdot\hat{q})\mathcal{F} - \frac{d\ln f_0}{d\ln q}\left[\dot{\Phi} + i\frac{\epsilon}{q}(\vec{k}\cdot\hat{q})\Psi\right] = \frac{C[f]}{f_0} \tag{1.44}$$

where the collision term on the right-hand side will depends on the particle species (massless neutrinos or photons). In the case of neutrinos, the right-hand side can be set to $C[f] = 0$ since their interactions are very weak. Photons instead interact with the electrons via Thomson scattering. We define functions $\mathcal{N}(\vec{k},\hat{q},\eta)$ for the neutrinos and $\Theta(\vec{k},\hat{q},\eta)$ for the photons by integrating out the q-dependence in the distribution function and expanding the angular dependence in a series of Legendre polynomials $\mathcal{P}_\ell(\hat{k}\cdot\hat{q})$

$$\Theta(\vec{k},\hat{q},\eta) \equiv \frac{1}{4}\frac{\int dq q^3 f_0(q)\mathcal{F}(\eta,x,q,\hat{q})}{\int dq q^3 f_0(q)} = \sum_{\ell=0}^{\infty}(-i)^\ell(2\ell+1)\Theta_\ell(\vec{k},\eta)\mathcal{P}_\ell(\hat{k}\cdot\hat{q}). \tag{1.45}$$

Θ corresponds to the fractional temperature $\delta T/T$. We also have a function $\Theta_P(\vec{k},\hat{q},\eta)$ for perturbations in the polarization of photons. The Boltzmann equations for the radiation components result in

$$\dot{\Theta} + ik\mu\Theta = -\dot{\Phi} - ik\mu\Psi - \dot{\tau}\left[\Theta_0 - \Theta + \mu v_b - \frac{1}{2}\mathcal{P}_2(\mu)\Pi\right] \tag{1.46}$$

$$\dot{\Theta}_P + ik\mu\Theta_P = -\dot{\tau}\left[\Theta_P - \frac{1}{2}(1-\mathcal{P}_2(\mu))\Pi\right] \tag{1.47}$$

$$\dot{\mathcal{N}} + ik\mu\mathcal{N} = -\dot{\Phi} - ik\mu\Psi \tag{1.48}$$

where $\mu = \hat{k}\cdot\hat{q}$, $\mathcal{P}_2$ is the second Legendre polynomial, $\Pi = \Theta_2 + \Theta_{P,0} + \Theta_{P2}$ and we omitted the functional dependencies for readability. In order to conveniently solve these equations numerically, we write equations for the Legendre coefficients $\Theta_\ell(k,\eta)$

$$\Theta_\ell(k,\eta) = \frac{1}{(-i)^\ell}\int_{-1}^{1}\frac{d\mu}{2}\mathcal{P}_\ell(\mu)\Theta(k,\mu,\eta). \tag{1.49}$$

The hierarchy of equations would be infinite, but it is common to define a truncation rule [11]. The hierarchy of equations reads:

$$\begin{aligned}
\dot{\Theta}_0 &= -k\Theta_1 - \dot{\Phi} && (1.50)\\
\dot{\Theta}_1 &= \frac{k}{3}\left[\Theta_0 - 2\Theta_2 + \Psi\right] + \dot{\tau}\left[\Theta_1 - \frac{i v_b}{3}\right] && (1.51)\\
\dot{\Theta}_2 &= \frac{k}{5}\left[2\Theta_1 - 3\Theta_3\right] + \dot{\tau}\left[\Theta_2 - \frac{\Pi}{10}\right] && (1.52)\\
\dot{\Theta}_\ell &= \frac{k}{2\ell+1}\left[\ell\Theta_{\ell-1} - (\ell+1)\Theta_{\ell+1}\right] + \dot{\tau}\Theta_\ell, \;\; \ell > 2 && (1.53)\\
\dot{\Theta}_{P,\ell} &= \frac{k}{2\ell+1}\left[\ell\Theta_{P,(\ell-1)} - (\ell+1)\Theta_{P,(\ell+1)}\right] + \dot{\tau}\left[\Theta_{P,\ell} - \frac{\Pi}{2}\left(\delta_{\ell,0} + \frac{\delta_{\ell,2}}{5}\right)\right] && (1.54)\\
\dot{\mathcal{N}}_0 &= -k\mathcal{N}_1 - \dot{\Phi} && (1.55)\\
\dot{\mathcal{N}}_1 &= \frac{k}{3}\left[\mathcal{N}_0 - 2\mathcal{N}_2 + \Psi\right] && (1.56)\\
\dot{\mathcal{N}}_\ell &= \frac{k}{2\ell+1}\left[\ell\mathcal{N}_{\ell-1} - (\ell+1)\mathcal{N}_{\ell+1}\right], \;\; \ell > 1 && (1.57)
\end{aligned}$$

where $\delta_{\ell,0} = 1$ if $\ell = 0$ and $\delta_{\ell,2} = 1$ when $\ell = 2$ and zero otherwise.

Gravity

The perturbed Einstein field equations in conformal Newtonian gauge result in

$$k^2\Phi + 3\frac{\dot{a}}{a}\left(\dot{\Phi} - \frac{\dot{a}}{a}\Psi\right) = -4\pi G a^2 \delta T^0{}_0 \quad (1.58)$$

$$k^2\left(\frac{\dot{a}}{a}\Psi - \dot{\Phi}\right) = 4\pi G a^2 i k^j \delta T^0{}_j \quad (1.59)$$

$$-\ddot{\Phi} + \frac{\dot{a}}{a}(\dot{\Psi} - 2\dot{\Phi}) + (2\frac{\ddot{a}}{a} - \frac{\dot{a}^2}{a})\Psi - \frac{k^2}{3}(\Phi + \Psi) = \frac{4\pi}{3} G a^2 \delta T^i{}_i \quad (1.60)$$

$$k^2(\Phi + \Psi) = 12\pi G a^2 (\hat{k}_i \hat{k}_j - \frac{1}{3}\delta_{ij})\Sigma^i{}_j \quad (1.61)$$

where $\Sigma^i{}_j = T^i{}_j - \delta^i{}_j T^k{}_k/3$. Only two equations are needed to close the system, so the other two are redundant. We choose to use Equations 1.58 and 1.61. After rewriting

δT^{μ}_{ν} in terms of densities, velocities and second order moments, this results in

$$k^2\Phi + 3\frac{\dot{a}}{a}\left(\dot{\Phi} - \frac{\dot{a}}{a}\Psi\right) = 4\pi G a^2 \left[\bar{\rho}\delta + \bar{\rho}_b\delta_b + 4\left(\bar{\rho}_\gamma\Theta_0 + \bar{\rho}_\nu\mathcal{N}_0\right)\right] \tag{1.62}$$

$$k^2(\Phi + \Psi) = -32\pi G a^2 \left(\bar{\rho}_\gamma\Theta_2 + \bar{\rho}_\nu\mathcal{N}_2\right). \tag{1.63}$$

Initial conditions

We have discussed the evolution of perturbations in the Universe, but have not described the source of primordial fluctuations. The theory of inflation, postulated in the 1980s [51] to solve several problems arising in cosmology (like the flatness problem, the curvature problem and the abundance of magnetic monopoles), provides a mechanism to produce super-horizon perturbations starting from quantum fluctuations.

Inflation introduces an early phase of accelerated expansion which flattens out the curvature of the Universe and creates causal contact between parts of the Universe that would otherwise not be causally connected. In its most simple description, the accelerated expansion is caused by a scalar field which is in a false vacuum state and then rolls slowly to its true vacuum state. One prediction of the simple single-field inflation is that it sources adiabatic perturbations. This means that the density contrast δ_i is the same for all species i [52].

1.2.2 Growth of structure

We have shown formally how the different fields evolve in time at linear order in an expanding spacetime metric. We now want to describe the asymptotic behaviour of the perturbations of different components and how this leads to the growth of structure in the Universe.

Gravitational potential

We start by describing the evolution of the gravitational potential, since it is coupled with all the other fields. At early times, the perturbations are outside the horizon and therefore cannot be affected by causal physics. As time passes, the horizon expands and perturbations start entering the horizon. As we have seen in Chapter 1.1, different components dominate the Universe at different timescales. Smaller perturbations enter the horizon during radiation domination and are damped due to the fact that pressure is very efficient at counteracting gravity in this epoch. Perturbations on larger scales enter the horizon in the matter dominated era and start growing linearly with the scale factor $a(t)$.

Dark and baryonic matter

Dark matter makes up the majority of matter in the Universe and drives the evolution of the matter perturbations. It is common to parametrize the evolution of dark matter perturbations using the tranfer function $T(k)$ and the growth factor $D(a)$ as

$$\delta(k,a) \propto k^2 \Phi_p(k) T(k) D(a) \tag{1.64}$$

where $\Phi_p(k)$ is the primordial gravitational potential. In this equation, $T(k)$ accounts for the scale-dependence of the dark matter perturbations while $D(a)$ parametrizes the scale-independent growth. The time evolution of dark matter perturbations follows that of the gravitational potential.

Baryon perturbations generally follow the growth of dark matter perturbations, but are initially strongly coupled to the photons. This leads to a characteristic oscillatory behaviour due to the electromagnetic interactions in their early time evolution, which is imprinted on the large-scale structure of the Universe as Baryonic Acoustic Oscillations (BAO).

Radiation

Initially photons are tightly coupled to the baryons through Thomson scattering. At the epoch of recombination $z_{rec} \simeq 1100$, protons and electrons combine to form neutral hydrogen so that the Universe becomes transparent to radiation. Recombination ends at the large scattering surface, when photons are free-streaming in the Universe.

The initial coupling between baryons and photons implies a competition between gravity and pressure and the creation of acoustic waves in the baryon-photon plasma. When the photons decouple, this wave remains imprinted in both the baryonic matter distribution and the radiation perturbations. In real space this corresponds to an enhanced correlation on a specific scale, corresponding to the sound horizon at recombination

$$r_s(\eta_{rec}) = \int_0^{\eta_{rec}} c_s(\eta')d\eta', \tag{1.65}$$

whereas in Fourier space this corresponds to oscillations in the perturbations.

1.2.3 Nonlinear structure formation

At early times and on sufficiently large scales, linear perturbation theory is sufficient to study the growth of structure in the Universe. However, perturbations at late times become nonlinear ($\delta \gg 1$) on certain scales and linear perturbation theory is no longer valid. The most common approach to studying the growth of structure on non linear scales is through numerical dark matter simulations. These are widely known as N-body simulations since the matter field is described by N discrete matter particles. N-body simulations track the interactions that act between the different particles from initial conditions and evolve the phase space of the particles in time until today [53]. Fitting formulas for the non-linear evolution are derived from the N-body simulations (e.g. [54–56]). Analytical approaches such as the Press Schechter formalism and the halo model exist and are valid in certain regimes [57–60]. The picture that emerges in the ΛCDM paradigm is that of a hierarchical formation of structures: smaller dark matter haloes form first and merge to form larger haloes at later times.

1.2.4 Galaxies in the cosmological context

Galaxies form from interstellar gas and dust hosted in the potential wells of dark matter haloes. While the dark matter is collisionless, baryons undergo collisions until the interplay between gravity and pressure leads to gravitational collapse. The collapse causes the heating of the gas, which then cools via different radiative mechanisms that depend on its temperature and density [61]. There are different evolutionary paths of galaxies, depending on the local conditions, as well as the interactions with other systems. These processes can be studied through observations and with the help of hydrodynamical simulations [62] or semi-analytical models (SAMs) [63].

Galaxies are often classified as early-type (quiescent or red) and late-type (star-forming or blue). This is a simplistic description, since a lot of variability is observed in the galaxy population and galaxies are made up of bulges, spiral arms, disks, bars and stellar halos, with different formation mechanisms. The most straightforward way to distinguish different galaxy types is their morphology: this is a combination of the intrinsic shape of the galaxy and its orientation relative to the line of sight. The first galaxy classification dates back to 1926, when Hubble assigned galaxies to a tuning fork diagram consisting of ellipticals, spirals (normal and barred), lenticulars and irregular galaxies. Ellipticals and lenticulars are redder, meaning that the stars in the system are older, whereas spirals are bluer and still star forming. Galaxy morphology can be made more quantitative by measuring the surface brightness as a function of radius $I_\lambda(R)$. This is described by the Sersic light profile

$$I_\lambda(R) = I_e \exp\left\{-b(n)\left[\left(\frac{R}{R_e}\right)^{\frac{1}{n}} - 1\right]\right\} \tag{1.66}$$

where I_e is the surface brightness at the effective radius R_e, n is the Sersic index and $b(n) \approx 2n - 1/3$ [64]. Elliptical galaxies and the bulges of spirals have Sersic indices $2 < n < 10$, whereas disks have $n \sim 1$ and bars $n \lesssim 0.5$.

Many properties of a galaxy can be measured through its Spectral Energy Distribution (SED) $f(\lambda)$, which is the total energy emitted per unit wavelength. This can be measured with a spectrometer and yields a wealth of information about metallicity and star formation history encoded in the emission and absorption lines. In an optical telescope we measure the flux of a galaxy, which is the SED integrated over the throughput of an optical filter and the area of the galaxy

$$F_i = \int_0^\infty f(\lambda) R_i(\lambda) d\lambda \tag{1.67}$$

where $R_i(\lambda)$ is the filter response of the system, including atmospheric transmission. For historical reasons, we often use magnitudes instead of fluxes to describe the light coming from galaxies

$$m_i = -2.5 \log\left(\frac{F_i}{F_0^{AB}}\right) \tag{1.68}$$

where F_0^{AB} is the flux of the standard source in the AB magnitude system corresponding to a bandpass-averaged spectral flux density of 3631 Jy. m_i is the apparent magnitude of the system, which depends on the intrinsic luminosity L and the luminosity distance squared d_L^2. It is thus convenient to define a distance independent quantity M which is known as the absolute magnitude and corresponds to the apparent magnitude of a source at a distance of 10 pc from the observer. The relation between absolute and apparent magnitude in the same filter band is

$$m - M = 5 \log\left(\frac{d_L}{10pc}\right) + K(z) \tag{1.69}$$

with $K(z)$ the K correction [65], allowing us to account for the difference between emitted and observed wavelengths due to redshift. The colour of a galaxy is the difference between magnitudes in two different bands. This provides information on the ages and metallicities of stars, but also depends on the dust extinction along the line of sight. As already hinted at, there is a bimodality in the colour distribution, which correlates with the early-type/late-type morphological classification [66]. The region in between red and blue galaxies is often referred to as green valley.

In order to describe the overall galaxy population, one can write a stellar mass function, which describes the number of galaxies at a certain redshift in ranges of stellar mass. This depends on the halo mass function of dark matter and the stellar-to-halo mass ratio. A more direct observable is the luminosity function of galaxies [67] describing the number of galaxies per comoving volume per unit luminosity or absolute magnitude at fixed redshift z

$$\phi(z, M) = \frac{dN}{dMdV}. \tag{1.70}$$

We can rewrite the comoving volume element dV as

$$dV = \frac{d_H d_M^2}{E(z)} d\Omega dz \tag{1.71}$$

where d_H is the Hubble distance, d_M the transverse comoving distance, Ω the solid angle and $E(z) = \left(\Omega_r(1+z)^4 + \Omega_m(1+z)^3 + \Omega_k(1+z)^2 + \Omega_\Lambda\right)^{\frac{1}{2}}$. A good parametrization of the luminosity function is the Schechter function

$$\phi(z, M) = \frac{2}{5} \ln 10 \phi^* 10^{\frac{2}{5}(M^* - M)(\alpha+1)} \exp\left(-10^{\frac{2}{5}(M^* - M)}\right) \tag{1.72}$$

where $M*$ is the characteristic magnitude defining the change between the power law behaviour for faint galaxies and the bright-end exponential cutoff, $\phi*$ is the normalization and α describes the faint-end slope. Both $M*$ and $\phi*$ can be constrained as a function of redshift. Furthermore, it is common to separately fit a Schechter function for star-forming and quiescent galaxies.

1.3 Probes of large-scale structure

In cosmology we want to study and describe the Universe as a whole. This implies that there is only one single realization that we can observe and we cannot run multiple experiments. To overcome this problem, we look at the Universe statistically. We are interested in ensemble averages over multiple realizations of our Universe, and we use ergodicity to transform them into spatial averages [61]. In this thesis, we focus on photometric galaxy surveys. In these surveys, galaxies are observed in the night sky through an optical telescope. The images are then processed, stacked and sources are detected in the images. In order to derive cosmological information from these observations, we need to link properties of galaxies (such as their positions and shapes) to the parameters of our cosmological model. To achieve this, we start by describing the statistics of a dark matter overdensity field.

1.3.1 Cosmological field statistics

The dark matter overdensity field $\delta(t,\mathbf{x})$ generated from inflation is a Gaussian random field, so that the only moments needed to describe it are the first moment $\langle\delta(\mathbf{x})\rangle = 0$ and the second moment $\langle\delta(\mathbf{x})\delta(\mathbf{x}')\rangle \equiv \xi(\mathbf{x},\mathbf{x}')$ which is the correlation function. Due to homogeneity and isotropy, the field is invariant to translations and rotations and we can write

$$\xi(\mathbf{x},\mathbf{x}') = \xi(|\mathbf{x}-\mathbf{x}'|) \tag{1.73}$$

so that the correlation function only depends on the distance between two points. It is often convenient to predict observables in Fourier space and, using Equation 1.39, one can show that the Fourier transform of the correlation function is the power spectrum

$$P(\mathbf{k},\mathbf{k}') = \langle\delta(\mathbf{k})\delta^*(\mathbf{k}')\rangle = P(k)\delta_D(\mathbf{k}-\mathbf{k}') \tag{1.74}$$

where * indicates the complex conjugate and we have used isotropy and homogeneity to obtain the final result (where $\delta_D(\mathbf{k}-\mathbf{k}')$ is the Dirac delta function). The matter power spectrum $P(k)$ appears in most large-scale structure observables. The real space variance of the field can be obtained from the power spectrum as

$$\sigma^2 \equiv \langle\delta^2(\boldsymbol{x})\rangle = \int \frac{d^3k}{(2\pi)^3}\int \frac{d^3k'}{(2\pi)^3} e^{i\boldsymbol{x}(\boldsymbol{k}-\boldsymbol{k}')}\langle\delta(\boldsymbol{k})\delta^*(\boldsymbol{k}')\rangle = \int \frac{d^3k}{(2\pi)^3}P(k) \tag{1.75}$$

and defines the normalization of the power spectrum. It is common to smooth the density field with a spherical top-hat window function to measure this normalization on a given scale, most commonly $R = 8\text{Mpc}/h$, where $H_0 = 100h(\text{km/s})/\text{Mpc}$, so that

$$\sigma_8^2 = \frac{1}{2\pi^2}\int k^2 P(k)\left[\frac{3j_1(kR)}{kR}\right]^2 dk \tag{1.76}$$

where $j_1(kR)$ is the spherical Bessel function of order one. The whole treatment presented so far can be generalized to any cosmological field f (for example the overdensity fields of baryons, photons or neutrinos).

We often work on the two-sphere $\mathbb{S}_2$, due to the geometry of our sky observations, and it is therefore useful to define the correlation function also on the sphere. In this case

$$\xi(\boldsymbol{\theta},\boldsymbol{\theta}') \equiv \langle f(\boldsymbol{\theta}) f(\boldsymbol{\theta}')\rangle = \xi(\boldsymbol{\Delta\theta}) \tag{1.77}$$

where $\boldsymbol{\Delta\theta} = |\boldsymbol{\theta}-\boldsymbol{\theta}'|$ and $f(\boldsymbol{\theta}) = \sum_\ell \sum_{|m|<\ell} f_{\ell m} Y_{\ell m}(\boldsymbol{\theta})$ where $Y_{\ell m}(\boldsymbol{\theta})$ is the spherical harmonic of degree ℓ and order m. The equivalent of the power spectrum on the sphere are the spherical harmonic power spectra $C(\ell)$

$$\langle f^*_{\ell m} f_{\ell' m'}\rangle = C(\ell)\delta_{\ell,\ell'}\,\delta_{m,m'} \tag{1.78}$$

where $\delta_{\ell\ell'}$ is the Dirac delta function.

1.3.2 Galaxy clustering

Galaxies reside in dark matter haloes and are therefore biased tracers of the underlying dark matter distribution. We can write

$$\delta_g(\mathbf{k},a) = b(\mathbf{k},a)\delta(\mathbf{k},a) \tag{1.79}$$

where $b(\mathbf{k},a)$ is the galaxy bias [68] and is often approximated as linear and scale-independent $b(\mathbf{k},a) = b_1$ on sufficiently large scales and for specific populations of galaxies. From here on, we will make use of this assumption.

The distribution of galaxies is a three dimensional observable so it is best traced in a spectroscopic survey, where the redshift is obtained by fitting the absorption and emission lines of the galaxy's SED. Spectroscopy is time consuming and costly though, and is limited to brighter samples than photometry. It also requires targeting of objects from an imaging survey and is prone to selection effects. In this work, we focus on photometric galaxy surveys, which can span wider areas and consist currently of hundreds of millions of objects. The galaxy positions in a photometric survey are used to measure the angular correlation function (or power spectrum) in tomographic bins. The reason for this, is the fact that the photometric determination of redshift yields larger uncertainties, so that we do not want to rely on the photometric redshifts of individual objects. We describe ways to measure photometric redshifts in Section 1.3.6.

From Equation 1.79, we can derive the galaxy-galaxy power spectrum

$$\langle \delta_g(\mathbf{k}, a)\ \delta_g^*(\mathbf{k}, a)\rangle = P_{gg}(k, a)\delta_D(\mathbf{k}-\mathbf{k}') = b_1^2 P(k, a)\delta_D(\mathbf{k}-\mathbf{k}'). \tag{1.80}$$

The dark matter power spectrum $P(k, a)$ must include the non linear evolution of dark matter, since galaxy clustering is a late-time probe. In order to obtain cosmological constraints from wide field surveys, we need to relate $P_{gg}(k, a)$ to the angular power spectrum $C_{gg}(\ell)$, by integrating over the line of sight. For simplicity, we use the Limber approximation [69] and consider two tomographic bins i and j, then

$$C_{gg}(\ell) = \int P\left(k = \frac{\ell + \frac{1}{2}}{\chi}, z\right)\left[W_\ell{}^i\left(\chi(z)\right) W_\ell{}^j\left(\chi(z)\right)\right]\frac{d\chi}{\chi^2} \tag{1.81}$$

where

$$W_\ell{}^i(\chi(z)) = n^i(z) b_1 \tag{1.82}$$

with $n(z)$ the redshift distribution of the galaxy sample normalized to unity.

1.3.3 Gravitational lensing

While the positions of galaxies trace the baryonic density field, their shapes directly reflect the total matter content of the Universe. This is due to the deep connection predicted by general relativity between the paths of photons travelling towards the observer and the inhomogeneities present along their way. The gravitational lensing effect can cause very dramatic distortions of the image of the object, such as multiple images and time delays, when the photon paths are deflected by a galaxy cluster or a very massive galaxy: this effect is called strong gravitational lensing [70, 71]. When we instead look at the cumulative and coherent effect that the overall large-scale structure distribution has on the statistical shapes of galaxies, we talk about cosmic shear. This weak lensing regime is the focus of wide field photometric surveys and yields stringent constraints on the cosmological parameters. Extensive treatments of gravitational lensing can be found in many reviews and books [36, 72–77]. In the following we provide a concise description of this effect.

As photons travel from a far away galaxy to the observer, their paths get distorted by the intervening matter but the surface brightness of the galaxy is conserved

$$I_{\mathrm{obs}}(\boldsymbol{\theta}) = I_{\mathrm{true}}(\boldsymbol{\theta}_S) \tag{1.83}$$

where $\boldsymbol{\theta}$ is the angle under which the observer sees the image of the galaxy and $\boldsymbol{\theta}_S$ is the true angular position of the galaxy. We consider two photons travelling from a source galaxy to the observer along null geodesics. The transverse comoving separation between the two light rays in a flat Universe is $\boldsymbol{x}_0 = \chi\boldsymbol{\theta}$ where χ is the comoving distance from the observer. Neglecting the anisotropic stress on large scales so that the two Newtonian gravitational potentials are equal $\Psi(a,\boldsymbol{x}) = \Phi(a,\boldsymbol{x})$, assuming $\Phi(a,\boldsymbol{x}) \ll c^2$ and using the Born approximation (the perturbed path of photons does not differ much from a straight line if we assume the weak lensing regime), one can compute the modification of the angular separation due to the lensing effect which results in

$$\boldsymbol{\theta}_S = \boldsymbol{\theta} - \frac{2}{c^2}\int_0^{\chi} \frac{\chi-\chi'}{\chi}\left[\nabla_\perp\Phi(\chi'\boldsymbol{\theta},\chi') - \nabla_\perp\Phi^{(0)}(\chi')\right] d\chi' \tag{1.84}$$

where $\Phi^{(0)}$ is the potential along the fiducial light ray and $\nabla_\perp$ is a derivative along the separation $\boldsymbol{x}$. By linearising the lens equation and defining the lensing potential

$$\psi(\boldsymbol{\theta}, \chi) = \frac{2}{c^2} \int_0^{\chi} \frac{\chi - \chi'}{\chi \chi'} \Phi(\chi' \boldsymbol{\theta}, \chi') d\chi' \tag{1.85}$$

one obtains the distortion matrix

$$\mathcal{A} = \frac{\partial \boldsymbol{\theta}_S}{\partial \boldsymbol{\theta}} = \delta_{ij} - \psi_{,ij} = \begin{pmatrix} 1 - \kappa - \gamma_1 & -\gamma_2 \\ -\gamma_2 & 1 - \kappa - \gamma_1 \end{pmatrix} \tag{1.86}$$

where $\psi_{,ij} = \partial^2 \psi / \partial\theta_i \partial\theta_j$ and we have introduced the two shear components $\gamma \equiv \gamma_1 + i\gamma_2$ and the convergence κ which are defined in terms of derivatives of the lensing potential. In the weak lensing regime, the shear components describe the elongation of a galaxy along and at 45° from the horizontal axis of the cartesian coordinate system centered at the galaxy and the convergence describes the magnification of the galaxy. By introducing the angular position ϕ of the source galaxy with respect to the x-axis, we can rewrite the shear γ in terms of tangential shear γ_t and cross shear $\gamma_\times$

$$\gamma_t = -\gamma_1 \cos(2\phi) - \gamma_2 \sin(2\phi) \tag{1.87}$$

$$\gamma_\times = \gamma_1 \sin(2\phi) - \gamma_2 \cos(2\phi). \tag{1.88}$$

The tangential component is usually referred to as E-mode, which is curl-free, and the cross shear is referred to as B mode, which is divergence free and vanishing since the shear field is sourced by a single scalar potential.

Using the Poisson equation

$$\Delta\Phi = 4\pi G a^2 \bar{\rho}_m \delta, \tag{1.89}$$

which is a late-time limit of Equation 1.58, we can rewrite the convergence κ in terms of the dark matter density perturbations and matter energy density:

$$\kappa(\boldsymbol{\theta}, \chi) = \frac{1}{2} \nabla^2 \psi(\boldsymbol{\theta}, \chi) = \frac{3 H_0^2 \Omega_m}{2c^2} \int_0^{\chi} \frac{(\chi - \chi')\chi'}{a(\chi')\chi} \delta(\chi' \boldsymbol{\theta}, \chi') d\chi'. \tag{1.90}$$

So far we have considered sources at a specific distance χ from the observer. In order to express the mean convergence from a population of source galaxies we integrate along the line of sight

$$\kappa(\boldsymbol{\theta}) = \int_0^{\chi_{\max}} n(\chi)\kappa(\boldsymbol{\theta},\chi)d\chi = \frac{3H_0^2\Omega_m}{2c^2}\int_0^{\chi_{\max}} \frac{g(\chi)\chi}{a(\chi)}\delta\left(\chi\boldsymbol{\theta},\chi\right)d\chi \tag{1.91}$$

where $n(\chi)$ is the source redshift distribution and

$$g(\chi) = \int_\chi^{\chi_{\max}} n(\chi')\left(\frac{\chi'-\chi}{\chi'}\right)d\chi' \tag{1.92}$$

is known as lensing efficiency. This means that the convergence is a 2D projection of the matter overdensity δ weighted by the lensing potential. We can now write the spherical harmonics power spectrum of the convergence of two tomographic bins i and j

$$C_{\kappa\kappa}^{ij}(\ell) = \left(\frac{3H_0^2\Omega_m}{2c^2}\right)^2 \int_0^{\chi_{\max}} P\left(k = \frac{\ell+\frac{1}{2}}{\chi},\chi\right)\frac{g^i(\chi)g^j(\chi)}{a(\chi)^2}d\chi \tag{1.93}$$

where we have used the Limber approximation to only consider modes perpendicular to the line of sight. Since cosmic shear is sourced by a single scalar field we have $C_{\gamma\gamma}(\ell) = C_{\kappa\kappa}(\ell) = C_{EE}(\ell)$ and $C_{BB}(\ell) = 0$.

We now have a connection between shear and cosmological parameters, but we still need to relate shear and galaxies' ellipticity. Assuming that galaxies are randomly oriented in space so that the instrinsic ellipticity $\langle\epsilon_{\text{int}}\rangle = 0$, we can use the measured ellipticities of galaxies to infer the underlying matter distribution. We first rewrite the distortion matrix

$$\mathcal{A} = (1-\kappa)\begin{pmatrix}1-g_1 & -g_2\\ -g_2 & 1-g_1\end{pmatrix} \tag{1.94}$$

where

$$g_i = \frac{\gamma_i}{1-\kappa} \tag{1.95}$$

is called reduced shear. This shows that the convergence κ only affects the size of the objects and the observable change in the ellipticities is controlled by the reduced shear. If we measure the shape of a galaxy using elliptical isophotes, the observed ellipticity can be related to the intrinsic ellipticity [78] through

$$\epsilon_{\text{obs}} = \frac{\epsilon_{\text{int}} + g}{1 + g^{*}\epsilon_{\text{int}}} \approx \epsilon_{\text{int}} + \gamma \tag{1.96}$$

where g^{*} is the complex conjugate of g and the approximation is valid in the weak lensing regime.

1.3.4 Galaxy galaxy lensing

Galaxy-galaxy lensing (GGL) correlates shapes of high-redshift galaxies (sources) with positions of galaxies at lower redshift (lenses). The spherical harmonics power spectrum is thus a cross-correlation of the two introduced in the previous sections

$$C^{ij}_{g\kappa}(\ell) = \frac{3H_0^2\Omega_m}{2c^2} \int P\left(k = \frac{\ell + \frac{1}{2}}{\chi}, \chi\right) n^{i}(\chi(z)) b_1 \frac{g^{j}(\chi)}{a(\chi)} \frac{d\chi}{\chi}. \tag{1.97}$$

$C^{ij}_{g\kappa}(\ell)$ can be measured from convergence and galaxy density maps. A more common observable of galaxy-galaxy lensing is the mean tangential shear as a function of angular scale $\langle\gamma_t(\boldsymbol{\theta})\rangle$. This requires finding lens-source pairs at given angular separations, measuring the tangential shear and averaging over angular bins. $\langle\gamma_t(\boldsymbol{\theta})\rangle$ can be related to the mean convergence in a circular aperture of angular radius θ and is thus a measurement of the total projected mass within the radius. Another related GGL observable is the excess of surface mass density $\Delta\Sigma$, where the surface mass density Σ is defined in terms of the convergence κ. The derivation of these quantities and a more extensive description can be found e.g. in [79].

1.3.5 3×2-point analysis

The combination of galaxy clustering, cosmic shear and galaxy-galaxy lensing is known as 3×2-point analysis. It is highly beneficial to combine the three large-scale structure observables to constrain cosmological parameters since different probes are subject to different systematic effects and biases. Galaxy clustering is strongly affected by the relation between dark matter and baryons encoded in the galaxy bias modeling, which is highly uncertain [68]. Cosmic shear, as mentioned, is a much more direct probe of the total matter distribution. Despite the theoretical simplicity, there are several systematic effects affecting cosmic shear. First of all, the shape measurement needs to be very accurate and the residual difference between the true and measured shear is often parametrized by an additive and a multiplicative shear bias [80], calibrated on simulations. One source of such biases is the Point Spread Function (PSF), which is the response of the optical system to a point source caused by the atmosphere and the optics of the telescope. The PSF is convolved with the intrinsic profile of a galaxy and smears it, so that accurate PSF reconstruction is crucial for shape measurement [81]. Furthermore, most cosmological signal in cosmic shear comes from small faint galaxies. For this reason, spectroscopic estimation of redshifts becomes infeasible and surveys need to rely on photometry in a few broad-bands for estimating the redshift distribution of the source sample. The sample is split in different tomographic bins, in order to increase the constraining power and reconstruct some of the three dimensional information [82]. We treat photometric redshift calibration in detail in the following subsection 1.3.6, as it is the focus of the second half of this thesis. Another effect that can bias cosmic shear measurements is the intrinsic alignment of galaxies [83], the fact that galaxy shapes are not uncorrelated but depend on complex astrophysical processes and the local environment. Finally, modelling choices, approximations and baryonic effects all have the potential to bias results from large-scale structure probes by shifting power between scales in the power spectrum, even though these effects should be negligible in current surveys. Galaxy-galaxy lensing inherits the systematics of cosmic shear and galaxy clustering but probes the relation between dark and baryonic matter, thereby helping to break degeneracies between cosmological parameters and the galaxy bias [84]. The 3×2-point analysis is a powerful probe of cosmological parameters describing the matter distribution of the local Universe, mostly Ω_m and σ_8, and also the dark energy parameters such as Ω_Λ and the dark energy equation of state parameter w.

1.3.6 Photometric redshift estimation

In the previous sections, we saw how the redshift distribution of galaxies $n(z)$ appears in the kernels of the different large-scale structure observables. As already mentioned, the most accurate way to retrieve redshift information would be through high resolution spectroscopy. This is infeasible for the large samples we are interested in measuring in current and future cosmological galaxy surveys, most notably for cosmic shear, which benefits from large samples and from small faint galaxies. Another promising approach is that of obtaining a low resolution spectrum by collecting many fluxes in narrow-band filters and with broad wavelength coverage. This is also costly, and currently available only for selected deep fields such as COSMOS [29, 85, 86] or to bright limiting magnitudes (for example in the PAU Survey [87]). In wide-field cosmological galaxy surveys we are thus relying on images in a few broad-band filters to constrain the redshift distribution of our samples, a technique known as photometric redshift (photo-z).

Many photo-z methodologies have been developed and compared over the years, as redshift requirements get more stringent [23, 88–93]. The precise determination of the redshift distribution is currently one of the main systematic errors of weak lensing and photometric clustering, possibly leading to biases in the determination of cosmological parameters. We can broadly distinguish two different needs for photo-zs: probes such as clustering (or galaxy evolution studies) benefit from the knowledge of individual redshifts of galaxies whereas often we can get away with the overall redshift distribution of an ensemble of galaxies, possibly divided in several tomographic redshift bins. To obtain individual redshifts there are two main approaches: SED fitting and machine learning. When constraining the overall redshift distribution, one can use techniques to stack or combine individual redshifts or methods specifically designed for redshift distributions, such as direct calibration, spatial cross-correlations and hierarchical Bayesian models or forward modelling. In the following, we give some details on each family of methods.

SED fitting methods

SED fitting methods rely on a family of known template spectra, coming either from spectroscopic observations or from stellar population synthesis models. The templates are either fitted individually or combined. Through a χ^2-minimization, the methods determine the redshifted template whose broad-band colours best match the observed galaxy, and determine simultaneously the galaxy redshift and type. The goal is to estimate the posterior $p(z|\text{fluxes})$ using Bayes' theorem, so that prior knowledge on the expected shape of the redshift distribution or the fraction of different galaxy types can be included. Wrong prior information or an incomplete template set can cause problems with these methods. Commonly used codes in this cathegory are: LePhare [85, 94], EAZY [95], BPZ [96], ZEBRA [97] and MIZUKI [98].

Machine learning methods

The idea behind the machine learning methods is to rely on training data, where the fluxes and the redshifts are known, either from spectroscopy or many-band photometry. Sometimes the methods are directly applied to images, other times features such as fluxes and colours are extracted from the data first. The algorithm of choice (usually an artificial neural network, gaussian process or random forest) will then minimize a score, the details of which depend on the specific method and on whether the result is the point estimate of the redshift or a probability distribution function. The weakness of these methods comes from the training sample: if this is not complete and representative of the target sample, the methods are prone to biases. Machine learning methods include ANNz [99], ANNz2 [100], DNF [101], GPz [102], TPZ [103], MLPQNA [104], METAPHOR [105] and DEmp [106] among many others.

Direct calibration and spatial cross correlations

Direct calibration methods rely on known redshift samples, that are reweighted to be representative of the sample of interest [107]. In the simplest scenario this implies a colour reweighting of the well-known redshifts in order to match the colour distribution of the galaxies with unknown redshifts. More sophisticated methods rely on dimensionality reduction in order to group galaxies with similar redshifts and properties. In recent years several surveys have used Self-Organizing Maps (SOM) [108], an unsupervised learning algorithm, to map photometric properties to cells of a two dimensional plane [103, 109–113]. Each SOM cell is calibrated with known redshifts and the redshift distribution of the sample of interest is derived probabilistically, taking into account selection effects.

Cross-correlation methods, often referred to as clustering redshifts, rely on the spatial correlation between the sample of interest and galaxies with well-known redshifts [114–119]. The secure redshifts are binned in narrow redshift bins and the two-point correlation function with the unknown redshifts are computed. The method relies on the knowledge of the redshift evolution of the galaxy bias. It is a complementary method, since it does not rely on photometric properties of the galaxies but only on their positions.

Forward modelling

Forward modelling in the context of photometric redshifts is the idea that it is possible to rely on the knowledge of the galaxy population, either explicitly or through observed distributions of intrinsic properties, in order to obtain the redshift distribution of a sample of galaxies [120, 121]. This makes use of forward simulations, so that the simulated data can be treated like real data, including observational effects and sample selection.

Hierarchical Bayesian methods infer both the posterior distribution for the redshifts of individual galaxies and the parameters of the model describing the population of galaxies [122–126]. These models are extensions of SED fitting where the prior is calibrated on the data. They differ in whether the inferred prior describes the intrinsic galaxy properties or the observed ones.

The forward modelling method that we employ and further develop in this thesis is called Monte Carlo Control Loops (MCCL) [27, 127]. The method relies on a parametric galaxy population model and on a simulator that renders realistic telescope images from a specific survey. The parameters of the model are constrained using the real telescope data via Approximate Bayesian Computation (ABC). Once the model is tuned, we can use it to simulate a whole survey and obtain the redshift distributions of galaxies [120] or to calibrate the cosmic shear measurement [127]. The redshift distribution of galaxies depends both on the intrinsic properties of the population and on the selection effects, which are carefully modelled. The robustness of the method can be tested by perturbing the simulator's configuration or by extending the forward model and the uncertainty is naturally included by running an ensemble of simulations.

1.4 Contributions

I wrote the introductory Chapter 1 based on my review of the cited literature.

CHAPTER 2

SYMBOLIC IMPLEMENTATION OF EXTENSIONS OF THE PYCOSMO BOLTZMANN SOLVER

> *There is a theory which states that if ever anyone discovers exactly what the Universe is for and why it is here, it will instantly disappear and be replaced by something even more bizarre and inexplicable.*
>
> *There is another theory which states that this has already happened.*
>
> — DOUGLAS ADAMS, The Restaurant at the End of the Universe

This chapter appeared in a similar form in Moser et al., 2022 [31].

2.1 Introduction

Our understanding of the Universe relies on the possibility to predict cosmological observables from theoretical principles. One of the key theoretical predictions is the evolution over time of the linear order perturbations of the constituents of the Universe, captured by the Einstein-Boltzmann equations (see, e.g., [11], [128]). The system of ordinary differential equations, due to its complexity and the coupling of the fields, needs to be solved numerically (see, e.g., [129]). For this purpose, several codes have been developed since the release of the pivotal Boltzmann code `COSMICS` [130], closely followed in time by `CMBFAST` [131], later ported to `C++` with the name `CMBEASY` [132]. The currently maintained Boltzmann solvers are `CAMB` [133], `CLASS` [134] and `PyCosmo` [12]. Both for `CAMB` and `CLASS` several codes have been written to include extensions beyond ΛCDM, for example `hi_class` [135] and `EFTCAMB` [136] for modified gravity theories, `CLASS_EDE` [137] for early dark energy and `CLASSgal` [138] to include general relativistic effects in the computation of galaxy number counts. `PyCosmo` was introduced by [12] as a novel `Python` library that uses symbolic representation of equations for generating efficient `C/C++` code. The framework includes both a Boltzmann solver as well as prediction tools for the computation of cosmological observables with several different fitting functions and approximations [139]. With these tools, `PyCosmo` offers similar utilities as, e.g., the Core Cosmology Library `CCL` [140] developed by the Dark Energy Science Collaboration (DESC).

An important feature of `PyCosmo` is the possibility to easily implement model extensions in symbolic form in the code, while taking advantage of the computational speed of the generated `C/C++` code. This feature has been improved by rewriting and refactoring the related code as a new `sympy2c` package presented in [141], which expands the idea of generating fast `C/C++` code from symbolic representations of equations. Both `sympy2c` and `PyCosmo` are publicly available in the `Python` Package Index (PyPI). In this work, we illustrate how the `PyCosmo` Boltzmann solver can be extended thanks to the symbolic framework by implementing several extensions. We introduce two extensions of the Standard Model of Cosmology: a constant dark energy equation of state and massive neutrinos. We also include a Radiation Streaming Approximation (RSA) for photons and massless relics, which approximates the evolution of radiation at late times and includes reionisation, following the treatment in [17] implemented in `CLASS`. This approximation speeds up the code by reducing the number of equations in the ODE system and avoiding the reflection of power caused by the truncation of the multipole expansion of the radiation equations.

We begin by giving an overview of the new features of the `PyCosmo` framework in

Section 2.2, where we discuss also the usage of the code and the precision settings used to compare with `CLASS`. We then present the equations of the models, implementation details and code comparisons with `CLASS`, both in terms of agreement and performance. In Section 2.3 we present the implementation of the constant dark energy equation of state, since it is a minimal modification of the Boltzmann system of equations for ΛCDM. We describe the inclusion of massive neutrinos, treated as non-interacting and non relativistic relics, in Section 2.4. In Section 2.5 we present the radiation streaming approximation, which requires `sympy2c` to handle a switch between two different equation systems and is then applied to all models. In Section 2.6 we discuss the results obtained by benchmarking the speed of the computations and comparing the numerical results to `CLASS`. We conclude in Section 2.7. This work heavily relies on the Boltzmann equations presented in [11]. To translate the `PyCosmo` notation to the Ma-Bertschinger and `CLASS` notation, we refer the reader to Appendix 2.A.1. Appendix 2.A.2 presents the Einstein-Boltzmann system of ODE in `PyCosmo` notation, using $\ln a$ as the independent variable. The adiabatic initial conditions for ΛCDM and all the other implemented models are shown in Appendix 2.A.3. We report in Appendix 2.A.4 the parameters of `PyCosmo` and `CLASS` that have been kept constant throughout the chapter. In Appendix 2.A.5, we provide a self contained summary of the computation of the total matter power spectrum, including general relativistic corrections, and using the $\mathcal{A}_s$ normalisation parameter.

2.2 PyCosmo framework

2.2.1 `C/C++` code generation

We reimplemented and improved the `C/C++` code generation related parts of previous versions of `PyCosmo` [12] as a separate `Python` package named `sympy2c` which we describe in detail in [141]. `sympy2c` translates symbolic representations of expressions and ordinary differential equations to `C/C++` code and compiles this code as a `Python` extension module.

`sympy2c` replaces the Backward-Differentiation-Formula (BDF) solver from the previous version of `PyCosmo` by the established and robust Livermore Solver for Ordinary Differential Equations (LSODA) solver [142] for improved step-size control and error diagnostics. LSODA detects stiff and nonstiff time domains automatically and switches between the nonstiff Adams method and the stiff BDF method. The BDF method solves a linear system derived from the Jacobian matrix of the differential equations at each time step. This affects runtime significantly for large systems. `sympy2c` leverages the symbolic form of the ODE and generates code to solve such systems efficiently by avoiding unnecessary computations based on the known sparsity structure of the involved Jacobian matrix.

To solve such linear systems `sympy2c` unrolls loops occurring in used LU factorization with partial pivoting (LUP) algorithm during code generation. This procedure depends on predetermined row permutations of the system, and the generated code includes checks for whether the considered permutation is appropriate for ensuring numerical accuracy. When solving the ODE, a new, not yet considered, permutation might arise. In this case, the solver delegates to a fall-back general LUP solver and records the new permutation. The result is a valid result but with a sub-optimal computation time. In this case, the warning message `"there are new permutations pending, you might want to recompile"` will be displayed together with the command necessary to recompile. Running the code-generator again will then also create optimized code for the newly recorded permutation(s), so that future runs of the solver will benefit from this. This approach starts with the identity permutation and could require several steps of solving the Boltzmann equations followed by code generation and compilation to achieve optimal performance. In our experiments not more than one such iteration is needed.

A large ODE system can result in `C/C++` functions with millions of lines of code which challenge the compiler and can cause long compilation times and high memory consumption, especially during the optimization phase of the compiler. To mitigate this, `sympy2c` can split the original matrix into smaller blocks and then generate code to implement blocked Gaussian elimination using Schur-complements. This affects the generated `C/C++` code by creating more but significantly shorter functions and thus supports the optimization step of the compiler. Another benefit of this approach is that runtime is improved by reducing the number of cache misses on the CPU.

Depending on the size and sparsity structure of the system, this code generation and compilation step can take seconds up to 30 minutes or even more. `PyCosmo` and `sympy2c` use caching strategies that consider previously generated code so that cached solvers are available within fractions of a second.

2.2.2 Usage

The equations for ΛCDM and the extended models are implemented symbolically in `PyCosmo`, both with and without radiation streaming approximation, in the `CosmologyCore_`*model*`.py` and `CosmologyCore_`*model*`_rsa.py` files. Supported models are currently `"lcdm"`, `"wcdm"` or `"mnulcdm"`.

The method `PyCosmo.build` initializes an instance of the `Cosmo` class for subsequent computations. `Cosmo` is the class that manages most of the functionalities of `PyCosmo` and on which all the other classes rely. `PyCosmo.build` requires the name of the model as well as all parameters which influence the code generation and compilation step. The argument `rsa` enables or disables the RSA and `l_max` specifies the maximum moment for truncating the photons and massless neutrinos hierarchies. The `"mnulcdm"` model also accepts parameters `l_max_mnu` and `mnu_relerr` which we describe later. Furthermore, parameters controlling the compiler, such as the optimization flag `-O`*n* and the `splits` to use to reduce memory consumption and compilation time, can be specified when calling `PyCosmo.build`. All the parameters that can be passed to `build` are specified in Table 2.1.

Parameters which do not affect code generation, such as cosmological parameters, precision settings, parameters specific for approximations and physical constants can be set or modified using the `Cosmo.set` method. Each cosmological model is equipped with a default set of such parameters, contained in a `default_`*model*`.ini` file. Listing 2.1 demonstrates how to create a cosmology and change parameters.

```python
import PyCosmo
cosmo = PyCosmo.build("lcdm", l_max=50, rsa=True)
cosmo.set(omega_m=0.28)
cosmo.print_params()
```

Listing 2.1: Creation of a ΛCDM cosmology using RSA and truncating photons and massless neutrino hierarchies at `l_max` = 50. We override the default value of Ω_m by 0.28 and print a report of all parameter values.

One parameter that is particularly relevant is `pk_type`, since it allows the user to switch between the Boltzmann solver (`pk_type` = `"boltz"`) and the approximations (`pk_type` = `"EH"` for the fitting function by Eisenstein and Hu [143] and `pk_type` = `"BBKS"` for the BBKS polynomial fitting function [144]). In this work we will always set `pk_type = "boltz"`. Other cosmological and precision parameters which are kept fixed throughout the chapter are reported in Appendix 2.A.4. Tutorials for the computation of cosmological observables and the usage of the Boltzmann solver can be found on the `PyCosmo Hub` (see [139]), a public platform hosting the current version of `PyCosmo`, along with `CLASS`, `CCL`, and `iCosmo` [145], an IDL predecessor of `PyCosmo`. The link to the `PyCosmo Hub` can be found at https://cosmology.ethz.ch/research/software-lab/PyCosmo.html.

2.2.3 Code comparisons setup

In order to validate the newly introduced models, we carry out detailed comparisons with `CLASS`[1]. We evaluate the accuracy in terms of relative difference:

$$\text{Relative difference}(\texttt{PyCosmo}, \texttt{CLASS}) = \frac{X_{\texttt{PyCosmo}} - X_{\texttt{CLASS}}}{X_{\texttt{PyCosmo}}}$$

where X is the cosmological observable we want to compare, for example the total matter power spectrum. Since this is a function of the wavenumber k, we compare

---

[1] We use `CLASS` v3.1.0 throughout the chapter, through the `Python` wrapper `classy`.

it visually by plotting as a function of k or, when specified, we look at the maximum relative difference in a k interval.

Comparing the two codes implies carefully setting the cosmological parameters and the precision settings. In the case of `CLASS`, we use the precision files shipped with the code: `cl_permille.pre` for fast and accurate computation and `pk_ref.pre` for high precision. `cl_permille.pre` guarantees a precision of 0.1% up to $k = 1h\mathrm{Mpc}^{-1}$ for the matter power spectrum, and `pk_ref.pre` a precision of 0.001% on scales $k < 50h\mathrm{Mpc}^{-1}$ [146]. Both `CLASS` precision settings use the radiation streaming approximation, described in detail in Section 2.5. They also include the tight coupling approximation and ultra-relativistic fluid approximation (see [17]), whereas only `cl_permille.pre` uses the fluid approximation for massive neutrinos (presented in [147]). The main parameters controlling precision in `PyCosmo` are the `l_max` parameter which defines the truncation of the multipole hierarchy for radiation fields (same for photons and massless neutrinos) and the `boltzmann_rtol` and `boltzmann_atol` parameters defining the relative and absolute tolerance of the LSODA ODE solver. Massive neutrinos also add two important precision parameters which will be described more in detail in section 2.4: `mnu_relerr` controlling the number of momenta q used for the massive neutrino integrals and `l_max_mnu`, controlling the truncation of the multipole expansion for massive neutrinos. We summarise all the precision parameters available in `PyCosmo` in Table 2.1. The two precision settings that we use in `PyCosmo` when comparing respectively to `CLASS` `cl_permille.pre` and `pk_ref.pre` are:

- *speed*: `l_max` = 17, `l_max_mnu` = 17, `rtol` = 10^{-5}, `atol` = 10^{-5}, `mnu_relerr` = 10^{-5}
- *precision*: `l_max` = 50, `l_max_mnu` = 50, `rtol` = 10^{-6} and `atol` =10^{-6}, `mnu_relerr` = 10^{-6}.

When using the RSA in `PyCosmo`, we set the RSA trigger parameters (detailed in Section 2.5) to:

- *speed*: `rsa_trigger_taudot_eta` = 5, `rsa_trigger_k_eta` = 45

Parameter	Description	Method	Default
`rsa`	Switch for the RSA	`build`	`False`
`compilation_flags`	GCC compiler's optimization flag	`build`	"-O3"
`splits`	Splittings of the ODE system	`build`	None
`reorder`	Whether to reorder the ODE system to speed up code generation when using `splits`	`build`	`True`
`l_max`	Hierarchy truncation of the relativistic relics	`build`	20
`l_max_mnu`	Hierarchy truncation of massive neutrinos	`build`	20
`mnu_relerr`	Relative error for massive neutrinos integral	`build`	1e-5
`sec_factor`	Safety factor for permuting rows in the LUP decomposition within LSODA	`set`	10
`boltzmann_rtol`	Relative tolerance of the LSODA solver	`set`	1e-5
`boltzmann_atol`	Absolute tolerance of the LSODA solver	`set`	1e-5
`boltzmann_max_bdf_order`	Maximum order used by the BDF integrator	`set`	5
`boltzmann_max_iter`	Max number of iterations of the LSODA solver	`set`	2e6

Table 2.1: Description of the precision parameters available in the PyCosmo Boltzmann solver, along with the method they are passed to and their default values.

- *precision*: `rsa_trigger_taudot_eta` = 100, `rsa_trigger_k_eta` = 240

which match the equivalent CLASS parameters in `cl_permille.pre` and `pk_ref.pre`. All the other precision parameters have default values, as in Table 2.1. We do not attempt to exactly match the precision parameters in the two packages, since CLASS includes a number of approximations that are not available in PyCosmo. It is possible to switch off most of the approximations but this would imply losing the precision guarantees of the default precision files.

The cosmological parameters that remain constant throughout the chapter are listed in Appendix 2.A.4, whereas the matter energy density Ω_m, the number of massless neutrinos N_ν, the number of massive neutrinos $N_{\nu,m}$, the total sum of neutrino masses Σm_ν and the dark energy equation of state parameter $w_{\rm de}$ (corresponding in `CLASS` to $\Omega_{cdm} = \Omega_m - \Omega_b$ since Ω_b is fixed, N_{ur}, N_{ncdm}, $m_{ncdm} = \Sigma m_\nu / N_{\nu,m}$ and $w_{0,fld}$) change in different models and are reported in the corresponding sections. Ω_Λ is computed by imposing the flatness condition $\sum_i \Omega_i = 1$.

2.3 A simple model: dark energy with a constant equation of state

2.3.1 Equations

In order to search for deviations from a cosmological constant, we consider here a dark energy equation of state with $w_{\rm de} = p_{\rm de}/\rho_{\rm de} \neq -1$. The equations for this model have been studied in detail in [148], and experimental constraints have been presented in, e.g. [149–152].

For a constant dark energy equation of state $w_{\rm de}$, the dark energy density is given by[2]

$$\rho(a) = \rho_0 \cdot a^{-3(1+w_{\rm de})}. \tag{2.1}$$

We immediately see that $\rho = $ const. for a cosmological constant with $w_{\rm de} = -1$. The Hubble parameter is given by

$$\frac{H(a)}{H_0} = \left[\Omega_r a^{-4} + \Omega_m a^{-3} + \Omega_\kappa a^{-2} + \Omega_\Lambda \cdot a^{-3(1+w_{\rm de})}\right]^{\frac{1}{2}}. \tag{2.2}$$

In this equation, Ω_r is the radiation density which includes photons and massless neutrinos, Ω_m the matter density, Ω_κ the curvature density (listed for completeness, even though the Boltzmann solver in `PyCosmo` currently only supports flat models) and Ω_Λ the dark energy density. In all cases, Ω_i is defined as the fraction of the energy density of the corresponding component today and the critical energy density $\rho_{\rm crit}$ of

---

[2] In the code, $w_{\rm de}$ is written as the parameter `w0`.

the Universe.

In the case of a cosmological constant, there are no dark energy perturbations. For $w_{\rm de} \neq -1$, we can write down the dark energy equations following [11, 128] and obtain

$$\begin{aligned}\dot{\delta}_{\rm de} &= -(1+w_{\rm de})(ku_{\rm de}+3\dot{\Phi}) - 3\frac{\dot{a}}{a}\left(\frac{\delta p_{\rm de}}{\delta \rho_{\rm de}} - w_{\rm de}\right)\delta_{\rm de} \\ \dot{u}_{\rm de} &= -\frac{\dot{a}}{a}\left(1-3w_{\rm de}\right)u_{\rm de} + \frac{\delta p_{\rm de}/\delta\rho_{\rm de}}{1+w_{\rm de}}k\delta_{\rm de} - k\sigma_{\rm de} + k\Psi,\end{aligned} \tag{2.3}$$

where all derivatives are with respect to the conformal time η and we use the conformal Newtonian gauge as in [12]. The anisotropic stress, $\sigma_{\rm de}$, vanishes, which deletes one term in the second perturbation equation.

In general, the sound speed $\tilde{c}^2_{\rm s,de} = \frac{\delta p_{\rm de}}{\delta \rho_{\rm de}}$ is a Gauge-dependent variable, which can be expressed in terms of the rest frame sound speed $c^2_{\rm s,de}$ and the adiabatic sound speed $c^2_{\rm a}$. The latter is equal to $w_{\rm de}$ for a constant dark energy equation of state [148]. Here, we use the expression [148]

$$\tilde{c}^2_{\rm s,de}\delta_{\rm de} = c^2_{\rm s,de}\delta_{\rm de} + 3\frac{d\ln a}{d\eta}(1+w_{\rm de})(c^2_{\rm s,de} - w_{\rm de})\frac{u_{\rm de}}{k}, \tag{2.4}$$

which is valid for a constant dark energy equation of state.

Inserting this expression in Eq. 2.3 we obtain

$$\begin{aligned}\dot{\delta}_{\rm de} &= -(1+w_{\rm de})(ku_{\rm de}+3\dot{\Phi}) - 3\frac{\dot{a}}{a}\left(c^2_{\rm s,de} - w_{\rm de}\right)\delta_{\rm de} - 9\left(\frac{\dot{a}}{a}\right)^2(1+w_{\rm de})(c^2_{\rm s,de} - w_{\rm de})\frac{u_{\rm de}}{k} \\ \dot{u}_{\rm de} &= -\frac{\dot{a}}{a}(1-3c^2_{\rm s,de})u_{\rm de} + \frac{c^2_{\rm s,de}}{1+w_{\rm de}}k\delta_{\rm de} + k\Psi.\end{aligned} \tag{2.5}$$

Compared to the system of equations in [12], the Einstein equations are modified to

$$k^2\Phi + 3\frac{\dot{a}}{a}\left(\dot{\Phi} - \frac{\dot{a}}{a}\Psi\right) = 4\pi G a^2\left[\rho_m\delta_m + 4\rho_r\Theta_{r0} + \rho_{\rm de}\delta_{\rm de}\right] \tag{2.6}$$

$$\dot{\Phi} - \frac{\dot{a}}{a}\Psi = -4\pi G\frac{a^2}{k}\left[\rho_m u_m + 4\rho_r\Theta_{r1} + \rho_{\rm de}(1+w_{\rm de})u_{\rm de}\right] \tag{2.7}$$

$$k^2(\Phi + \Psi) \quad = \quad -32\pi G a^2 \rho_r \Theta_{r2}. \tag{2.8}$$

In order to evolve the perturbation equations, we also need to define the initial conditions for these. We choose the adiabatic initial conditions from `CLASS`, outlined in [148] in the synchronous gauge, and then transform them into the conformal Newtonian gauge. We report the initial conditions for all the fields in Appendix 2.A.3.

2.3.2 Numerical implementation

In order to implement the new equations outlined in the previous section, we generated a new file for the symbolic Boltzmann equations, called `CosmologyCore_wcdm.py`. This can be used instead of the default equations file `CosmologyCore.py` for the ΛCDM model, by setting the `model` to `"wcdm"` as shown in Listing 2.1.

In `PyCosmo`, all derivatives are written with respect to $\ln a$ (in the code `lna`) instead of η. The perturbation equations for a ΛCDM model in this notation are detailed in 2.A.2. In the previous section, we presented the dark energy perturbation equations with respect to η. Using the conversion $\frac{d\ln a}{d\eta} = aH$, we can rewrite the two dark energy perturbation equations from Eq. 2.5 as

$$\begin{aligned}
\frac{d\delta_{\mathrm{de}}}{d\ln a} =& -(1+w_{\mathrm{de}})\left(\frac{k u_{\mathrm{de}}}{aH(a)} + 3\frac{d\Phi}{d\ln a}\right) \\
& - 3(c^2_{\mathrm{s,de}} - w_{\mathrm{de}})\left(\delta_{\mathrm{de}} + 3(1+w_{\mathrm{de}})\frac{aH(a)u_{\mathrm{de}}}{k}\right) \\
\frac{du_{\mathrm{de}}}{d\ln a} =& -(1-3c^2_{\mathrm{s,de}})u_{\mathrm{de}} + \frac{c^2_{\mathrm{s,de}}}{1+w_{\mathrm{de}}}\frac{k\delta_{\mathrm{de}}}{aH(a)} + \frac{k}{aH(a)}\Psi.
\end{aligned} \tag{2.9}$$

Then the two dark energy perturbation equations in Eq. 2.9 can be expressed with `SymPy` as

```
ddelta_de_dlan = - (1 + w0) * (k / (a * H) * u_de + 3* dPhi_dlan)
          - 3 * (cs_de2 - w0) * delta_de
```

```
          - 9 * (1 + w0) * (cs_de2 - w0) * (a * H) / k * u_de

du_de_dlan = - (1 - 3* cs_de2) * u_de
          + cs_de2 / (1 + w0) * k * delta_de / (a * H)
          + k / (a * H) * Psi
```

Listing 2.2: The two dark energy perturbation equations as implemented in `CosmologyCore_wcdm.py`.

`CosmologyCore_wcdm.py` also contains the background equations and the initial conditions for the linear perturbations.

2.3.3 Code comparisons

In Figure 2.1 we show the dark energy perturbations $\delta_{\rm de}$ and $u_{\rm de}$ as a function of scale factor a for three wavenumbers $k = 0.005, 0.05$ and 5 Mpc^{-1}, plotted both with `PyCosmo` as well as `CLASS`. We also display the relative difference between the evolution of the perturbations obtained by the two codes. The cosmology we consider is

- **wCDM**: $\{\Omega_\Lambda, \Omega_m, N_\nu, N_{\nu,m}, \Sigma m_\nu, w_{\rm de}\}$ = {0.69992, 0.3, 3.044, 0, 0, -0.9}

with all other parameters as specified in 2.A.4 and using the *precision* settings from 2.2.3. In general, we find good agreement between the codes. When the fields are highly oscillating around zero, we observe a degradation of the agreement, as expected, given the impact of step-size control and numerical precision of the solver in that regime. We also notice a discrepancy at initial time for small values of k, which is caused by the tight coupling approximation in `CLASS`. In Figure 2.2 we show the wCDM total matter power spectrum computed with the two codes for 200 log-spaced k values between 10^{-4} and 10 Mpc^{-1} at redshifts $z = 0$, $z = 1$ and $z = 5$. In general, we observe that the results for wCDM show the same level of agreement with `CLASS` as ΛCDM. The discrepancies tend to grow on large scales for $z = 5$ but remain below the 10^{-3} level. The same is observed for higher redshifts. In Section 2.6 we summarize the

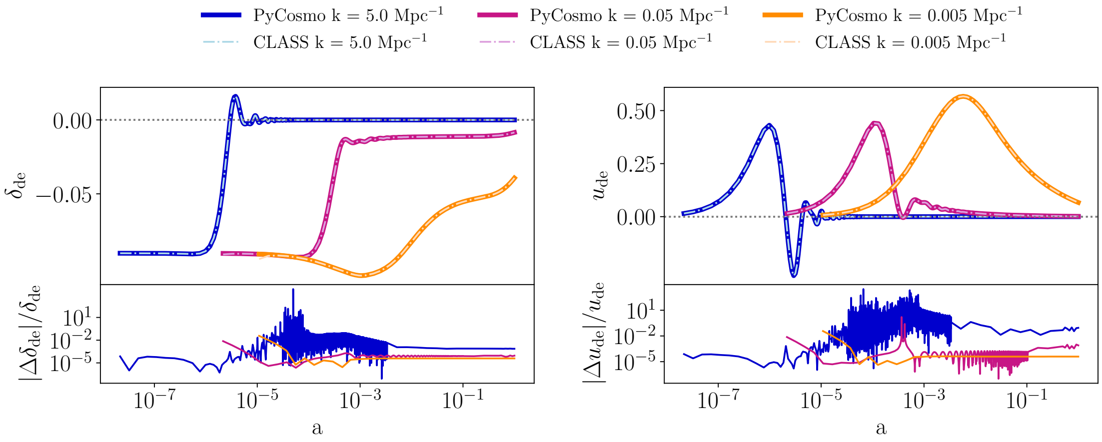


Figure 2.1: In the top left panel, we show the dark energy perturbation $\delta_{\rm de}$ as a function of scale factor a for a dark energy equation of state with $w = -0.9$ and sound speed $c^2_{\rm s,de} = 1$ at three values of the wavenumber $k = 0.005, 0.05$ and 5 Mpc^{-1}. In the top right panel, we show dark energy perturbation $u_{\rm de}$ as a function of scale factor a for the same wavenumbers. Perturbations computed with `PyCosmo` are displayed by thick lines, while `CLASS` values by dash-dotted lines. The bottom panels display relative differences between the two codes.

comparisons in terms of computing time and power spectrum relative difference for different k ranges at $z = 0$.

2.4 A complex model: massive neutrinos

2.4.1 Equations

Oscillation experiments provide evidence that neutrinos have mass (see e.g. [153–155] and also [156] for a recent global fit of neutrino oscillation data) and since their masses are imprinted onto cosmological observables, we need to include the evolution of light massive relics in the system of equations. This allows cosmological probes to constrain the properties of neutrinos, in particular the sum of the neutrino masses (see e.g. [16, 157, 158] for reviews on neutrino cosmology).

In this section, we present the implementation of the Einstein-Boltzmann equations

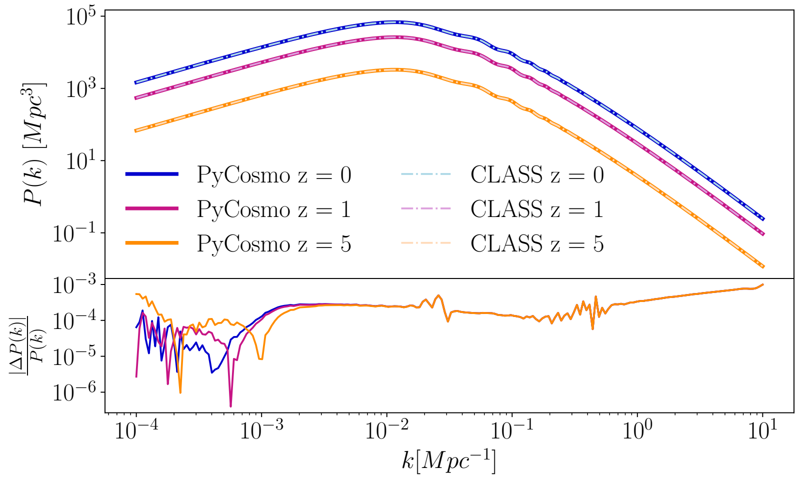


Figure 2.2: Total matter power spectrum at redshifts $z = 0$, $z = 1$ and $z = 5$ for a dark energy equation of state with $w = -0.9$ and sound speed $c^2_{\mathrm{s,de}} = 1$, plotted both with `PyCosmo` (thick solid lines) and `CLASS` (dash-dotted lines) for 200 log-spaced k values between 10^{-4} and 10 Mpc^{-1}. On the bottom panel we display the relative difference between the two codes.

for massive neutrinos into the `PyCosmo` Boltzmann solver. Massive neutrinos modify both the background evolution and the linear order perturbations. Qualitatively, massive neutrinos undergo a phase transition: they behave like radiation at early times, when they are fully relativistic, and shift to a matter-like behaviour at late times (see, e.g., [16, 159]). The transition happens smoothly through cosmic time and the dependence on mass, scale factor and momentum in the evolution of the distribution function prevents from integrating out the momentum dependence. This can be done only when considering approximations. For this reason, the inclusion of massive neutrinos in the Boltzmann equations is highly non trivial and has a strong impact on the size of the ODE system.

In this section we write the equations for $N_{\nu,m}$ massive neutrinos with degenerate masses and total neutrino mass sum Σm_ν, where the sum goes over the three neutrino mass eigenstates. A generalization to neutrinos with different masses is simply achieved by suppressing the $N_{\nu,m}$ in front of the equations and writing separate equations for each neutrino species with mass $m_{\nu,i}$. In `PyCosmo` we introduce degenerate massive neutrinos. The introduction of neutrino hierarchies is left as future development.

The massive neutrino density can be written as

$$\rho_{\nu,m}(a) = 4\pi a^{-4}\int_0^\infty dq q^2 f_0(q)\epsilon, \tag{2.10}$$

where $q = ap$, with p the proper momentum, related to the 4-momentum by $P_i = a(1-\Phi)p_i$ in the Newtonian conformal gauge. Then $\epsilon = \sqrt{q^2+a^2m^2}$ is the proper energy measured by a comoving observer multiplied by the scale factor and $f_0(q)$ is the Fermi-Dirac distribution

$$f_0(q) = \frac{g_s}{(2\pi\hbar)^3}\frac{1}{e^{q/k_b T_0}+1}, \tag{2.11}$$

where g_s is the spin degeneracy factor that equals $2N_{\nu,m}$ in the case of degenerate neutrinos. T_0 is the temperature of the Cosmic Neutrino Background today expressed in units of the CMB temperature $T_0 = T_{\nu,m}T_{\mathrm{CMB}}$, $T_{\nu,m} = \left(\frac{4}{11}\right)^{1/3}$ for neutrinos that undergo instantaneous decoupling and k_b is the Boltzmann constant. The Friedmann equation results in

$$\frac{H(a)}{H_0} = \left[(\Omega_r + \Omega_{\nu,m}(a))a^{-4} + \Omega_m a^{-3} + \Omega_\kappa a^{-2} + \Omega_\Lambda\right]^{\frac{1}{2}} \tag{2.12}$$

where Ω_r still includes massless neutrinos if present (denoted with the subscript ν such that $\Omega_r = \Omega_\gamma + \Omega_\nu$), whereas the massive neutrino energy density is $\Omega_{\nu,m}(a)$. Note that we factor out the a^{-4} term from $\Omega_{\nu,m}(a)$ for similarity with the other energy densities, but $\Omega_{\nu,m}(a)$ still contains a dependency on the scale factor a, differently from the other species, since the proper energy ϵ contains a factor of a that we cannot integrate out (see equation 2.10, $\Omega_{\nu,m}(a)a^{-4} = \rho_{\nu,m}(a)/\rho_{\mathrm{crit}}$).

The massive neutrino perturbations arise from a linear expansion of the distribution function $f(x,q,\hat{q},\eta) = f_0(q)\left[1+\mathcal{M}(x,q,\hat{q},\eta)\right]$ around the Fermi-Dirac distribution $f_0(q)$. The function $\mathcal{M}(x,q,\hat{q},\eta)$ is Fourier transformed and expanded in a Legendre series as

$$\mathcal{M}(\vec{k},\hat{q},q,\eta) = \sum_{\ell=0}^{\infty} (-i)^{\ell}(2\ell+1)\mathcal{M}_{\ell}(\vec{k},q,\eta)P_{\ell}(\mu) \tag{2.13}$$

with $\mu = \hat{k}\cdot\hat{q}$, $\mathcal{P}_{\ell}$ the Legendre polynomials and $\mathcal{M}_{\ell}$ defined as

$$\mathcal{M}_{\ell} = \frac{1}{(-i)^{\ell}}\int_{-1}^{1}\frac{d\mu}{2}\mathcal{P}_{\ell}(\mu)\mathcal{M}(\mu). \tag{2.14}$$

The massive neutrino Boltzmann equations are then derived similarly to those of the ultra-relativistic fields, setting the collision term to 0, since neutrinos are only weakly interacting. Using the definition of $\mathcal{M}_{\ell}$ to express the Boltzmann equations as a hierarchy of moments, we obtain

$$\dot{\mathcal{M}}_0 = -\frac{qk}{\epsilon}\mathcal{M}_1 - \dot{\Phi}\frac{d\ln f_0}{d\ln q} \tag{2.15}$$

$$\dot{\mathcal{M}}_1 = \frac{qk}{3\epsilon}(\mathcal{M}_0 - 2\mathcal{M}_2) - \frac{\epsilon k}{3q}\Psi\frac{d\ln f_0}{d\ln q} \tag{2.16}$$

$$\dot{\mathcal{M}}_{\ell} = \frac{qk}{(2\ell+1)\epsilon}\left[\ell\mathcal{M}_{\ell-1} - (\ell+1)\mathcal{M}_{\ell+1}\right], \ \ \ell \geq 2. \tag{2.17}$$

The hierarchy is truncated at a multipole ℓ_{max} (`l_max_mnu` in the code) when solving the system of equations numerically with a hierarchy truncation from [11], which is analogous to that for photons and massless neutrinos

$$\dot{\mathcal{M}}_{\ell_{max}} \simeq \frac{qk}{\epsilon}\mathcal{M}_{\ell_{max}-1} - \frac{(\ell_{max}+1)}{\eta}\mathcal{M}_{\ell_{max}}. \tag{2.18}$$

Massive neutrinos also modify the Einstein equations due to the extra terms in the stress-energy tensor. These now read

$$\begin{gathered} k^2(\Psi+\Phi) = -12\left(\frac{H_0}{a}\right)^2\left(\Omega_{\gamma}\Theta_2 + \Omega_{\nu}\mathcal{N}_2 + \frac{N_{\nu,m}}{(2\pi)^2\rho_{\text{crit}}}\int q^2 dq\frac{q^2}{\epsilon}\frac{\mathcal{M}_2}{e^{q/k_bT_0}+1}\right) \qquad (2.19) \\ k^2\Phi + 3\frac{\dot{a}}{a}\left(\dot{\Phi} - \frac{\dot{a}}{a}\Psi\right) = \frac{3}{2}(H_0a)^2\left[\Omega_m\delta_m a^{-3} + 4a^{-4}\left(\Omega_r\Theta_{r0} + \frac{N_{\nu,m}}{(2\pi)^2\rho_{\text{crit}}}\int q^2 dq\frac{\epsilon\mathcal{M}_0}{e^{q/k_bT_0}+1}\right)\right], \end{gathered}$$

where $\Omega_m\delta_m$ is a shortcut for $\Omega_{dm}\delta + \Omega_b\delta_b$, $\Omega_r\Theta_{r0} = \Omega_{\gamma}\Theta_0 + \Omega_{\nu}\mathcal{N}_0$ and ρ_{crit} is the critical energy density of the Universe. The massive neutrino quantities we are interested in are the density fluctuation, the fluid velocity and the shear stress which are computed by integrating over the moments $\mathcal{M}_\ell$:

$$\begin{aligned}\delta_{\nu,m} &= \frac{\int q^2 dq \epsilon f_0(q) \mathcal{M}_0}{\int q^2 dq \epsilon f_0(q)} \\ u_{\nu,m} &= \frac{\int q^2 dq q f_0(q) \mathcal{M}_1}{\int q^2 dq \epsilon f_0(q) + \frac{1}{3} \int q^2 dq \frac{q^2}{\epsilon} f_0(q)} \\ \sigma_{\nu,m} &= \frac{2}{3} \frac{\int q^2 dq \frac{q^2}{\epsilon} f_0(q) \mathcal{M}_2}{\int q^2 dq \epsilon f_0(q) + \frac{1}{3} \int q^2 dq \frac{q^2}{\epsilon} f_0(q)}.\end{aligned} \tag{2.20}$$

Note that in the PyCosmo implementation, we substitute q with $q' = \frac{q}{k_b T_0}$ for convenience.

2.4.2 Numerical implementation

The implementation of the massive neutrino equations uses `sympy2c` similarly to the PyCosmo implementation of ΛCDM and wCDM. The background integral over momentum q is computed using indefinite numerical integration from `sympy2c` , whereas the integrals at perturbation level use a Gauss-Laguerre quadrature integration scheme in our symbolic representation of the ODE system, since we need to evolve a finite number of equations. This follows the approach used in [147]. The number of discrete q values is governed by the parameter `mnu_relerr`, which sets the relative difference between the Gauss-Laguerre integration of a test function ($\sum_{n=2}^{n=4} q^n f_0(q)$) with respect to its analytical result. This parameter is specified using `PyCosmo.build` since it affects the size of the ODE system and thus also code generation. We introduce an additional parameter influencing code generation in addition to the truncation parameter `l_max` of the photon and massless neutrino hierarchies: `l_max_mnu` that truncates the Legendre series $\mathcal{M}_\ell$ as described above.

The implementation of massive neutrino cosmologies results in a large systems of equations and thus C functions with millions of lines of generated code, challenging the optimizer of the used C compiler. To mitigate significantly compilation time and memory requirements, we enable `sympy2c` to use the matrix splitting feature (enabled by specifying the `splits` parameter and `reorder=True` in `PyCosmo.build`) described in Section 2.2. The user can also pass a `compilation_flags` parameter, that enables or disables compiler optimizations of the `C` code and has a diametrical effect on compilation vs runtime.

We implement initial conditions for massive neutrinos that match the adiabatic initial conditions in `CLASS` and `COSMICS` and can be triggered using the `initial_conditions` parameter. We report the equations in Appendix 2.A.3.

2.4.3 Code comparisons

One of the key effects of massive neutrinos on cosmological observables is the suppression of small scale matter overdensities due to neutrino free streaming [157]. In Figure 2.3 we show the total matter power spectrum obtained with `PyCosmo` and `CLASS` for 200 log-spaced k values between 10^{-4} and 10 Mpc^{-1} at redshifts $z = 0$, $z = 1$ and $z = 5$ with the following cosmological parameters:

- **degenerate $\Sigma \mathbf{m}_\nu$ = 60 meV**: $\{\Omega_m, N_\nu, N_{\nu,m}, \Sigma m_\nu, w_{\mathrm{de}}\} = \{0.29869, 0.00440, 3, 0.06, -1\}$[3].

All other parameters are set to default values (see 2.A.4) and we use the *precision* settings for the two codes (see 2.2.3). In the bottom panel we display the relative difference between `PyCosmo` and `CLASS`. In Appendix 2.A.5 we outline the equation of the total matter power spectrum, following a fully general relativistic treatment in the presence of massive neutrinos [138, 160–163]. In general, we observe a very good agreement, with a maximum relative discrepancy of 5×10^{-4} on intermediate scales. The redshift evolution does not impact the agreement. We also show the suppression of the total matter power spectrum with respect to the power spectrum with massless neutrinos in Figure 2.4 both for Σm_ν = 60 meV and 120 meV (Ω_m=0.29737), when

---

[3] Ω_m and N_ν are determined by fixing $\Omega_{m,tot} = \Omega_m + \Omega_{\nu,m} = 0.3$ and $N_{\mathrm{eff}} = N_\nu + {T_{\nu,m}}^4 \left(\frac{4}{11}\right)^{-\frac{4}{3}} N_{\nu,m} = 3.044$ in order to look at the effects of neutrino mass on the total matter power spectrum in Figure 2.4.

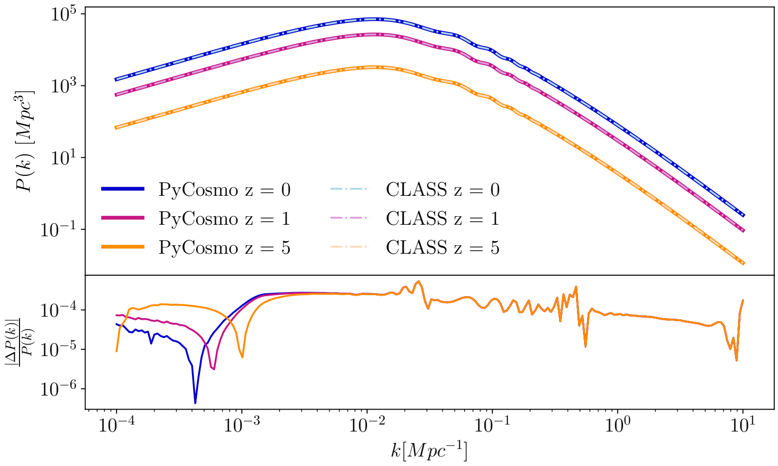


Figure 2.3: Total matter power spectrum at redshifts $z = 0$, $z = 1$ and $z = 5$ for a cosmology with three massive neutrinos with $\Sigma m_\nu = 60$ meV, plotted both with `PyCosmo` and `CLASS` for 200 log-spaced k values between 10^{-4} and 10 Mpc^{-1}. In the bottom panel, we display the relative difference between the two codes.

keeping $\Omega_{m,tot} = \Omega_m + \Omega_{\nu,m}$ fixed to $\Omega_{m,tot} = 0.3$ and $N_{\mathrm{eff}} = 3.044$. The cosmological parameters for the ΛCDM model are set to:

- **ΛCDM**: $\{\Omega_m,\ N_\nu,\ N_{\nu,m},\ \Sigma m_\nu,\ w_{\mathrm{de}}\} = \{0.3, 3.044, 0, 0., \text{-}1\}$.

On the left panel of the figure, we show the suppression of the power spectrum computed with `PyCosmo` and `CLASS`, using the default *precision* settings and the relative difference between the two codes. The suppression has a large discrepancy at low k values when it is approaching and crossing zero. Furthermore, there is a ~ 1% difference on small scales (large k values), where the effects of the hierarchy truncation and the approximations are most dominant. We verify that this discrepancy is reduced when matching the `l_max` parameters in the two codes (`l_max_g = l_max_pol_g = l_max_ur = l_max_ncdm = 50` in `CLASS`, `PyCosmo` remains in *precision* settings) and suppressing the ultra-relativistic fluid approximation in `CLASS`, as shown in the right panel of Figure 2.4.

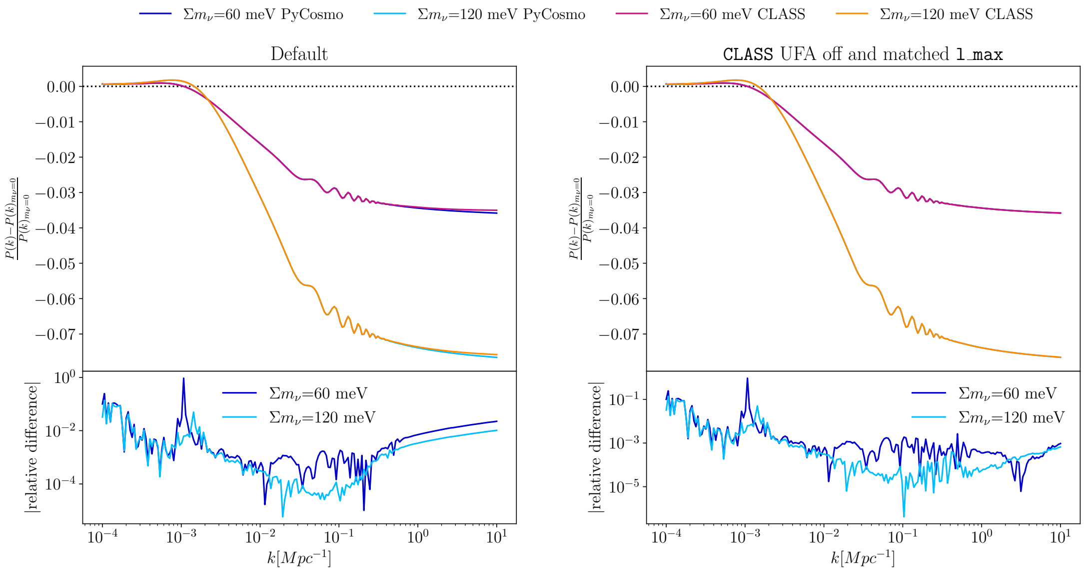


Figure 2.4: Suppression of the total matter power spectrum for two cosmologies with three neutrinos and $\Sigma m_\nu = 60$ and 120 meV, respectively, compared to a ΛCDM cosmology with massless neutrinos only computed with `PyCosmo` and `CLASS`. For all cosmologies, we set the total matter density to $\Omega_{m,tot} = 0.3$ and $N_{\rm eff} = 3.044$. The left panel displays the power spectrum and the relative difference with the default *precision* settings for `CLASS` and `PyCosmo`. In the right plot, we display the same quantities when setting `l_max = 50` for all the radiation fields and massive neutrinos and switching off the UFA in `CLASS`.

2.5 An approximation scheme: radiation streaming approximation

2.5.1 Equations

After decoupling, photons and massless neutrinos behave approximately like test particles free-streaming in the gravitational field determined by the massive components, making it possible to derive a non-oscillatory solution of the inhomogeneous Boltzmann equations inside the Hubble radius. This approximation, called the Radiation Streaming Approximation (RSA), was introduced in the Newtonian gauge by [164] and in the synchronous gauge by [17]. This treatment allows both to avoid unphysical oscillations resulting from the hierarchy truncation and to speed up the integration, which is slowed down by fast late time oscillations of the radiation fields especially on small scales. At late times, the approximation does not need to be precise, since the contribution of the radiation energy density to the overall energy density is negligible. This approximation only impacts the linear perturbations. The evolution of the relativistic fields can be written as

$$\begin{aligned}
\Theta_0 &= \Phi + \frac{\dot{\tau} u_b}{k} \\
\Theta_1 &= -\frac{2}{k}\dot{\Phi} + \frac{\dot{\tau}}{k}\left(\frac{aH}{k}u_b - c_b^2\delta_b + \Phi\right) \\
\mathcal{N}_0 &= \Phi \\
\mathcal{N}_1 &= -\frac{2}{k}\dot{\Phi} \\
\Theta_2 &= \Theta_{P0,1,2} = \mathcal{N}_2 = 0,
\end{aligned} \tag{2.21}$$

with all the higher order multipoles set to 0. This approximation is switched on when two conditions are satisfied, following the `CLASS` [17] scheme:

- `rsa_trigger_k_eta` $\leq k\eta$
- `rsa_trigger_taudot_eta` $\leq -(\dot{\tau}\eta)^{-1}$,

corresponding to decoupled radiation within the horizon.

2.5.2 Numerical implementation

In order to implement the radiation streaming approximation within `PyCosmo` we use a functionality from `sympy2c` to switch between two different ODEs at a dynamically computed time point.

This requires:

- Two ODE systems specified as `sympy2c OdeFast` objects. In our case these will be the symbolic representations of the model of interest (ΛCDM, ΛCDM with massive neutrinos or wCDM) and its equivalent RSA system. Note that the two systems can have different dimensions.

- A `switch_time` function which determines at which time to switch from the first system of equations to the second. In the RSA implementation we use the switching conditions described above.

- A `switch` function, computing the initial conditions for the second system of equations from the state of the first system before and at the switching time. In the RSA implementation, this function just discards the matrix entries for the fields Θ_i, Θ_{Pi} and $\mathcal{N}_i$, $i \geq 0$ since Θ_0, Θ_1, $\mathcal{N}_0$ and $\mathcal{N}_1$ have analytical expressions and thus do not need initial conditions.

- A `merge` function which specifies how to combine the matrix valued results from both numerical solutions into a final matrix. Our RSA implementation keeps the full matrix from the solution of the full system of equations and extends the matrix from the `RSA` solution using the analytical formula for Θ_i and $\mathcal{N}_i$ for $i \leq 1$. All other entries are set to 0 for $i \geq 2$ and for all the polarization terms Θ_{Pi}.

In order to use the RSA in `PyCosmo` one needs to pass the `rsa = True` flag to `PyCosmo.build` for any of the models. This will switch the `CosmologyCore_`*model*`.py` equation file to a `CosmologyCore_`*model*`_rsa.py` file containing the RSA equations for the radiation fields and all the other equations of the system.

2.5.3 Internal code consistency

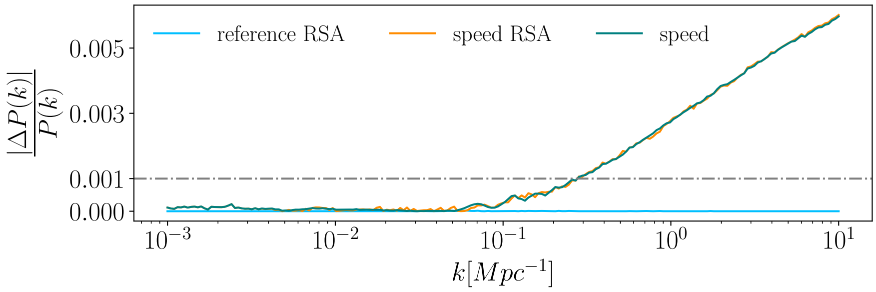
The main achievement obtained by implementing the RSA is a significant reduction of the computation time for the fields, especially for high values of k, as we will discuss in detail in the next section. This is especially true when choosing the ΛCDM and wCDM models, where most of the system of ODEs consists of radiation perturbations. In Figure 2.5, we show the effects of the RSA and of the hierarchy truncation on the

ΛCDM matter power spectrum (**ΛCDM** cosmological parameters as specified in the Section 2.4.3). To show the effects of the RSA alone we display in light blue in both panels of Figure 2.5 the relative difference between a *reference* power spectrum with `l_max` = 200 and `atol` = `rtol` = 10^{-6} and the same power spectrum computed with RSA (`rsa_trigger_taudot_eta` = 100, `rsa_trigger_k_eta` = 240) for 200 log-spaced k values between 10^{-3} and 10 Mpc^{-1}. We observe that the oscillations around zero correspond to a relative difference of the order of 0.001%, but the computation time is reduced by approximately 80%. In Figure 2.5 we also display the relative difference between the power spectrum computed using `PyCosmo` *speed* (panel a) and *precision* (panel b) settings from Section 2.2.3, both with and without RSA, and the *reference* power spectrum just described.

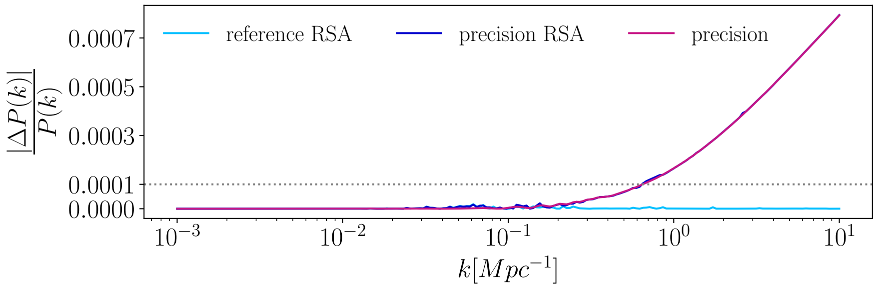
Reducing the `l_max` parameter to 50 (*precision* settings) increases the discrepancy with the reference power spectrum to ~0.08% for the full system, regardless of whether the RSA is turned on or not. For the *speed* settings, the discrepancy is ~0.6% both when using or not using the RSA. The computation time is again reduced by roughly 80% when using RSA compared to solving the full equation system with the same `l_max`, meaning that the approximation is essential to reduce the computational time of the perturbations for mid to high k values without sacrificing the accuracy. The unphysical reflection of power caused by the hierarchy truncation dominates on small scales, making the inaccuracy introduced by the approximation completely negligible.

2.6 Agreement and performance comparison with `CLASS`

In this section, we present benchmarks of the `PyCosmo` Boltzmann solver for all the models described in the previous sections, both in terms of relative difference to `CLASS` and in terms of computing times. The cosmological parameters that are modified in each model are specified in the Λ**CDM**, **wCDM** and **degenerate** $\mathbf{\Sigma m_\nu}$ **= 60 meV** parameter settings in the previous sections, while the fixed parameters are reported in Appendix 2.A.4. In order to perform a comparison purely on the Boltzmann solver, we read in the `CLASS` recombination files for each model and set the same initial



(a) Relative difference between `PyCosmo` *reference* and *speed* settings with (orange) and without (green) RSA.



(b) Relative difference between `PyCosmo` *reference* and *precision* settings with (blue) and without (purple) RSA.

Figure 2.5: Relative difference between the ΛCDM matter power spectrum computed with `l_max` = 200 and `atol` = `rtol` = 10^{-6} (*reference*) and the same power spectrum computed with `PyCosmo` *speed* (top panel) and *precision* (bottom panel) settings with and without RSA. We also display in both panels in light blue the relative difference between the *reference* matter power spectrum with and without RSA. The grey horizontal lines correspond to a 10^{-3} (dash-dotted) and 10^{-4} (dotted) precision level.

conditions[4]. All the computations are carried out on a single core on a full node of the ETH Zurich Euler cluster[5], by disabling parallel execution in `CLASS` and not enabling parallel computation for the `PyCosmo` power spectrum. We run on a full node on the cluster, instead of on a laptop, in order to only run the Boltzmann solver and not get impacted by other processes being executed by the operating system. In the case of `PyCosmo`, we only report the time necessary for the power spectrum

---

[4] Note that `initial_conditions=class` in `PyCosmo` corresponds to adiabatic initial conditions in `CLASS`.

[5] Euler is a HPC Cluster of ETH Zurich, a description of the hardware in an Euler VI node can be found at https://scicomp.ethz.ch/wiki/Euler (two 64-core AMD EPYC 7742 processors).

Model	k_{max}	*Speed* settings, time [s]			*Precision* settings, time [s]		
		`PyCosmo`	`PyCosmo*`	`CLASS`	`PyCosmo`	`PyCosmo*`	`CLASS`
ΛCDM	1 Mpc^{-1}	1.26	0.23	0.42	3.80	1.05	2.02
ΛCDM	10 Mpc^{-1}	8.80	0.44	0.80	20.5	2.20	5.28
wCDM	1 Mpc^{-1}	1.32	0.65	1.29	3.84	1.55	2.86
wCDM	10 Mpc^{-1}	9.08	0.82	4.91	20.93	2.72	10.18
M_νCDM	1 Mpc^{-1}	54.54	29.04	10.19	237.26	154.93	105.87
M_νCDM	10 Mpc^{-1}	357.24	98.52	13.78	1337.32	471.22	417.95

Table 2.2: Best execution time from three executions on a full Euler VI node. `PyCosmo*` stands for `PyCosmo` with RSA.

computation, which does not include the compilation time for the first time the `C/C++` code is generated for each model. Most models also require a second compilation (recompilation) that applies permutations to the existing equations in order to enable the use of specialized solvers instead of the standard solver as explained in Section 2.2.1. The runtime we report is the best time obtained in three executions.

We begin by comparing the runtime between the computation of the matter power spectrum in `PyCosmo`, with and without turning on the RSA, and `CLASS`, both for the *speed* and *precision* settings (defined in Section 2.2.3). In Table 2.2, we display the time necessary to compute the power spectrum with `PyCosmo` and `CLASS` for a 100 log-spaced k values between k_{min} and k_{max}. We fix $k_{min} = 10^{-4}$ Mpc^{-1}, and set k_{max} first to 1 Mpc^{-1} and then to 10 Mpc^{-1}, where the effects of the radiation streaming approximation are most evident.

We observe that, while `PyCosmo` achieves a slower runtime than `CLASS` before introducing the RSA, the RSA reverts the situation for all the models, except for massive neutrinos. This happens despite the presence of further approximations in the `CLASS` implementation. Previous versions of `PyCosmo` achieved a comparable execution time with `CLASS` without physical approximations [12]. This was due to the reduction of the dynamic range of the time step based on a consistency relation of the Einstein equations. The adaptive control of the time step was highly optimized for ΛCDM and proved difficult to extend to more general models. The new approach has also the

Model	precision	k range [Mpc^{-1}]				
		$10^{-4}-10^{-3}$	$10^{-3}-10^{-2}$	$10^{-2}-10^{-1}$	$10^{-1}-1$	$1-10$
ΛCDM	*speed*	0.0006	0.0007	0.0007	0.0031	0.0077
	precision	0.0001	0.0003	0.0005	0.0003	0.0008
wCDM	*speed*	0.0006	0.0007	0.0008	0.0031	0.0078
	precision	0.0002	0.0003	0.0005	0.0003	0.0008
M_νCDM	*speed*	0.0004	0.0010	0.0010	0.0046	0.0146
	precision	0.0001	0.0003	0.0005	0.0002	0.0001

Table 2.3: Maximum relative differences in k ranges between `PyCosmo` and `CLASS` for given models and precision settings.

advantage of reducing the number of permutations necessary to create the optimized code.

In the case of massive neutrinos, the size of the system determines a considerable decline in the performance of both codes. `PyCosmo` is significantly slower than `CLASS` when using the *speed* settings, mainly due to the presence of the fluid approximation for non cold dark matter in `CLASS`. For the *precision* settings, the time needed for the computation is comparable for `CLASS` and `PyCosmo`. The ODE system used by `CLASS` is larger in this case, since the `tol_ncdm` parameters are set to 10^{-10} in `pk_ref.pre`, and the three neutrinos are treated independently despite having the same mass. We do not decrease `mnu_relerr`, equivalent to `tol_ncdm`, in `PyCosmo` since we do not deem it necessary to achieve the desired precision. The error caused by sampling less q values is always subdominant compared to the hierarchy truncation, due to the small contribution of massive neutrinos to the overall matter density.

We compare the numerical results of `PyCosmo` and `CLASS` in Table 2.3. We report the maximum relative difference between the matter power spectra in `PyCosmo` and `CLASS` in five k ranges for the different models and precision settings. The RSA is not relevant here because it leads to a 10^{-5} relative difference, when compared to the full computation with the same `l_max`, as shown for ΛCDM in the previous section.

We start by noting that the size of the relative differences is comparable in all models when using the same k ranges, with the exception of the model with massive neutrinos for $k > 0.1$ Mpc^{-1}. We also see that the difference in precision between *speed* and *precision* settings is dominant on scales $k > 0.1$ Mpc^{-1}, where the truncation effects have the largest impact. The *precision* settings lead to a relative difference to `CLASS` that is better than 0.1% on all scales $k < 10$ Mpc^{-1}, while the *speed* settings lead to a relative difference of order 0.5 – 1% beyond $k = 0.1$ Mpc^{-1}. This is acceptable, especially since `cl_permille.pre` comes with no guarantees for scales $k > 1$ hMpc^{-1}.

2.7 Conclusion

In this chapter, we demonstrated how the `PyCosmo` Boltzmann solver can be easily modified to include extensions of the ΛCDM cosmological model and approximation schemes, by taking advantage of the `SymPy` symbolic implementation of equations. The symbolic expressions are translated into optimized `C/C++` code by the `sympy2c` package presented in [141]. In this way, `PyCosmo` combines the speed of `C/C++` with the user-friendliness of symbolic `Python`.

We first presented two cosmological model extensions: dark energy with a constant equation of state, which is a minimal modification of ΛCDM, and massive neutrinos, which enlarge considerably the ODE system and comprise numerical integrations, constituting a more complex extension. The inclusion of these models makes `PyCosmo` more widely applicable for constraining cosmology. We also implemented an approximation scheme, the radiation streaming approximation. In order to trigger the approximation, `sympy2c` includes a functionality to switch between two different ODE systems when a condition is verified. The radiation streaming approximation makes the solution of the ODE system considerably faster, up to an 80% speed-up, since it suppresses oscillations of the radiation fields for large k values. The errors introduced by the approximation are largely sub-dominant compared to the artificial power reflection induced by the hierarchy truncation. For convenience, we presented a conversion table between common conventions for cosmological perturbations (Appendix 2.A.1) and a clarification of the computation of the total matter power spectrum with $\mathcal{A}_s$ normalization (Appendix 2.A.5).

We compared the numerical results obtained by computing the total matter power spectrum with `CLASS` and found an agreement better than 0.1% with high precision settings. With more relaxed precision settings, we found an agreement of 0.5% for scales $k < 1\mathrm{Mpc}^{-1}$ for all models. The `PyCosmo` Boltzmann solver achieves precision and speed that is comparable to `CLASS`, while not relying on physical approximations (such as tight coupling and ultra relativistic fluid approximation) other than the radiation streaming approximation introduced in this work. In the future, we plan to include more beyond ΛCDM models in `PyCosmo`. Possible extensions include time-varying dark energy, early dark energy [137, 165], curvature [166], dark matter models [167], such as axions and fuzzy dark matter, and extensions of the neutrino sector. We believe that our symbolic implementation will be applicable for most model extensions, with some refactoring needed when a model introduces an ODE system already at background level (for example in the case of scalar field models).

2.A Appendix of chapter 2

2.A.1 Notation table

Ma-Bertschinger	`PyCosmo`	Meaning
a	a	scale factor
k	k	wavenumber of Fourier mode
P_i	P_i	conjugate momentum to x_i
p_i	p_i	proper momentum
q_i	q_i	ap_i
ϵ_i	ϵ_i	$(q_i^2 + am^2)^{1/2}$
δ_γ	$4\Theta_0$	photons overdensity
θ_γ	$3k\Theta_1$	photons velocity divergence
σ_γ	$2\Theta_2$	photons shear stress
$F_{\gamma\ell}$	$4\Theta_\ell$	ℓth Legendre component of photons perturbations
δ_{ur}	$4\mathcal{N}_0$	massless neutrinos overdensity
θ_{ur}	$3k\mathcal{N}_1$	divergence of massless neutrinos velocity
σ_{ur}	$2\mathcal{N}_2$	shear stress of massless neutrino fluid
ψ, ϕ	$\Psi, -\Phi$	Newtonian gravitational potentials
τ	η	conformal time
$\tau_c{=}1/\kappa'$	$-\dot{\tau}$	Thomson scattering rate
θ_c	$ku = ikv$	velocity divergence of dark matter
θ_b	$ku_b = ikv_b$	velocity divergence of baryons
$\Psi_1, \Psi_2, \Psi_\ell$	$\mathcal{M}_1, \mathcal{M}_2, \mathcal{M}_\ell$	Legendre components of massive neutrinos perturbations
δ_h	$\delta_{\nu,m}$	massive neutrinos overdensity
θ_h	$ku_{\nu,m}$	massive neutrinos velocity divergence
σ_h	$\sigma_{\nu,m}$	massive neutrinos shear stress

Table 2.A.1: Relations between the notation in Ma-Bertschinger and `PyCosmo`. Ma-Bertschinger notation is used in both the Boltzmann solvers `COSMICS` and `CLASS`. Some exceptions include θ_{ur} that is denoted θ_r in `COSMICS` and all the massive neutrino quantities that are denoted δ_{ncdm}, θ_{ncdm} and σ_{ncdm} in `CLASS` (`ncdm` = non cold dark matter) and δ_n, θ_n and σ_n in `COSMICS`.

2.A.2 Linear perturbations in ln(a)

In `PyCosmo` we use $\ln(a)$ (from here on simply lna) as the independent variable of the Einstein-Boltzmann ODE system. The relation between $\ln(a)$ and η is simply: $\frac{d\ln(a)}{d\eta} = aH$. We report the Einstein-Boltzmann equations in ΛCDM, as used in the code, in the following:

- Einstein equations:

$$\begin{aligned}
\Psi &= -\Phi - 12\left(\frac{H_0}{ka}\right)^2\left(\Omega_\gamma\Theta_2 + \Omega_\nu\mathcal{N}_2\right)\\
\Pi &= \Theta_2 + \Theta_{P,0} + \Theta_{P,2}\\
\frac{d\Phi}{d\ln a} &= \Psi - \left(\frac{k}{aH}\right)^2\frac{\Phi}{3} + \frac{1}{2}\left(\frac{H_0}{H}\right)^2\left(\left(\Omega_{dm}\delta + \Omega_b\delta_b\right)a^{-3} + 4a^{-4}\left(\Omega_\gamma\Theta_0 + \Omega_\nu\mathcal{N}_0\right)\right)
\end{aligned}$$

- Dark matter:

$$\begin{aligned}
\frac{d\delta}{d\ln a} &= -\frac{k}{aH}u - 3\frac{d\Phi}{d\ln a}\\
\frac{du}{d\ln a} &= -u + \frac{k}{aH}\Psi
\end{aligned}$$

- Baryonic matter:

$$\begin{aligned}
\frac{d\delta_b}{d\ln a} &= -\frac{k}{aH}u_b - 3\frac{d\Phi}{d\ln a}\\
\frac{du_b}{d\ln a} &= -u_b + \frac{k}{aH}\Psi + \frac{\dot{\tau}}{RaH}\left(u_b - 3\Theta_1\right) + \frac{k}{aH}{c_s}^2\delta_b
\end{aligned}$$

- Photons temperature:

$$\begin{aligned}
\frac{d\Theta_0}{d\ln a} &= -\frac{k}{aH}\Theta_1 - \frac{d\Phi}{d\ln a}\\
\frac{d\Theta_1}{d\ln a} &= \frac{k}{3aH}\left(\Theta_0 - 2\Theta_2 + \Psi\right) + \frac{\dot{\tau}}{aH}\left(\Theta_1 - \frac{u_b}{3}\right)\\
\frac{d\Theta_2}{d\ln a} &= \frac{k}{5aH}\left(2\Theta_1 - 3\Theta_3\right) + \frac{\dot{\tau}}{aH}\left(\Theta_{P,0} - \frac{\Pi}{2}\right)
\end{aligned}$$

For $\ell > 2$

$$\frac{d\Theta_\ell}{d\ln a} = \frac{k}{aH(2\ell+1)}\left(\ell\Theta_{\ell-1} - (\ell+1)\Theta_{\ell+1}\right) + \frac{\dot{\tau}}{aH}\Theta_\ell$$

- Photons polarization:

$$\begin{aligned}
\frac{d\Theta_{P,0}}{d\ln a} &= -\frac{k}{aH}\Theta_{P,1} + \frac{\dot{\tau}}{aH}\left(\Theta_{P,0} - \frac{\Pi}{2}\right)\\
\frac{d\Theta_{P,1}}{d\ln a} &= \frac{k}{3aH}\left(\Theta_{P,0} - 2\Theta_{P,2}\right) + \frac{\dot{\tau}}{aH}\Theta_{P,1}\\
\frac{d\Theta_{P,2}}{d\ln a} &= \frac{k}{5aH}\left(2\Theta_{P,1} - 3\Theta_{P,3}\right) + \frac{\dot{\tau}}{aH}\left(\Theta_{P,2} - \frac{\Pi}{10}\right)
\end{aligned}$$

For $\ell > 2$

$$\frac{d\Theta_{P,\ell}}{d\ln a} = \frac{k}{aH(2\ell+1)}\left(\ell\Theta_{P,\ell-1} - (\ell+1)\Theta_{P,\ell+1}\right) + \frac{\dot{\tau}}{aH}\Theta_{P,\ell}$$

- Massless neutrinos:

$$\begin{aligned}
\frac{d\mathcal{N}_0}{d\ln a} &= -\frac{k}{aH}\mathcal{N}_1 - \frac{d\Phi}{d\ln a}\\
\frac{d\mathcal{N}_1}{d\ln a} &= \frac{k}{3aH}\left(\mathcal{N}_0 - 2\mathcal{N}_2 + \Psi\right)
\end{aligned}$$

For $\ell \geq 2$

$$\frac{d\mathcal{N}_\ell}{d\ln a} = \frac{k}{(2\ell+1)aH}\left(\ell\mathcal{N}_{\ell-1} - (\ell+1)\mathcal{N}_{\ell+1}\right)$$

The hierarchy truncations for relativistic species read:

$$\begin{aligned}
\frac{d\Theta_{\ell_{max}}}{d\ln a} &= \frac{1}{aH}\left(k\Theta_{\ell_{max}-1} - \left(\frac{(\ell_{max}+1)}{\eta} - \dot{\tau}\right)\Theta_{\ell_{max}}\right)\\
\frac{d\Theta_{P,\ell_{max}}}{d\ln a} &= \frac{1}{aH}\left(k\Theta_{P,\ell_{max}} - \left(\frac{(l_{max}+1)}{\eta} - \dot{\tau}\right)\Theta_{P,\ell_{max}}\right)\\
\frac{d\mathcal{N}_{\ell_{max}}}{d\ln a} &= \frac{1}{aH}\left(k\mathcal{N}_{\ell_{max}-1} - \frac{\ell_{max}+1}{\eta}\mathcal{N}_{\ell_{max}}\right)
\end{aligned}$$

The equations for the extended models are easily obtained with the same change of variables from the equations in η, reported in the corresponding sections of the chapter.

2.A.3 Adiabatic initial conditions

The default setting of `PyCosmo` is to use the adiabatic initial conditions from `CLASS` [17, 168, 169]. We present the initial conditions for ΛCDM, wCDM and massive neutrinos in the following. We start by introducing auxiliary notation which is useful to define the initial conditions:

$$\eta_0 = \eta(a_0) \qquad \frac{\dot{a}}{a} = aH(a_0) \qquad \omega = \frac{\Omega_m H_0}{\sqrt{\Omega_{r,tot}}} \qquad \frac{\rho_m}{\rho_r} = \frac{\Omega_m a_0}{\Omega_r}$$

$$F_\nu = \frac{\Omega_{\nu,tot}}{\Omega_{r,tot}} \qquad F_{cdm} = \frac{\Omega_{cdm}}{\Omega_m} \qquad F_b = 1 - F_{cdm} \qquad F_g = 1 - F_\nu$$

where $\Omega_{\nu,tot} = \frac{7}{8}N_{\text{eff}}\left(\frac{4}{11}\right)^{\frac{4}{3}}\Omega_\gamma$ which includes massless and massive neutrinos (assumed to be relativistic at early times), $\Omega_{r,tot} = \Omega_\gamma + \Omega_{\nu,tot}$ and η_0 is defined as the minimum between $0.001/k$ and $0.1h$ with k expressed in hMpc^{-1}. a_0 is computed from η_0 assuming radiation domination as $a_0 = \eta_0\sqrt{\Omega_{r,tot}}/H_0$. We use these conservative definitions of initial times to avoid using an iterative shooting algorithm. The initial perturbations are computed in synchronous gauge as

$$\begin{aligned}
\delta_g &= -\frac{(k\eta_0)^2(1-\frac{\omega\eta_0}{5})}{3} \\
\theta_g &= -\frac{k(k\eta_0)^3}{36}\left(1 - \frac{3(1+5F_b-F_\nu)}{20(1-F_\nu)}\omega\eta_0\right) \\
\delta_b &= \frac{3\delta_g}{4} = \delta_c \\
\theta_b &= \theta_g \\
\delta_{ur} &= \delta_g \\
\theta_{ur} &= -\frac{k(k\eta_0)^3}{36(4F_\nu+15)}\left(4F_\nu + 23 - \frac{3(8F_\nu^2+50F_\nu+275)}{20(2F_\nu+15)}\eta_0\omega\right)
\end{aligned}$$

$$\begin{aligned}
\sigma_{ur} &= \frac{2(k\eta_0)^2}{45+12F_\nu}\left(1+\frac{(4F_\nu-5)}{4(2F_\nu+15)}\eta_0\omega\right)\\
l3_{ur} &= \frac{2(k\eta_0)^3}{7(12F_\nu+45)}\\
\eta_{\rm sync} &= 1-\frac{(k\eta_0)^2}{12(15+4F_\nu)}(5+4F_\nu-\frac{(16F_\nu^2+280F_\nu+325)}{10(2F_\nu+15)}\eta_0\omega).
\end{aligned}$$

In the wCDM case we add the generalized initial adiabatic conditions from [148] for the dark energy fields:

$$\begin{aligned}
\delta_{\rm de} &= -\frac{1}{4}(1+w_{\rm de})\frac{4-3c_{\rm s,de}^2}{4-6w_{\rm de}+3c_{\rm s,de}^2}(k\eta_0)^2\\
\theta_{\rm de} &= -\frac{kc_{\rm s,de}^2}{4(4-6w_{\rm de}+3c_{\rm s,de}^2)}(k\eta_0)^3.
\end{aligned}$$

We then need to introduce the transformation from synchronous to conformal Newtonian gauge that uses the following quantities

$$\begin{aligned}
\delta_{tot} &= \frac{F_g\delta_g+F_\nu\delta_{ur}+\frac{\rho_m}{\rho_r}(F_b\delta_b+F_{cdm}\delta_c)}{1+\frac{\rho_m}{\rho_r}}\\
v_{tot} &= \frac{\frac{4}{3}(F_g\theta_g+F_\nu\theta_{ur})+\frac{\rho_m}{\rho_r}F_b\theta_b}{(1+\frac{\rho_m}{\rho_r})}\\
\alpha &= \frac{\eta_{\rm sync}+\frac{3}{2}\left(\frac{\dot a}{a}\right)^2\frac{1}{k^2}(\delta_{tot}+3\frac{\dot a}{a}\frac{v_{tot}}{k^2})}{\frac{\dot a}{a}}
\end{aligned}$$

and reads

$$\begin{aligned}
-\Phi &= \eta_{\rm sync}-\frac{\dot a}{a}\alpha\\
\delta_c &= \delta_c({\rm sync})-3\frac{\dot a}{a}\alpha & \theta_c &= k^2\alpha\\
\delta_b &= \delta_b({\rm sync})-3\frac{\dot a}{a}\alpha & \theta_b &= \theta_b({\rm sync})+k^2\alpha\\
\delta_g &= \delta_g({\rm sync})-4\frac{\dot a}{a}\alpha & \theta_g &= \theta_g({\rm sync})+k^2\alpha
\end{aligned}$$

$$\delta_{ur} = \delta_{ur}(\text{sync}) - 4\frac{\dot{a}}{a} \qquad\qquad \theta_{ur} = \theta_{ur}(\text{sync}) + k^2\alpha$$
$$\delta_{\rm de} = \delta_{\rm de}(\text{sync}) - 3(1+w_{\rm de})\frac{\dot{a}}{a}\alpha \qquad\qquad \theta_{\rm de} = \theta_{\rm de}(\text{sync}) + k^2\alpha.$$

The fields where we did not specify a conversion are gauge invariant, including the massive neutrinos' perturbations of the distribution function. The initial conditions imposed to the multipoles of $\mathcal{M}(\mu)$ are

$$\begin{aligned}
\mathcal{M}_0 &= -\frac{1}{4}\delta_\nu \frac{d\ln f_0}{d\ln q} = \frac{q}{k_b T_0}\frac{\mathcal{N}_0}{e^{-q/k_b T_0}+1}\\
\mathcal{M}_1 &= -\frac{\epsilon}{3qk}\theta_\nu \frac{d\ln f_0}{d\ln q} = \frac{\epsilon}{k_b T_0}\frac{\mathcal{N}_1}{e^{-q/k_b T_0}+1}\\
\mathcal{M}_2 &= -\frac{1}{2}\sigma_\nu \frac{d\ln f_0}{d\ln q} = \frac{q}{k_b T_0}\frac{\mathcal{N}_2}{e^{-q/k_b T_0}+1}\\
\mathcal{M}_3 &= -\frac{1}{4} l3_{ur} \frac{d\ln f_0}{d\ln q} = \frac{q}{k_b T_0}\frac{\mathcal{N}_3}{e^{-q/k_b T_0}+1}
\end{aligned}$$

and are set after transforming the other fields to the conformal Newtonian gauge. We can then relate the fields from CLASS to those of PyCosmo with the following conversions (already introduced in Appendix 2.A.1):

$$u = \theta_c/k \qquad u_b = \theta_b/k \qquad u_{\rm de} = \theta_{\rm de}/k \qquad \Theta_0 = \delta_g/4 \qquad \mathcal{N}_0 = \delta_{ur}/4$$
$$\Theta_1 = \theta_g/3k \qquad \mathcal{N}_1 = \theta_{ur}/3k \qquad \mathcal{N}_2 = \sigma_{ur}/2 \qquad \mathcal{N}_3 = l3_{ur}/4.$$

2.A.4 Fixed cosmological parameters

We report all the parameters of the configuration file of PyCosmo that are kept fixed throughout the chapter:

```
[cosmology]
h = 0.7
```

```
omega_b = 0.06
flat_universe = True
```

```
Tcmb = 2.725
Yp = 0.24
wa = 0.0
cs_de2 = 1.0
T_mnu = 0.71611

[recombination]
recomb = ‘class’

[linear_perturbations]
pk_type = ‘boltz’
pk_norm_type = ‘A_s’
pk_norm = 2.1e-9
k_pivot = 0.05

[internal:boltzmann_solver]
initial_conditions = ‘class’
dt_0 = 1.5e-2
sec_factor = 10.0
boltzmann_max_bdf_order = 5
boltzmann_max_iter = 10000000
fast_solver = True

[internal:physical_constants]
kb = 8.617342790900664e-05
evc2 = 1.7826617580683397e-36
G = 6.67428e-11
hbar = 6.582118991312934e-16
mpc = 3.085677581282e22
mp = 938.272013425824
msun = 1.98855e30
sigmat = 6.6524616e-29
```

Note that the physical constants are not set to the default values in `PyCosmo`, but to default values from `CLASS` (found in the header files `thermodynamics.h` and `background.h`).

The same parameters are passed to `CLASS` and are also fixed throughout this work:

```
output = ‘mPk’
T_cmb = 2.725
Omega_b = 0.06
h = 0.7
T_ncdm =  0.71611,0.71611,0.71611
ksi_ncdm = 0, 0, 0
Omega_fld = 0
wa_fld = 0
cs2_fld = 1
reio_parametrization = ‘reio_none’
YHe = 0.24
gauge = ‘newtonian’
A_s = 2.1e-09
n_s = 1
alpha_s = 0
k_pivot = 0.05
```

Note that some parameters (for instance $T_{\nu,m}$ or $c^2_{s,\mathrm{de}}$) are specified only when necessary.

2.A.5 Power spectrum computation

We compute the gauge invariant real space matter power spectrum, using a fully general relativistic treatment [160–163], accounting for real space matter fluctuations and volume distortions, similarly to `CLASS` [138]. Note that general relativistic corrections are not included in other observables in `PyCosmo`, but are left as future development. Prior versions of `PyCosmo` separated transfer function and growth factor and used the Poisson equation to relate the Newtonian gauge matter density perturbation to the Newtonian gravitational potential. This is still the case when setting `pk_norm_type` to `deltah` and using the power spectrum fitting functions.

We define $\Omega_{m,tot} = \Omega_{dm} + \Omega_b + \Omega_{\nu,m}$ as the total matter energy density, including massive neutrinos and $P_{m,tot} = P_{\nu,m}$, since the massive neutrino component is the only matter ingredient that has a non-zero pressure term. Then the gauge invariant matter density reads

$$\begin{aligned}\delta_{m,tot} = \frac{\delta\rho_{m,tot}}{\bar{\rho}_{m,tot}} + 3\frac{aH}{k^2}\theta_{m,tot} = &\frac{\Omega_{dm}\delta + \Omega_b\delta_b + \Omega_{\nu,m}\delta_{\nu,m}}{\Omega_{m,tot}} \\ &+ 3\frac{aH}{k}\frac{\Omega_{dm}u + \Omega_b u_b + (\Omega_{\nu,m} + P_{\nu,m})u_{\nu,m}}{\Omega_{m,tot} + P_{m,tot}},\end{aligned} \tag{22}$$

where we omitted the a dependencies for brevity.

The power spectrum is defined in terms of the primordial power spectrum of gauge invariant curvature perturbations, $P_{\mathcal{R}}(k) = \frac{2\pi^2}{k^3}\mathcal{A}_s\left(\frac{k}{k_p}\right)^{n_s-1}$ with n_s the tilt of the primordial power spectrum and k_p the pivot scale with corresponding $\mathcal{A}_s$ amplitude , as

$$P(k,a) = 2\pi^2\mathcal{A}_s\frac{k^{n_s-4}}{k_p^{n_s-1}}\delta_{m,tot}(k,a)^2, \tag{23}$$

valid in the case of adiabatic initial conditions for which the initial curvature $\mathcal{R}$ is normalized to 1. These are the only initial conditions currently implemented in `PyCosmo`. `PyCosmo` also allows to output $P_{cb}(k,a)$, the power spectrum of dark and baryonic matter, where massive neutrinos are excluded.

2.B Contributions

For Chapter 2, I was the main contributor. I translated the equations to the correct formalism and implemented the massive neutrinos and radiation streaming approximation equations in the `PyCosmo` framework. I also contributed to the documentation of the code for the release on `PyPi`, drawing and interpretation of the results and writing the manuscript. The work on massive neutrinos was initiated as part of my Master's thesis, of which this chapter constitutes a major extension.

Further contributors: Christiane S. Lorenz, Uwe Schmitt, Alexandre Refregier, Janis Fluri, Federica Tarsitano and Lavinia Heisenberg. They contributed in particular to the interpretation of the results and manuscript preparation. Christiane S. Lorenz derived in the appropriate formalism and implemented the equations for the wCDM model and Uwe Schmitt developed the `sympy2c` framework [141] for symbolic manipulation of equations that is used throughout the chapter. Janis Fluri set up the `PyCosmo`Hub. The mentioned contributors are also the co-authors of the published work [31] upon which this chapter is based.

CHAPTER 3

SIMULATION-BASED INFERENCE OF DEEP FIELDS: GALAXY POPULATION MODEL AND REDSHIFT DISTRIBUTIONS

> *Space is big. You just won't believe how vastly, hugely, mind-bogglingly big it is. I mean, you may think it's a long way down the road to the chemist's, but that's just peanuts to space.*
>
> — DOUGLAS ADAMS, The Hitchhiker's Guide to the Galaxy

This chapter appeared in a similar form in Moser et al., 2024 [32].

3.1 Introduction

Cosmological probes allow us to investigate the structure and components of our Universe, by posing constraints on a cosmological model. The standard model of cosmology, known as ΛCDM, comprises of three main components: dark energy, dark matter and baryons. These components can be traced in a large-scale galaxy survey, by measuring the positions and shapes of galaxies and their correlations. In recent years, state-of-the-art experiments such as the Dark Energy Survey[1] (DES; [18]), the Kilo-

---

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

Degree Survey[2] (KiDS; [19]) and the Hyper Suprime-Cam Subaru Strategic Program [3] (HSC; [20]) have reported their constraints resulting from galaxy clustering, cosmic shear and galaxy-galaxy lensing and their combination, known as 3×2 point analysis [170–172]. Precise determination of the redshift distribution $n(z)$ of samples of galaxies is critical for obtaining cosmological constraints from galaxy surveys. Redshift information allows the separation of source and lens sample, and the computation of cosmological observables. Spectroscopy is prohibitively time-consuming as a technique to provide accurate redshifts of all galaxies in a wide survey, and is furthermore subject to selection biases, especially for faint samples. Surveys thus rely on integrated measurements in a limited number of broad-bands in order to determine the redshift distribution of the sample of interest (for review, see [173, 174]). This has proven to be a challenging task, especially since the relationship between redshift and colour in a limited wavelength range is subject to degeneracies [174]. The characterization of photometric redshift (photo-z) distributions is one of the key systematics affecting cosmic shear measurements since errors in the calibration of redshift distributions and their uncertainties can lead to biases in the retrieved cosmological parameters [174–181]. Traditional photo-z approaches include template fitting (for example LePhare [85, 94], BPZ [96], ZEBRA [97] and EAZY [95]) and machine learning methods (for example ANNz [99], ANNz2 [100, 182] and DNF [101]).

In a cosmological survey it is common to employ methods to constrain the overall redshift distribution of the sample of interest rather than the redshifts of single objects, either by an empirical reweighting of a well measured redshift sample [107, 111, 183], by using spatial cross-correlations [184], or by a combination of the two [112, 185]. Another approach, which has been introduced in recent years, is simulation-based inference (SBI) [120, 121, 127]. SBI relies on forward modelling the survey of interest: the redshift distribution of a sample of galaxies is the result of the statistical properties of the observed galaxy population, the observing conditions and limitations of the detector, and the selection applied to define the target sample. Accurate modelling of magnitude, colour and (at second-order) size distributions of galaxies as a function of redshift, taking into account the observational and instrumental effects, enables

---

[2] http://kids.strw.leidenuniv.nl/

[3] https://hsc.mtk.nao.ac.jp/ssp/survey/

robust determination of the redshift distribution of the sample, as well as a straightforward estimation of its uncertainty. Furthermore, the constraints on the galaxy population model provide some insights on the statistical nature of galaxies and allow a robust treatment of complicated selection functions. The methodology used in this chapter has been developed in [120] and further extended and applied in [127] and [186, 187]. We refer to this forward modelling framework as Monte Carlo Control Loops (MCCL) [27]. The method relies on an empirical galaxy population model to describe the intrinsic properties of galaxies and stars. We render the objects with photon shooting methods using an image simulator called Ultra Fast Image Generator (UFIG [188]) in a set of broad-bands described by the filter throughputs of the telescope used. Moreover, we simulate observational and instrumental effects such as sky and detector noise, point spread functions (PSF), reddening and saturation. In this way, we can post-process the simulations and the real images in the same way. We run SEXTRACTOR [189] on both to obtain catalogs of objects and apply the same selection functions to both the simulated and the real catalog, which simplifies the treatment of selection biases. The model parameters are constrained using the observed data via Approximate Bayesian Computation (ABC), using distance measures that ensure that the photometric properties of the objects in simulations statistically agree with real data. The method has been used to simulate and perform a cosmic shear measurement of the Dark Energy Survey Year 1 [127, 190, 191], for redshift calibration on Subaru data [120], and to obtain the luminosity functions of blue and red galaxies at different redshifts with Canada-France-Hawaii Telescope Legacy Survey (CFHTLS [192]) data [186]. Furhermore, it has been applied to simulate the narrow band imaging of Physics of the Accelerating Universe survey (PAUS [193]) [187, 194] and galaxy spectra from the Sloan Digital Sky Survey (SDSS [195]) CMASS sample [196, 197].

In this chapter, we use HSC Deep/UltraDeep (DUD) data [28] and accurate many-band photometric redshifts from COSMOS2020 [29] in order to obtain tight constraints on the model parameters at high redshift. Previous constraints to the model parameters extended to a redshift of $z \sim 1$, whereas in this work we explore the regime of Stage IV surveys. [198] suggests that HSC data is the most powerful for constraining the Schechter parameters of the luminosity function and thus the redshift distribution of galaxies, because of its exquisite depth. As done in previous work, we use an ABC framework, with several important practical improvements. After tuning the model, we use the obtained posterior parameters to simulate HSC DUD data in the COSMOS field and validate the $n(z)$ for different magnitude cuts.

The chapter is structured as follows. Section 3.2 describes the HSC DUD data and the COSMOS2020 catalog used both to tune the model and validate our results. In Section 3.3, we describe the methodology and introduce the changes compared to previous work. Section 3.4 reports the results of our analysis. We conclude the chapter in Section 3.5. We assume a standard ΛCDM cosmology with h=0.7, Ω_m=0.3, and Ω_Λ=0.7 throughout the chapter.

3.2 Data

In this section, we present the data used to constrain our model of the galaxy population and to validate the obtained redshift distributions. We rely on data from the Deep and Ultradeep layers of the third data release (PDR3) of the Hyper Suprime-Cam Subaru Strategic Program (HSC) [28]. In order to provide our model with additional redshift information and validate the $n(z)$, we complement the HSC data with accurate photo-z estimates from the COSMOS2020 panchromatic photometric catalog [29].

3.2.1 Deep/UltraDeep data from HSC PDR3

HSC is a large multi-band imaging survey conducted with the 8.2-metre Subaru telescope. It comprises of three layers: Wide, Deep and UltraDeep. The Wide layer covers 1470 deg^2, considering partially observed areas in five broad-band filters (g,r,i,z,y). The Deep/UltraDeep (DUD) layers cover ~ 36 deg^2 in the five broad-band filters and four additional narrow-band filters. In this work, we use the publicly available coadded broad-band DUD images with local sky subtraction from PDR3[4]. There are four different fields: COSMOS, DEEP2-3, SXDS+XMM-LSS and ELAIS-N1. Each field is separated in tracts which are equi-area rectangular regions on the sky, divided in 9×9

---

[4] https://hsc-release.mtk.nao.ac.jp/doc/index.php/available-data__pdr3/

patches comprising of 12 arcmin per side corresponding to 4200 pixels and overlapping by 100 pixels. We use all the available DUD patches for tuning the model. The dataset consists initially of roughly 1500 patches. We blacklist patches where:

- more than 30% of the image area is flagged as `NO_DATA`,
- more than 50% of the image area is covered by the `BRIGHT_OBJECT` mask,
- the image overlaps for more than 30% of the area with another patch.

`NO_DATA` and `BRIGHT_OBJECT` masks correspond to flags 8 and 9 in the mask layer of the data. The overlap, on the other hand, is computed using the footprint of the images and by masking the pixels on the top or upper edge of each coadd that are also covered by another patch. After blacklisting, we retain a total of 746 patches.

3.2.2 COSMOS2020 catalog

The COSMOS2020 catalog [29] consists of nearly 1 million high quality photometric redshifts derived via template fitting of many broad and narrow band observations ranging from UV to IR wavelengths. There are four different publicly available catalogs, which differ in the method used for extracting photometry (SExtractor, used in Classic, and The Farmer) and for the photometric redshift template fitting code (`LePhare` [85, 94] and EAZY [95]). Since we use SExtractor for the photometric measurement on the HSC data in our pipeline, we also work with the COSMOS2020 Classic catalog. We use both `LePhare` and EAZY photo-zs. We remove areas where the photometry is unreliable or with partial coverage by means of the `FLAG_COMBINED` parameter thus reducing the area to $1.27 deg^2$. We select objects that have `MAG_AUTO` < 99, `LePhare` (`lp_zBEST`) or EAZY (`ez_z_phot`) photo-z between 0 and 8 (removing `Nan` values), `LePhare` object type galaxy (`lp_type=0`) and SExtractor `FLAGS` < 4. The COSMOS2020 catalog is used both for providing redshift information while constraining the model and for validation. The validation sample is explained in the following section, while the reweighting procedure used during the ABC analysis is detailed in Section 3.3.4.

3.2.3 Validation sample

In order to build our validation sample, we compare the COSMOS2020 `COMBINED` footprint with the HSC DUD data in the COSMOS field and find 63 overlapping patches, out of which 56 are almost fully covered. We perform SEXTRACTOR forced photometry on these coadds using the i band for detection and match the obtained catalog with objects in the COSMOS2020 catalog by position and magnitude (using the magnitude `MAG_APER` measured in a $3''$ diameter apeture). The `BRIGHT_OBJECT` masks from HSC PDR3 are very conservative and cause a loss of roughly one third of the COSMOS2020 objects, since COSMOS2020 uses the less conservative HSC PDR2 masks. We compare the simulated final redshift distributions in the COSMOS field to both `LePhare` and EAZY photo-zs. We take into account sample variance in COSMOS as described in Section 3.3.4.

3.3 Method

The backbone of our forward modelling framework is an empirical parametric model of the galaxy population, used to generate distributions of intrinsic properties of galaxies. Once a galaxy catalog is generated given a set of model parameters, we simulate an image of the survey of interest, in our case the HSC DUD fields. In order to obtain a realistic simulation, we include the effects of the instrument and the known observational systematics that impact the photometric measurement. The end-to-end process from a set of model parameters to a realistic telescope image is implemented in the Ultra Fast Image Generator (UFIG [188]). UFIG has been developed as a simulator for MCCL, with speed as one of the primary features. Computational speed is critical for this task, since a large number of simulations is required to tune the parameters of the model. The inference is performed by running an Approximate Bayesian Computation (ABC), where the realism of our simulated images is increased by minimizing a set of distance measures. In the following, we describe the galaxy population model, how we extend UFIG to reproduce realistic HSC DUD images, how we include redshift information from COSMOS2020 reducing the impact of cosmic variance and the details of our ABC scheme. We focus on the novelties introduced in this work.

3.3.1 Galaxy population model

We include in our model two different populations of galaxies: red and blue, often referred to as quiescent and star-forming galaxies. We sample absolute magnitudes M and redshifts z from Schechter luminosity functions

$$\phi(z, M) = \frac{2}{5}\ln 10\phi^*(z)10^{\frac{2}{5}(M^*(z)-M)(\alpha+1)}\exp\left(-10^{\frac{2}{5}(M^*(z)-M)}\right), \tag{3.1}$$

where the parameters $M^*(z)$ and $\phi^*(z)$ are functions of redshift. We then assign a spectral energy distribution (SED) to each galaxy as a linear combination of 5 spectral templates from KCORRECT [195]

$$SED(\lambda) = \sum_{i=0}^{4} c_i T_i(\lambda).$$

The coefficients of the templates are also different for blue and red galaxies and evolve with redshift. We show the five KCORRECT template spectra for reference in Figure 3.1. We assign sizes to galaxies using a log-normal distribution for the half light radius and a Sersic light profile. The ellipticities are sampled from a Beta distribution. Furthermore, we add stars to our simulations using the Besançon model of the Milky Way [199]. The magnitudes from the catalog of pseudo-stars are sampled with replacement, and the positions are assigned randomly within the HEALPIX pixel (`nside=8`, see [120, 127] for more details). The positions of the bright end of the star distribution is taken from the Gaia DR3 catalog [200, 201] and abundance-matched to the Besançon model. For an extensive description of the galaxy population model, see [120, 186].

In the following, we highlight the modifications to the galaxy population model compared to [120, 127, 186]: (i) modification of the luminosity function parametrization, (ii) addition of new parameters in the morphology sector to allow different characteristics for blue and red galaxies, (iii) small changes to the parametrization of ellipticities

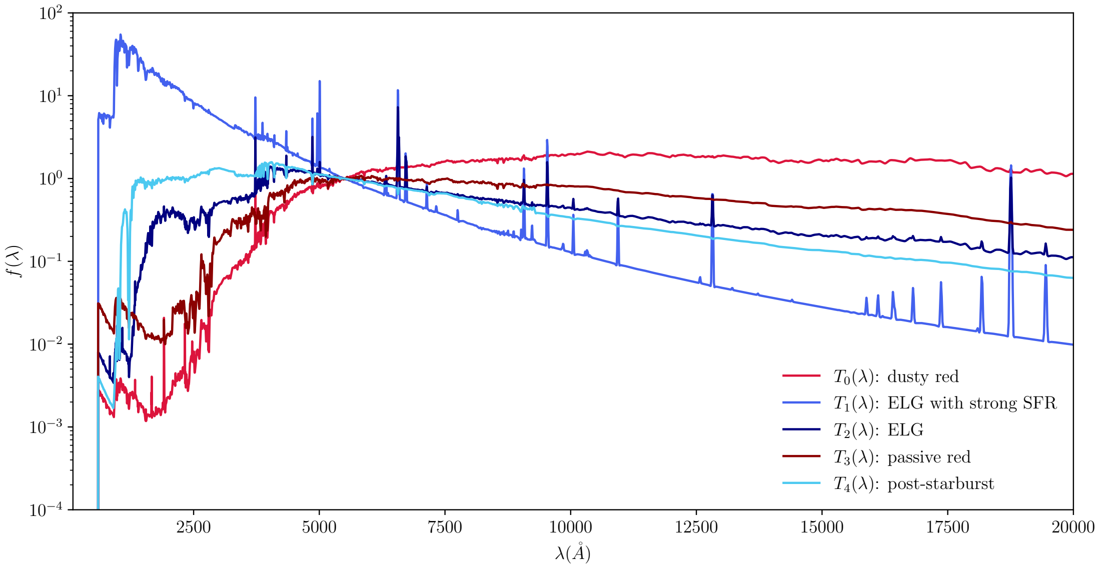


Figure 3.1: The five KCORRECT templates [202] which are combined linearly to obtain the spectral energy distributions of galaxies. They are renormalized so that $f(\lambda) = 1$ at $\lambda = 5500(\AA)$.

and Sersic indices, (iv) modification of parametrization for template coefficients of the SED.

Luminosity function parametrization

We modify the redshift evolution of the luminosity function parameters, in accordance with galaxy evolution models [67]. In the Pure Luminosity Evolution (PLE) scenario, massive galaxies assemble and form most of their stars at high redshifts. They then evolve without merging. This results in the functional form

$$M^*(z) = M^*_{\text{intcpt}} + M^*_{\text{slope}} \log(1+z) \tag{3.2}$$

for the evolution of the characteristic absolute magnitude with different parameters M^*_{intcpt} and M^*_{slope} for blue and red galaxies. In the Pure Density Evolution (PDE) scenario, galaxies undergo mergers so that they are more massive but less numerous at lower redshifts. This scenario can be modelled through the evolution of the normalization of the luminosity function as

$$\phi^*(z) = \phi^*_{\mathrm{ampl}}(1+z)^{\phi^*_{\mathrm{exp}}} \tag{3.3}$$

where ϕ^*_{ampl} and ϕ^*_{exp} also depend on the galaxy population. We vary these 8 parameters during the ABC. Furthermore, α_{blue} and α_{red}, describing the steepness of the faint-end slope of the luminosity function, are also varied in this analysis, differently from previous work.

Updated galaxy morphology

We added new parameters in the morphology section. The relation between galaxy half light radius r_{50} and absolute magnitude is described by three parameters: $\log r_{50}^{\mathrm{intcpt}}$, $\log r_{50}^{\mathrm{slope}}$ and $\log r_{50}^{\mathrm{std}}$ [120]. We sample $\log r_{50}$ from a normal distribution with mean

$$\log r_{50}^{\mathrm{mean}} = \log r_{50}^{\mathrm{slope}} M + \log r_{50}^{\mathrm{intcpt}}$$

and standard deviation $\log r_{50}^{\mathrm{std}}$, where M is the absolute magnitude of the galaxy. In our updated model we have a separate set of these parameters for red and blue galaxies. We also vary the $\log r_{50}^{\mathrm{std}}$ parameter, which was fixed in previous work.

Ellipticity and Sersic indices

The parameterization of ellipticity $p(e)$ using a Beta distribution is slightly modified; parameters e_{mode}, e_{spread} correspond to the mode and concentration of the Beta distribution respectively ($e_{\mathrm{spread}} = a + b$, where a and b are Beta parameters). This change makes the parameters easier to interpret and allows for designing simpler priors. We also modified the prescription for modelling the distribution of Sersic indices. We use a Betaprime distribution with free parameter n_s, which is the mode of the distribution. It is related to parameter α of the Betaprime distribution: $\alpha = n_s(\beta+1)+1$. Parameter β is responsible for the scatter, and fixed throughout the analysis, to the following values: $\beta_{\mathrm{blue}} = 5$ and $\beta_{\mathrm{red}} = 50$. These values were chosen so that the distributions of

Sersic indices match that of [203].

Spectral templates parametrization

In our model, the coefficients of the SED templates are drawn from a five-dimensional Dirichlet distribution separately for blue and red galaxies, similarly to [120, 127, 186]. The Dirichlet distribution is used because the samples drawn from it sum to 1 and the spectrum can then be rescaled to match the absolute magnitude of the galaxy. The parameters of the Dirichlet distribution evolve with redshift: we use two separate sets of parameters for z=0 and z=3, with parameters for other redshifts being an interpolation between them

$$\alpha_i(z) = \left(\alpha_{i,0}\right)^{1-\frac{z}{3}} \times \left(\alpha_{i,3}\right)^{\frac{z}{3}} .$$

In previous work the α_i were constrained at redshifts $z = 0$ and $z = 1$; we now use $z = 3$ since the functional form is fixed (so that we do not need to have a large sample of galaxies at $z = 3$ to pose limits on the parameters' values) and this allows us to enforce prior bounds at higher redshifts.

Previously, the prior on this distribution was also a Dirichlet variable with unity weights, multiplied by a uniform number between [5,15], which accounted for the variance. This way, the α_i parameters were affecting both the mean and variance of the Dirichlet variable. We change the model to capture the mode and variance in separate parameters. Furthermore, [120] derived weights to apply to each template using the New York University Value-Added Galaxy Catalog [202] thus effectively using different template spectra for blue and red galaxies. We removed the weights and reparametrized the template spectra to be normalized to 1 at wavelength 5500 Å. We use a redundant parametrization with modes of the Dirichlet distributions $\bar{\alpha}_i,\ i = 0,\ldots,4$ and two new $\alpha_{\mathrm{std},0/3}$ parameters. The parameters $\alpha_{\mathrm{std},0/3}$ correspond to the standard deviation of the 5-dimensional Dirichlet coefficients with equal concentrations at redshifts $z = 0$ and $z = 3$ and evolve in redshift the same way as the template coefficients α_i. We enforce the normalization $\sum_i \bar{\alpha}_i = 1$, and the final Dirichlet coefficients are calculated as

$$\alpha_i = 1 + \bar{\alpha}_i \cdot \left[\frac{1}{N} \left(1 - \frac{1}{N} \right) \alpha_{\text{std}}^{-2} - N - 1 \right]$$

with N=5. This new parametrization reduces the number of local minima in the problem and makes the $\bar{\alpha}_i$ variables more interpretable. Finally, the template spectra are the same for red and blue galaxies and we encapsulate the information about the different galaxy populations in the ABC prior on the template coefficients. The ABC prior on template coefficients was obtained in a preprocessing step where we performed an ABC on catalog level using the COSMOS2015 catalog [86], as described in Appendix 3.A.1. The final model has 46 parameters, out of which 4 are redundant (see Table 3.A.1 in Appendix 3.A.1).

3.3.2 Image simulations of HSC DUD fields

The catalogs of intrinsic galaxy properties are used to create simulated HSC DUD images. The image generation procedure, including realistic observational and instrumental effects, is as follows. We input the metadata provided by the HSC database[5] about size of the image in pixels, pixel scale ($0.168''$/pixel) and sky coordinates of the images. We perform our simulations using the g, r, i, z and y broad-bands. HSC replaced the r and i filters with more uniform filters $r2$ and $i2$ which have been coadded together with r and i. In our simulations, we use the filter throughputs from $r2$ and $i2$, after checking that the magnitude shifts are small. In order to compute the apparent magnitude of a galaxy in a specific broad-band, we integrate over its SED and the filter throughput taking into account k-corrections and reddening due to galactic extinction. The computation of arbitrary magnitudes in the AB system is described in Section 3.2.3 of [120] and the wavelength dependent extinction to account for reddening in Appendix D of the same paper. The magnitude zeropoint is set to 27 mag/ADU for the HSC coadds.

We simulate PDR3 coadded images directly, as introduced in Section 3.2.1. In order to

---

[5] https://hsc-release.mtk.nao.ac.jp/datasearch/

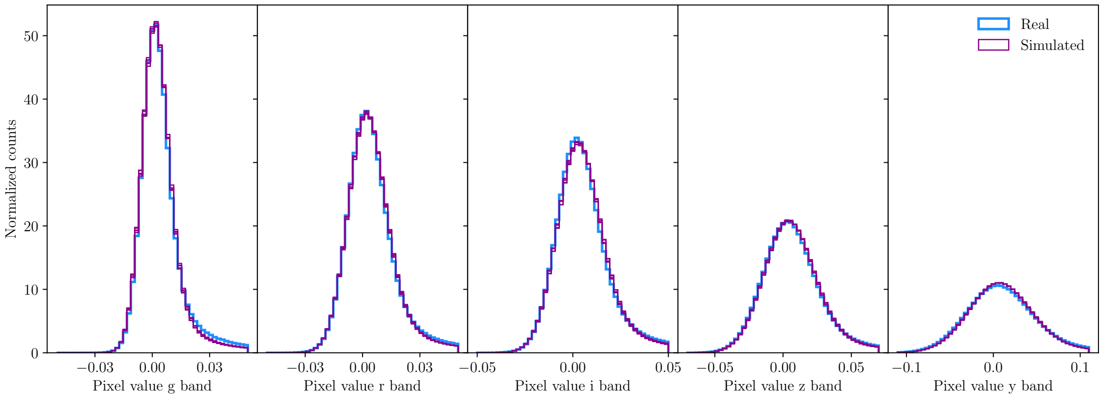


Figure 3.2: Example of pixel histogram around the background values from `tract` 9813 patch 702. The simulated background is shown from 5 different ABC configurations.

simulate the coaddition process, we use systematic maps derived from the metadata. We create a map of the exposure times and number of exposures per pixel for each patch. The CCD gain of a single exposure multiplied by the number of exposures per pixel gives us a rough estimate of the effective gain to convert between ADUs and number of photons. Galaxies are randomly distributed on the image and rendered by sampling individual photons according to the galaxy's Sersic profile. This procedure naturally includes Poisson noise [188, 190]. The Point Spread Function (PSF) is rendered as a distortion to the light profile of the galaxy. In order to estimate the impact of the PSF in the real images, we use a Convolutional Neural Network (CNN) as presented in [204] and updated in [127]. The PSF is estimated at the position of stars matched with Gaia DR3 [200, 201] with magnitudes included between 18 and 22 in the i band, which have SExtractor `FLAGS` 0 or 16 and are not at the image boundary. We perform this selection because stars with apparent magnitudes lower than 18 in the i band are included in the bright objects masks of HSC. The matching is done with a Balltree with a maximum distance of 1.5 pixels. We reserve 15% of the selected stars for validation. Each PSF parameter is then interpolated across the coadd using a Chebyshev polynomial basis of maximum order 4 (see Appendix C of [127] for details).

In order to simulate the background noise in an image, we first derive a map of the root-mean-square of the noise from the real data using the `Background2D` function with `SExtractorBackground` estimator and 3σ sigma-clipping from `photutils` [205]. This map is derived from the data individually for each patch. We then add Gaussian noise to the simulation with mean read off from the image header and standard deviation taken from the map (different for each pixel). Since the standard deviation of the noise that we apply is already different in each pixel and is estimated from background subtracted images, we do not need to perform any background subtraction, including local background subtraction which has the most impact in the surroundings of bright objects, which are anyway masked. The resulting simulated background is in good agreement with the background in the real data. We show an example of the pixel histogram of an image for real data and simulations for low pixel values in Figure 3.2. We observe that there is a slight overestimation of the background level due to the lack of background subtraction.

An alternative approach to background estimation would amount to adding a Gaussian background using the parameters in the image headers (both mean and standard deviation) and then applying global and local background subtraction. In our tests this procedure worsened the agreement between data and simulations. We create masks of the areas with no data and surrounding bright stars using the bit flags 8 and 9 from the second layer of the fits files.

The final steps of the simulation process convert photons to ADUs by dividing out the effective gain and saturate pixels that are above the maximum value of the real data. This is a simplistic estimate of the saturation limit, which is good enough in practice since the saturated areas are always masked. We show an example of a simulated image compared to real data in Figure 3.3. The most noticeable difference between the real image and the simulation is the presence of some large galaxies in the simulation. These are not ruled out by our distance measures and need further investigation. The lack of local background subtraction in the simulated image is noticeable around bright objects.

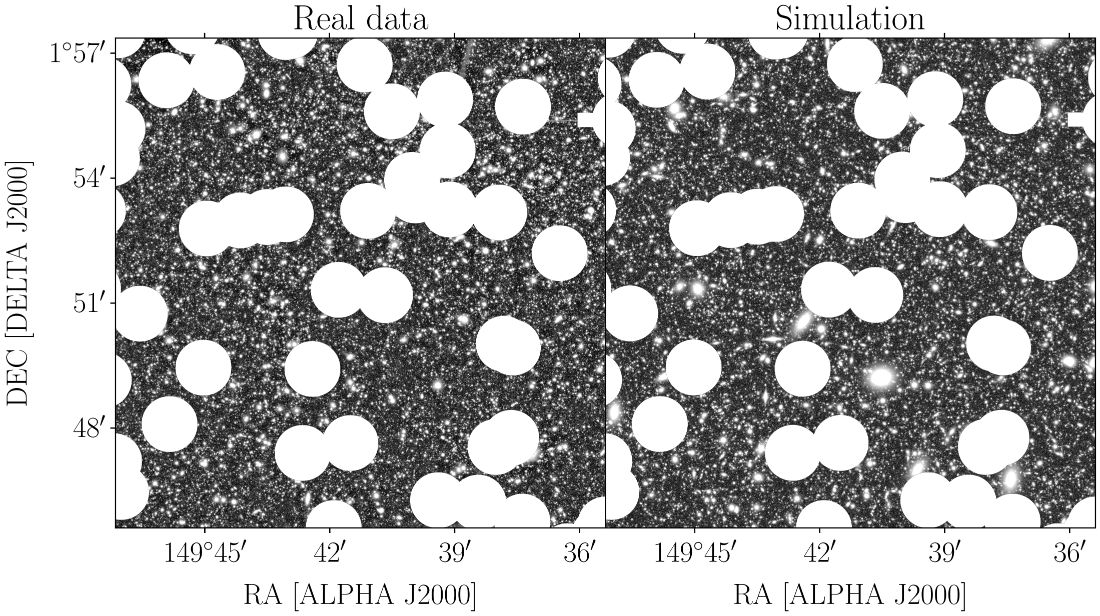


Figure 3.3: A comparison between a real and a simulated coadd for tract 9813, patch 702 in the HSC COSMOS field is shown. The bright star mask, as derived from the data, is applied to both the real data and simulation.

3.3.3 Source extraction and matching

In order to proceed in our analysis, we run SExtractor in dual-image mode with the same settings (reported in appendix 3.A.3) on real images and simulations. We use the i band image as detection image. In the simulations, the SExtractor detections are matched by position and magnitude to the true properties of the injected galaxies. This procedure can have an impact on the resulting photometric properties of galaxies, on the ABC posterior and on the redshift distribution since not all SExtractor detections are matched to an injected galaxy or matched correctly. We use the segmentation map produced by SExtractor and find, for each detection, the overlapping simulated object that minimizes the sum of the differences Δmag between `MAG_AUTO` and true magnitude in all bands

$$\Delta\mathrm{mag} = \sum_{b \in g,r,i,z,y} |\mathrm{mag_b} - \mathtt{MAG\_AUTO\_b}|. \tag{3.4}$$

This matching procedure improves on our previous technique of matching each detected object to the closest detected object below a predefined magnitude difference, especially in the case of a crowded field. Nevertheless, the procedure to match true and detected objects and the impact of blending will need further investigation in the future. While in the matching procedure we use `MAG_AUTO` (which is closest to the UFIG true magnitude), in the following we always use `MAG_APER` in a $3''$ aperture (referred to as `MAG_APER3` from here on, with related quantities `MAGERR_APER3`, `FLUX_APER3` and `FLUXERR_APER3`), unless otherwise specified. This induces photometric biases for bright large objects (which are larger than the fixed $3''$ aperture) but provides more reliable colours and reduces photometric biases for faint objects. `FLUX_AUTO` is the sum of the pixel values assigned to the object and thus depends on the adaptive determination of the object's size. By injecting the same catalog in a deep and ultradeep image, we observed a dependence of `MAG_AUTO` on exposure time, since more pixels of an object rise above the background. A selection in `MAG_AUTO` is undesirable in our case, since we calibrate the model on deep images and then extrapolate it to ultradeep images.

3.3.4 COSMOS2020 redshift assignment to HSC deep fields

Before describing our ABC scheme, we show how we incorporate redshift information from COSMOS2020 in the SEXTRACTOR catalogs obtained from other deep fields. We apply the reweighting technique described in Section 4.2 of [179]. We start from the validation sample introduced in Section 3.2.3 where we have galaxy photometry from our own SEXTRACTOR run in the HSC COSMOS field overlapping with COSMOS2020 and `LePhare` and EAZY photo-zs derived from position matching the COSMOS2020 catalog. In order to assign a redshift estimate to a target galaxy in another deep field, we first add Gaussian noise to the COSMOS galaxies' fluxes until the noise level is equal to that of the target galaxy (the images in the COSMOS field are UltraDeep and thus less noisy than in the other fields). We discard COSMOS galaxies that have larger flux errors than the target galaxy. We then match a COSMOS2020 galaxy to the target galaxy by minimizing the flux χ^2

$$\chi^2 \equiv \sum_b \left(\frac{f_b - f_b^{\mathrm{COSMOS}}}{\sigma_b} \right)^2 \tag{3.5}$$

where $b \in g, r, i, z, y$ and f_b is the `FLUX_APER3` of an object in band b and σ_b its `FLUXERR_APER3`. In this way, we reweight the COSMOS2020 $n(z)$ to match the colour distribution of galaxies in the image we are considering. We verify that a COSMOS2020 galaxy is never matched more than 5 times in the same image (multiple matches only happen for very bright galaxies). In the following subsections, we present two contributions to the uncertainties of the redshift distributions from COSMOS2020, beyond the photo-z errors on individual objects. These are taken into account when validating the $n(z)$ derived from our forward modelling approach against COSMOS2020 photo-zs.

Sample variance in COSMOS

The COSMOS field only spans 2 deg^2 of the sky. This means that, while the volume spanned by COSMOS observations is large due to the considerable depth [206], there are notable sample variance effects at low redshifts. In order to estimate the impact of sample variance, we assign COSMOS2020 photo-zs to all galaxies in the other deep fields and look at the offset in mean redshift. This ensures that we span a larger area and the difference in depth is negligible when only considering galaxies with `MAG_APER3` below 25 in the i band. In order to also measure the scatter due to sample variance, we produce 10 subfields the size of COSMOS (56 images) and measure the standard deviation of the 10 mean redshifts. We prefer this approach to a standard Bootstrap in order to preserve the locality of the effect, which is due to the inhomogeneity of large-scale structure on small scales. The redshift offsets are reported in Table 3.3.1. We observe a mean redshift offset between the COSMOS field and the other deep fields of $\Delta z = \langle z_{\mathrm{COSMOS}} \rangle - \langle z_{\mathrm{deep}} \rangle \approx 0.015$ when we cut at i band magnitude of 23, meaning that sample variance causes a bias for the brightest sample.

Value	i band mag cut 23	i band mag cut 24	i band mag cut 25
Δz SV `LePhare`	0.014 ± 0.012	0.001 ± 0.019	0.002 ± 0.019
Δz SV EAZY	0.015 ± 0.012	0.001 ± 0.019	0.003 ± 0.018
$\Delta z = \langle z_{\mathrm{EAZY}} \rangle - \langle z_{\mathtt{LePhare}} \rangle$	0.014	0.018	0.0006

Table 3.3.1: Shifts in photometric redshift due to sample variance (SV) and difference between the two photo-z codes used in COSMOS2020.

Offset between EAZY and `LePhare`

We notice that, when applying a simple magnitude cut in the i band and the selection described in Section 3.2.2, there is a systematic offset between EAZY and `LePhare` photometric redshifts from COSMOS2020. Similarly to sample variance, this effect has a stronger impact on the brightest sample where EAZY predicts systematically higher redshifts than `LePhare`. The systematic offset is $\Delta z = \langle z_{\mathrm{EAZY}} \rangle - \langle z_{\mathtt{LePhare}} \rangle = 0.014$ for a magnitude cut at `MAG_APER3_i`=23 and 0.018 for a magnitude cut at `MAG_APER3_i`=24. The offset is negligible when cutting at `MAG_APER3_i`=25. We report these offsets in Table 3.3.1.

3.3.5 Factorised ABC inference

We constrain the 46 parameters of our galaxy population model using the HSC deep data and the COSMOS2020 catalog described in Section 3.2. This data combination constitutes a unique sample to precisely constrain our galaxy population model at high redshift, given its completeness up to high magnitudes. We perform simulation-based inference (SBI) to derive a posterior distribution of the parameters of the model, since the likelihood of the observables is unknown, but we have the ability to sample from it through simulations. Our ABC scheme is similar to the one used in [186], and involves prior-to-posterior iterations. The base idea behind ABC is that the model posterior $p(\theta|x)$, where x is the observed data and θ the parameters of the model, can be approximated by $p(\theta|\rho(x, y) < \epsilon)$, where ρ is a distance metric, y is the simulated data and ϵ is a threshold.

The unique property of our problem is that the dataset comprises of a large number of images, which can be considered as semi-independent. We divide our full dataset $\mathbf{d}$ into N smaller parts $\mathbf{d} = d_1, \dots, d_N$. This way we can factorize the posterior on the full dataset into posterior from its parts:

$$p(\theta|\mathbf{d}) \sim \prod_{i=1}^{N} p(\theta|d_i) p(\theta), \tag{3.6}$$

where $p(\theta)$ is the prior on the model parameters θ. Then we use the posterior of one part of the data as prior for another part:

$$p(\theta|d_1) \sim p(d_1|\theta) p(\theta), \tag{3.7}$$

$$p(\theta|d_{i+1}) \sim p(d_{i+1}|\theta) p(\theta|d_i). \tag{3.8}$$

This factorization allows for efficient application of the simple rejection ABC algorithm, which allows for very low complexity of our high performance computing implementation. In practice, we begin by sampling 10000 points from the model's prior $p(\theta)$ and using each model parameter configuration to simulate a part of the data d_1. We accept the parameters θ where the combined distance metric computed from the simulations falls in the 20th percentile. We then resample the obtained distribution and iterate the procedure.

We modify the ABC inference engine compared to [120, 127, 186] as follows: (i) updated distance metrics, (ii) modifications to the ABC iteration engine and posterior modelling, (iii) modification of model's priors.

Distance metrics

We use the Maximum Mean Discrepancy (MMD), a kernel two-sample test for high dimensional probability distributions [207], as our primary distance measure

$$d_{\mathrm{MMD}} = \frac{1}{N(N-1)} \sum_{i,j} k(x_i, x_j) + k(y_i, y_j) - k(x_i, y_j) - k(y_i, x_j) \tag{3.9}$$

where x_i is a property of the ith object in the real data and y_j is a property of the jth object in the simulated data. The kernel we use is Gaussian

$$k(x_i, y_j) = \exp\left(-\frac{||x_i - y_j||^2}{2\sigma}\right) \tag{3.10}$$

with free parameter σ. We describe how we choose the value of σ in Appendix 3.A.2. We extend the input vector of the MMD compared to previous work to include magnitudes `MAG_APER3`, sizes `FLUX_RADIUS`, the two photo-z estimates assigned as described in Section 3.3.4, ellipticities calculated from windowed moments `**_WIN_IMAGE`, and a new variable called the *flux fraction* f_b, calculated as:

$$f_b = \frac{\texttt{FLUX\_APER3\_b}}{\sum_j \texttt{FLUX\_APER3\_j}} \tag{3.11}$$

where `FLUX_APER_b` is the SEXTRACTOR flux in band b in a circular $3''$ aperture. Flux fractions capture similar information as colours. This information is technically also present in the magnitudes, but these mostly impact constraints on the luminosity function parameters. We found that the addition of the flux fractions improved our capacity to constrain the $\bar{\alpha}_i$ parameters of the SED template coefficients distribution. Since our main goal is redshift calibration, we decide to mostly focus on colours, magnitudes and redshifts. For this reason, we include `MAG_APER3` and f_b in all 5 bands, but only include the ellipticity and `FLUX_RADIUS` in the reference i band. Our model does not account for colour gradients in the galaxy size, so it is a reasonable approximation to only constrain the size model in the reference band. The MMD input vector is then 14-dimensional. Before calculating the MMD distance, each column is scaled, so that its mean is close to zero and its standard deviation to 1. The scaling is obtained from the real data and used throughout the analysis for both real and simulated data.

Since the fraction of outliers in the COSMOS2020 catalog for galaxies above magnitude 25 is extremely large ($\simeq 25\%$), we only select objects with `MAG_APER3` < 25 in the i band. We also include the same MMD distance with a cut at i band `MAG_APER3` of 23, in order to upweight this sample of galaxies that is predominant in current large-scale structure surveys. Other selection cuts to remove stars and objects with bad measurements are described in Appendix 3.A.2.

As the MMD does not capture the differences in number counts, which has an important impact on the normalization of the luminosity function, we combine it with a fractional difference in number of objects

$$d_{\mathrm{ng}} = \frac{|N_{\mathrm{SIM}} - N_{\mathrm{HSC}}|}{N_{\mathrm{HSC}}} \tag{3.12}$$

where N_{SIM} is the number of objects in the simulation and N_{HSC} in the real data. This distance is also included for both magnitude cuts at `MAG_APER3_i` < 23 and `MAG_APER3_i` < 25.

The distance metrics in an iteration n are computed for all patches that are included in that iteration ($|d_n|$). We aggregate each distance using the median, which is robust to outliers, so that we have four distances ($d_{\mathrm{ng},25}$, $d_{\mathrm{ng},23}$, $d_{\mathrm{MMD},25}$ and $d_{\mathrm{MMD},23}$) for each parameter configuration. We then rescale the distribution of each distance across ABC points so that it has minimum equal to zero and median equal to 1 and finally add the distances with weights:

$$d_{\mathrm{comb}} = 0.1 \cdot d_{\mathrm{ng},25} + 0.1 \cdot d_{\mathrm{ng},23} + 0.6 \cdot d_{\mathrm{MMD},25} + 0.2 \cdot d_{\mathrm{MMD},23}. \tag{3.13}$$

The weights are chosen to rebalance the sample and upweight bright galaxies (below i magnitude of 23), which are the target of Stage III cosmological surveys, and would otherwise only account for 20% of the sample.

Iteration engine and posterior modelling

We use sets of HSC patches randomly selected without replacement and increase their number in each iteration: the first set has $|d_1| = 10$ patches. Using fewer images at the beginning of the ABC allows us to eliminate very unlikely areas of parameter space without wasting computing time. We then add 1 patch at every subsequent iteration.

In each iteration, we simulate the $|d_n|$ patches for 10000 parameter samples and keep the 2000 samples with the lowest combined distance as the posterior. Then, we perform a density estimation for the posterior using a Gaussian Mixture Model (GMM) with 20 components. We draw the new 10000 samples from this GMM to create a resampled posterior, and pass it as a prior to the next iteration. The GMM fitting is performed in a transformed space to make it easier for the GMM to capture non-Gaussian distributions. We check that the GMM fitting accurately resamples the posterior and monitor the evolution of the distance measures at each iteration. When there is no more improvement in any of the distance metrics, we stop iterating the algorithm. This stopping condition is similar to that of [186], where we look at the evolution of each distance separately because the combined distance is rescaled differently at each iteration. We ran 23 iterations of the algorithm. The details of these iterations are summarised in Table 3.A.4 in Appendix 3.A.2. Appendix 3.A.1 introduces the prior that we used for the ABC run, gives an overview of all model parameters and reports the resulting mean and standard deviation of each parameter in the posterior.

3.4 Results

In this section, we present the results obtained from tuning our galaxy population model. We iteratively performed ABC inference on randomly selected batches of images taken from the HSC deep fields and complemented with reweighted COSMOS2020 many-band photometric redshifts as described in sections 3.3.4, 3.3.5 and appendix 3.A.2. We show the resulting posterior distribution of the model parameters in the following section. Then, we use samples from the posterior to run simulations in the COSMOS field and compare the photometric properties obtained by running SExtractor with the same settings on simulations and real data. We choose to use these patches as validation set, since we have redshift information for individual objects from the photo-z codes `LePhare` and EAZY. We conclude with a comparison of the obtained redshift distributions with the COSMOS2020 catalog at different magnitude cuts.

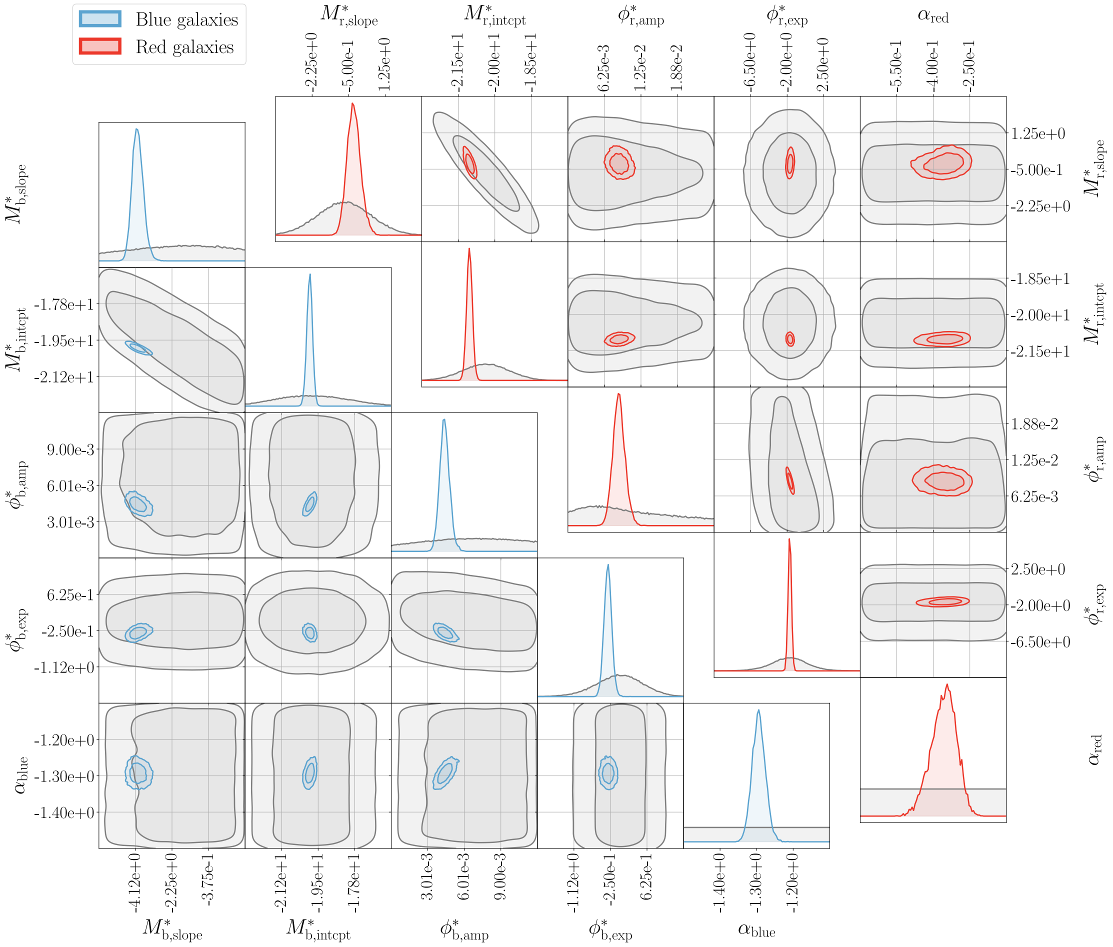


Figure 3.4: ABC posterior for the luminosity function parameters. The parameters for blue and red galaxies are shown in light blue and red, respectively. The ABC prior is shown in grey.

3.4.1 ABC posterior

Figures 3.4, 3.5, and 3.6 show the posterior distribution obtained after 23 iterations of the ABC algorithm, where we fulfill our stopping condition (see Section 3.3.5). The model has 46 parameters, that we show divided into three categories for clarity: luminosity function parameters, SED template coefficients and galaxy morphology parameters.

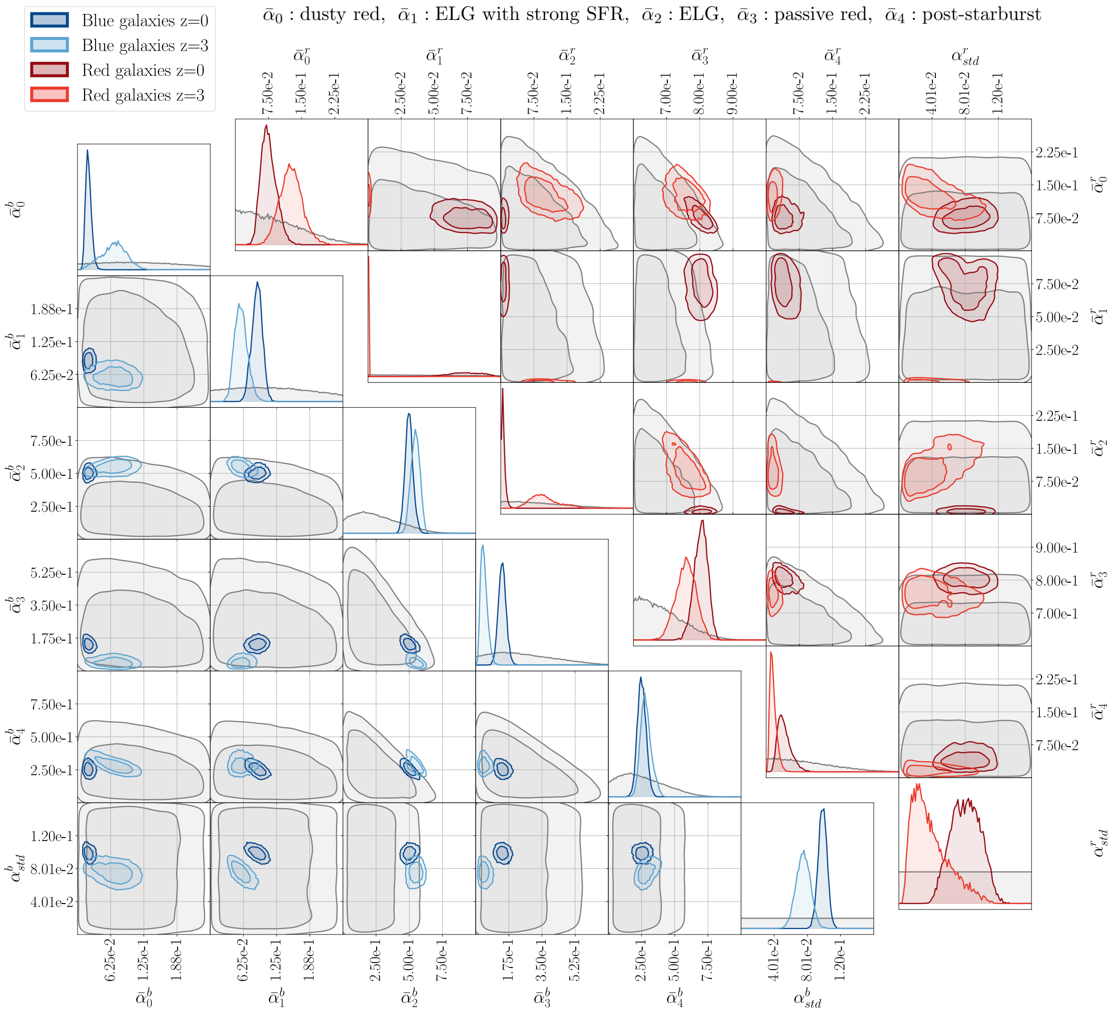


Figure 3.5: ABC posterior for the parameters controlling the spectral energy distributions. The parameters for blue galaxies at redshift 0 are displayed in light blue and in dark blue at redshift 3. Similarly, the parameters controlling red galaxies at redshift 0 are in red and at redshift 3 in dark red. The modes of the Dirichlet coefficients are encoded in the first five parameters and one final parameter controls the standard deviation of the Dirichlet distribution. The ABC prior is shown in grey.

The luminosity function parameters are the most relevant for determining the redshift distribution. In Figure 3.4 we show the parameters for blue galaxies in the lower left triangle and for red galaxies in the upper right triangle. We also plot the prior distribution in grey. We notice that the parameters for blue galaxies are better constrained than those for red galaxies, possibly because blue galaxies are more abundant. The parameters M^*_{intcpt} are generally very well constrained by the ABC, and we notice several strong correlations between the luminosity function parameters, most notably between M^*_{intcpt} and M^*_{slope} and between ϕ^*_{amp} and ϕ^*_{exp}.The parameter α_{blue} we obtain is very close to the fiducial value of -1.3 whereas $\alpha_{\mathrm{red}} \approx -0.35$ is slightly higher than the fiducial -0.5.

In Figure 3.5 we show the constraints on the SED templates coefficients. We show the prior, obtained from the catalog level ABC run described in Appendix 3.A.1, in grey. The posterior distribution of the parameters controlling the blue galaxy population is shown in the lower triangle of Figure 3.5 (in dark blue at redshift z=0 and light blue at redshift z=3) and that controlling red galaxies in the upper triangle (in dark red for redshift z=0 and light red for redshift z=3). We notice that the red galaxies are less constrained, especially at high redshift. The passive galaxy template (T_3) is dominating the SED of red galaxies, as imposed by the prior and there is a significant contribution of the dusty red template (T_0), more prominent at $z = 3$. The contribution of the ELG template (T_2) increases with redshift, whereas that of the ELG with strong star formation template (T_1) decreases with redshift. This contamination indicates that our galaxy population coming from the red luminosity function is not completely passive (probably includes galaxies from the green valley). The post-startburst template (T_4) contributes little and decreases with redshift. The blue galaxies are better constrained and dominated by the ELG template (T_2). The post-starburst template (T_4) is also present in the blue galaxy population at all redshifts, whereas the contribution of the ELG template with strong star formation (T_1) slightly decreases with redshift. The contribution of the dusty red template (T_0) increases with redshift, whereas that of the passive red template (T_3) decreases with redshift. We should not overinterpret the mixture of modes of the template coefficients, since our model allows large freedom due to the scatter parameters and the mixture of templates with redshift evolving coefficients. The KCORRECT templates (shown in Figure 3.1) are derived from SDSS data [208] and do not provide an accurate characterization of galaxies at high redshifts. We also need to consider that there is a smooth transition between red and blue galaxies and that these two categories might not provide a good description of high redshift galaxies.

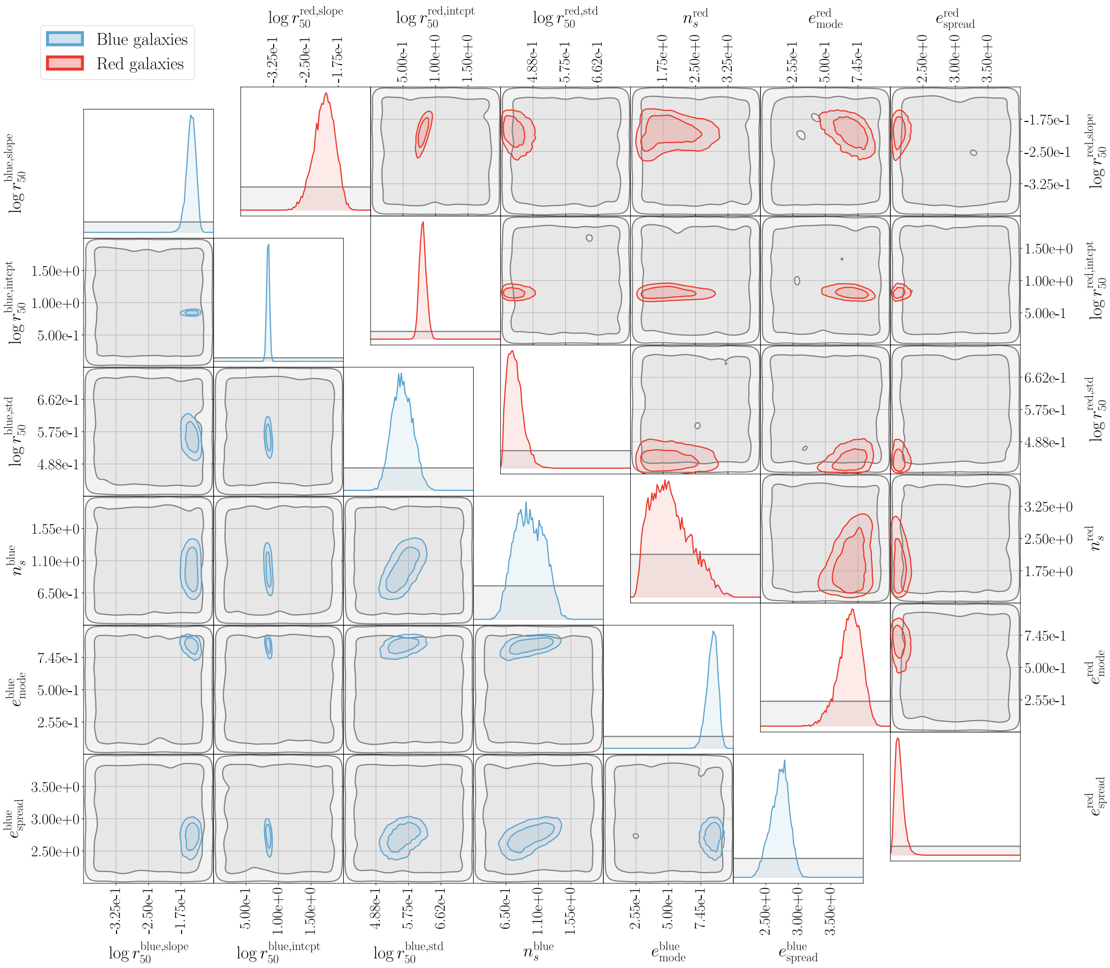


Figure 3.6: ABC posterior for the parameters controlling galaxy morphology. In light blue we show the parameters for blue galaxies and in red for red galaxies. The ABC prior is shown in grey.

In Figure 3.6, we show the parameters describing galaxy morphology. These are considerably more difficult to constrain for our distance metrics, since the effects of the parameters are often degenerate and describe the full intensity profile of the galaxies, which is not sufficiently captured by ellipticity and radius alone. We again show the prior in grey and the posterior distributions for blue and red galaxies in red and blue respectively. We note that the best constrained parameter is $\log r_{50}^{\text{intcpt}}$ for both blue and red galaxies, whereas other parameters are less constrained. The distribution of Sersic indices for red galaxies remains very broad, encompassing values between 1.5 and 3, whereas the distribution for blue Sersic indices is centered at ≈ 1. Sersic indices lack redshift evolution in our model, which is observationally measured (high redshift galaxies have smaller Sersic indices [209]).

3.4.2 Comparison of simulations and real data

We sample 30 parameter configurations from the ABC posterior at random and simulate the 56 HSC images overlapping with the COSMOS2020 COMBINED footprint. We run SExtractor consistently in dual-image mode using the i band image for detection on the 30 simulations and the real data and compare the obtained photometric properties of galaxies. We show 2D contours and 1D histograms of selected photometric properties of simulated galaxy samples and HSC real data up to `MAG_APER3_i` of 25 in Figure 3.8. In the lower triangle we show `MAG_AUTO` in the r, i, z bands and `FLUX_RADIUS` in the i band. The magnitudes show excellent agreement both in the 1D projections and in the 2D contours. The sizes, on the other hand, are more discrepant: there is a tail of large galaxies in the simulations and also a population of galaxies smaller than the smallest galaxies in the data. This is an indication of limitations in the modelling of galaxy morphology. Our model does not include size evolution with redshift at fixed absolute magnitude, which can cause a model bias in the size distribution. The upper triangle of Figure 3.8 shows the colours of galaxies and their correlations. Colours are very important, since they strongly correlate with redshift. We observe a rather good agreement of the colour distributions between the simulations and the real data, with some differences in the tails of the distributions.

In order to highlight the agreement and the discrepancies between colours, sizes and magnitudes, and also display their evolution with redshift, we include a scatter plot in Figure 3.7. We show the relation between i band `MAG_AUTO` and `FLUX_RADIUS`, $r - i$ and $i - z$ colours and redshift (from the `LePhare` photo-z code for the real data). We observe that many trends are present both in the data and simulations: in particular the colour-redshift degeneracies are well reproduced by the simulations up to $z \approx 4$. Since simulating realistic galaxy colours is generally a challenging task, this highlights

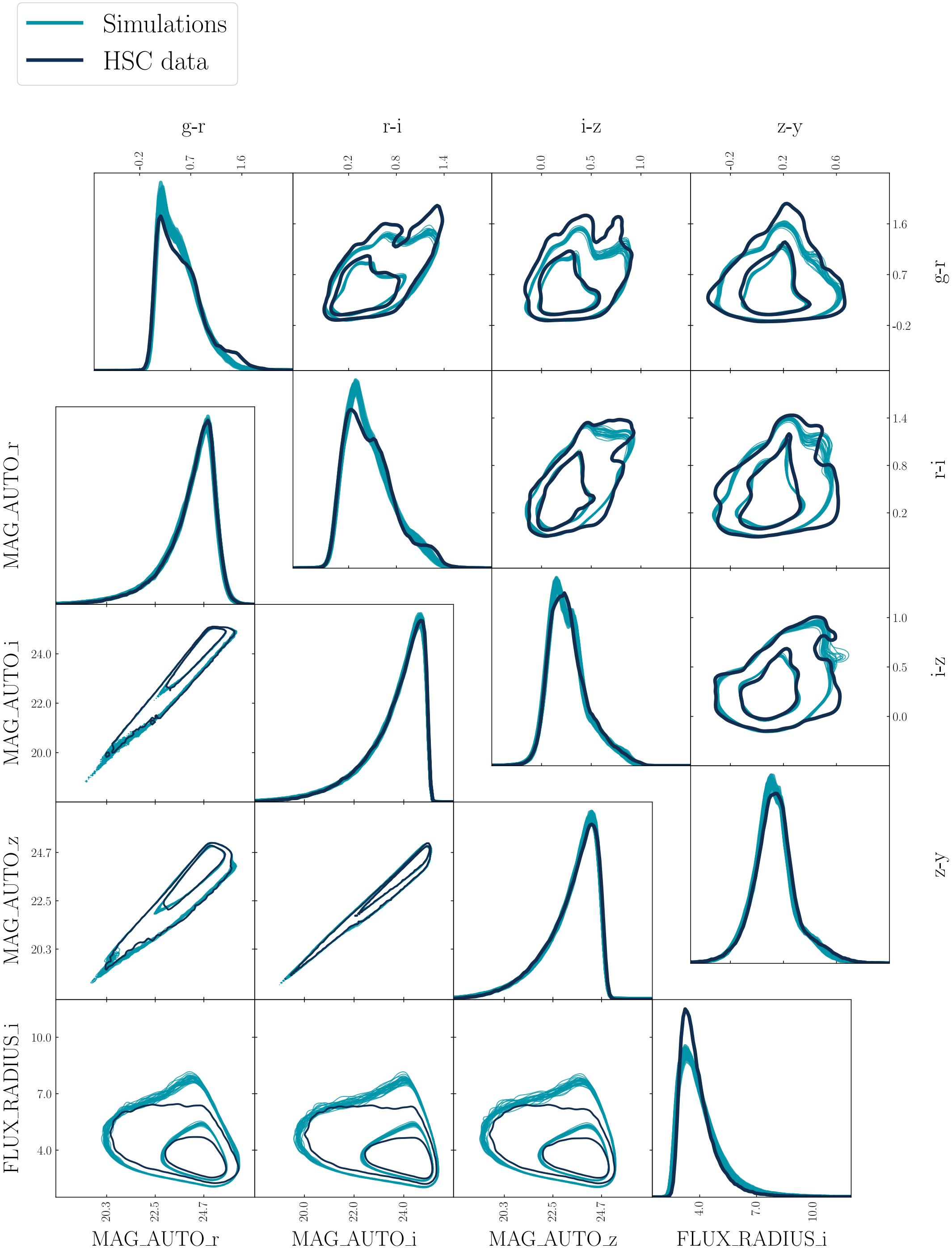
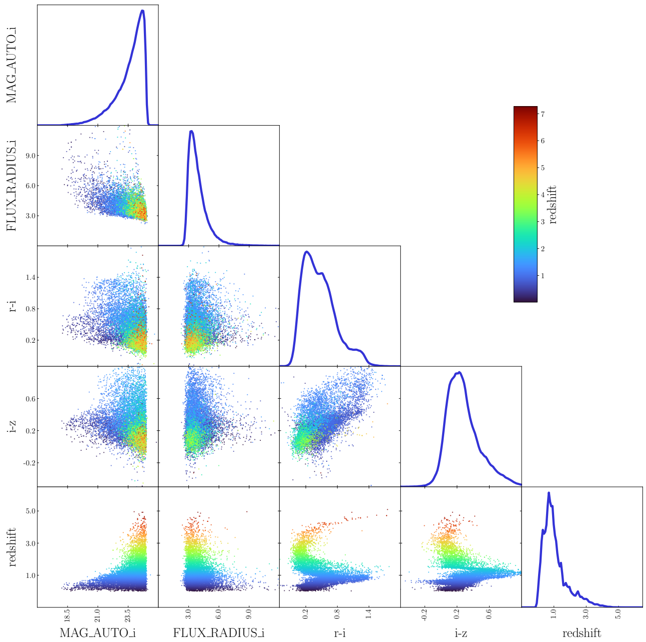
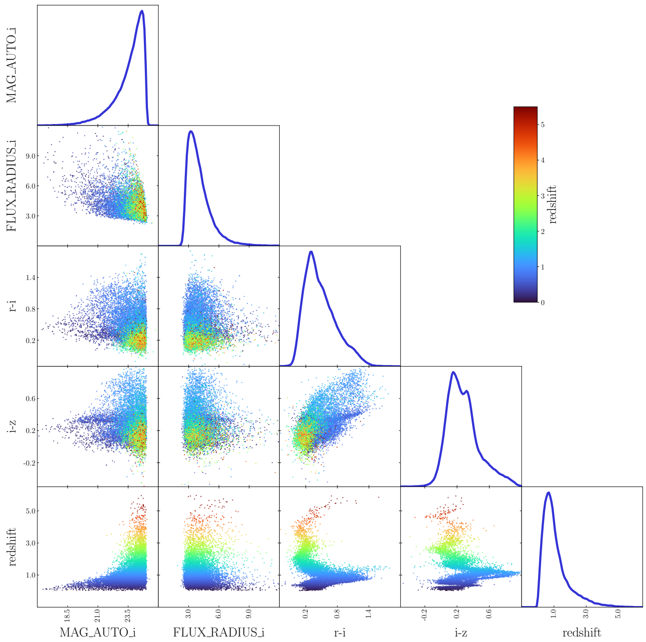


Figure 3.8: Comparison of selected photometric properties (`MAG_AUTO` in r, i and z band, `FLUX_RADIUS` in the i band and colours) of HSC real data (in dark blue) and 30 simulations (in teal) in the COSMOS field.



(a) Scatter plot of selected photometric properties of the data in the COSMOS field. Redshift refers to `LePhare` photo-zs.



(b) Scatter plot of selected photometric properties of one simulation in the COSMOS field.

Figure 3.7: Scatter plots of real data and one of the simulations in the COSMOS field. We include `MAG_AUTO` and `FLUX_RADIUS` in the i band, $r-i$ and $i-z$ colours and redshift (from COSMOS2020 CLASSIC `LePhare` in the case of the real data). Each point corresponds to a randomly selected galaxy from the catalog.

the effectiveness and precision of this method. We note that the simulations include an excess of galaxies at very high redshifts ($z > 4.5$) that are not observed in the data. Due to the relation between absolute magnitude and size (and the evolution of angular diameter distances with redshift), these objects are large enough to be detected. This heavy high-z tail is likely to bias the redshift distribution. The current parametrization of the luminosity function implies that the characteristic absolute magnitude $M^*(z)$ becomes brighter at higher redshifts (with a $\log(1+z)$ redshift dependence). In a Universe that grows hierarchically, this is not the case at high enough redshifts. One possibility would be to truncate the growth of $M^*(z)$ at a redshift z_{max} so that $M^*(z) = \text{const}$ for $z > z_{max}$ but constraining this parameter would be difficult since it would affect only a small fraction of the objects. We also observe an abundance of large objects in the simulations at all redshifts (as already observed in the 1D `FLUX_RADIUS` histogram and in the comparison between real and simulated images). Our model does not account for the evolution of galaxy sizes with redshift at fixed absolute magnitude, which is seen in observations where high redshift galaxies are up to five times smaller than local galaxies (see [209] and references therein). We leave high redshift model refinements for future work.

3.4.3 Redshift distributions

Since the photometric properties of the simulations are in statistical agreement with the data, we derive the posterior redshift distributions. In Figure 3.9 we show the $n(z)$s from 30 simulations in the COSMOS field together with those from the COSMOS2020 photometric redshift catalog in the same field. We also include the redshift distributions obtained by assigning COSMOS2020 photo-zs to 10 sets of 56 continuous patches selected at random in the HSC deep fields, using the reweighting procedure presented in 3.3.4 to reduce the impact of sample variance. We show three different magnitude cuts in the i band `MAG_APER3` in the three columns (`MAG_APER3_i` < 23, 24, 25) and the two different photo-z codes used in COSMOS2020 (`LePhare` and EAZY) in the two rows. We notice that the redshift distributions obtained from MCCL are smooth due to the absence of clustering in the simulations.

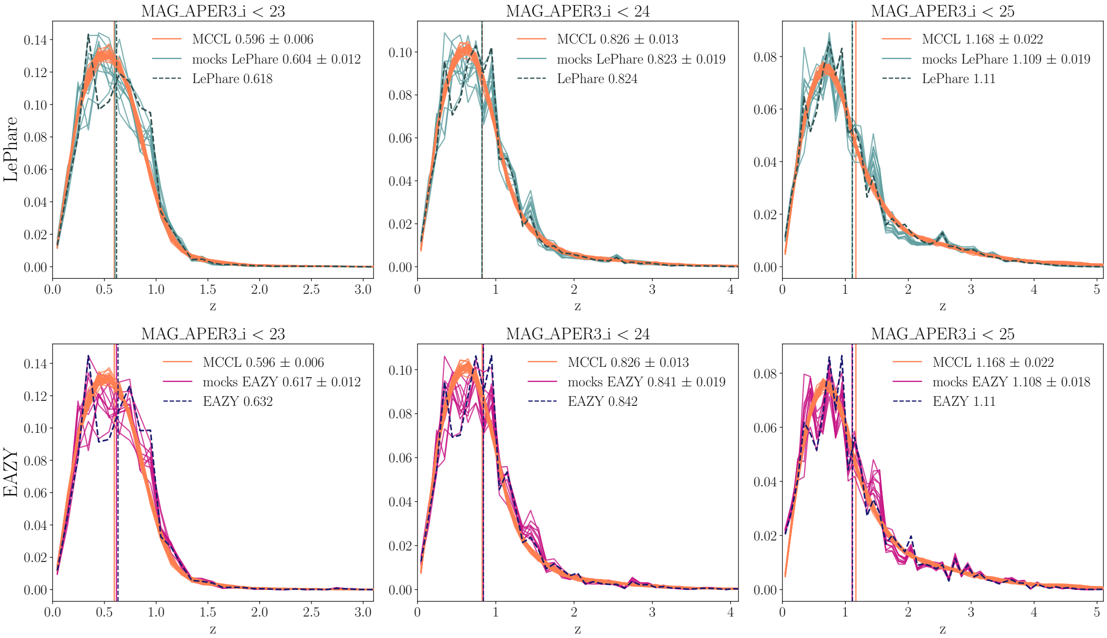


Figure 3.9: Redshift distributions in the COSMOS field, from 30 simulations from the ABC posterior (in orange), from COSMOS2020 photo-zs and from the COSMOS mocks generated by reweighting the COSMOS2020 photo-zs based on HSC deep fields photometry (as explained in Section 3.3.4) in 10 sets of 56 patches. The top row shows `LePhare` photo-zs (in dark grey for COSMOS and sea green for the mocks) and the bottom row EAZY photo-zs (in dark blue for COSMOS and purple for the mocks).

We report mean redshift errors for our simulations, corresponding to the standard deviation of the means of the 30 simulations. We estimate an error on the mean redshift of 0.002 at all magnitude cuts from COSMOS2020 using Bootstrap. Considering sample variance and the systematic offsets between the two photo-z estimates in COSMOS2020 reported in Section 3.3.4 and 3.3.4, we expect a minimum error of 0.02 in the mean redshift from COSMOS2020. By summing this in quadrature with the MCCL errors, we obtain a rough estimation of the combined errors $\sigma_{23} \approx 0.02$, $\sigma_{24} \approx 0.025$ and $\sigma_{24} \approx 0.03$ for the three magnitude cuts at `MAG_APER3_i`$= 23, 24, 25$. We report the mean redshifts and errors in Table 3.4.2. Our estimated redshift distribution for i band magnitude cut of 23 has low mean and is more concentrated around the mean than the $n(z)$ of COSMOS2020, especially when compared to EAZY photo-zs. This can be partly explained by the fact that the simulations are only affected by shot noise due to the limited number of objects, but do not include sample variance, since the

	`MAG_APER3_i` < 23	`MAG_APER3_i` < 24	`MAG_APER3_i` < 25
$\bar{z}_{\text{MCCL}}$	0.596	0.826	1.168
$\bar{z}_{\text{LePhare}}$	0.618	0.824	1.110
$\bar{z}_{\text{LePhare, mocks}}$	0.604	0.823	1.109
$\bar{z}_{\text{EAZY}}$	0.632	0.842	1.110
$\bar{z}_{\text{EAZY, mocks}}$	0.617	0.841	1.108
σ_{MCCL}	0.006	0.013	0.022
σ_{combined}	0.02	0.025	0.03

Table 3.4.2: We report the mean redshifts obtained from MCCL in the COSMOS field, together with the `LePhare` and EAZY mean redshifts from COSMOS2020 at three different magnitude cuts (`MAG_APER3_i` < 23, 24, 25). We also report the `LePhare` and EAZY mean redshifts obtained when reweighting according to HSC deep photometry 10 sets of 56 images (mocks), the MCCL errors and a rough estimation of the combined errors.

objects are randomly distributed in space without accounting for clustering. When considering the $n(z)$ distributions reweighted according to deep field photometry in Figure 3.9, we observe the shift of ≈ 0.015 towards lower redshift reported in Table 3.3.1. To make sure that the systematic offset is not due to the reweighting methodology, we also create mocks from the simulations by assigning the redshifts from a simulated COSMOS field to equal area sets of patches from other simulated HSC deep fields. The obtained redshift shifts are negligible ($\Delta z \approx 0.001$ at all magnitude cuts). The sample variance reduction procedure leads to a 1σ agreement between the redshift distribution of the simulations and the COSMOS2020 data at magnitude cut of 23 in the i band. The redshift distribution for objects below i band magnitude of 24 is in excellent agreement with COSMOS2020. The presence of heavier high redshift tails in the simulations, originating from the extrapolated redshift growth of $M^*(z)$ in the luminosity function parametrization, as explained in Section 3.4.2, reduces the mean redshift agreement for the `MAG_APER3_i` < 25 sample to 2σ. Removing all simulated objects at $z > 4.5$ reduces the MCCL mean redshift estimate of this sample to $\bar{z}_{\text{MCCL,25}} = 1.131 \pm 0.014$, in better agreement with COSMOS2020.

3.5 Conclusions

Redshift calibration is one of the key systematics affecting cosmic shear measurements. Shifts in the mean of the $n(z)$ lead to biased cosmological constraints from large-scale structure surveys [174–181]. It is therefore important to explore and combine a wide range of different methodologies to infer accurate photometric redshift distributions.

In this work, we presented a simulation-based inference approach to obtain redshift distributions from coadded telescope images, extending the result from [120] to deep HSC data and increasing the accuracy of the method. We developed several extensions of the methodology, both in terms of modelling and inference. We calibrated the parameters of our galaxy population model using photometric properties from galaxies in the HSC deep fields and accurate photometric redshifts from COSMOS2020, and obtained realistic simulations. We report the resulting parameters of the model in Table 3.A.1. We compared our results with photometric properties and photo-zs from the COSMOS2020 catalog, simulating the same area with the Subaru telescope in five broad bands, and found good agreement. We showed how sample variance in COSMOS has a strong impact on bright magnitude limited samples. We found a systematic redshift offset in the COSMOS field for objects below magnitude 23 in the HSC i band, common to both photometric redshift methods (`LePhare` and EAZY). Previous work [112, 180] found that the use of the COSMOS field high quality photo-zs alone for redshift calibration could bias low the mean retrieved redshift. This is not in contrast with our results, since redshift calibration strongly depends on the selection function and we are only estimating the sample variance in the COSMOS field itself, and not assessing the impact of using COSMOS2020 as a calibration sample within a methodology. Once this effect is taken into account, our simulations achieve 1σ agreement with the mean of the redshift distribution of COSMOS2020 up to `MAG_APER3_i=24`, and 2σ agreement up to `MAG_APER3_i=25`. The overall shape of the $n(z)$ agrees well. The presence of a high redshift tail at $z > 4.5$ requires further investigation and is an indication of model bias in the luminosity function.

Forward modelling has several advantages that can benefit cosmological large-scale structure surveys in different ways. On one hand, realistic simulations can be used to optimize the survey strategies and model the selection function to the needed level of accuracy. On the other hand, as done in this work, simulations can be used for calibration and to study effects that are difficult to model otherwise, for example the impact of blending and how unrecognized blends can affect the shear measurement. The MCCL method has good prospects of applicability to data from upcoming Stage IV surveys such as the Legacy Survey of Space and Time (LSST) [21] and Euclid [22, 210], which will have depths comparable to the HSC deep fields. The error on the mean redshift per tomographic bin required by these surveys is $\Delta z < 0.001(1+z)$, about an order of magnitude tighter than the current work. In order to make this possible, a number of extensions and improvements are desirable. First of all, imaging of wider deep fields with many band photometry in order to reduce sample variance or deeper spectroscopy with clean selection cuts would greatly benefit photometric redshift calibration in general and forward modelling methods in particular. Secondly, it will be necessary to investigate the evolution of the luminosity function at high redshifts and find a suitable parametrization in order to avoid an excess of high redshift galaxies. This also includes designing good distance metrics to constrain the tail of the distribution using simulation-based inference. It would be beneficial to extend our galaxy population model to include effects that are well understood but not currently modelled, for example the size evolution of galaxies with redshift at fixed absolute magnitude and a relation between the absolute magnitude and colour of red galaxies (more massive galaxies are redder [211]). Another possible improvement of the model of morphologies is the inclusion of bulges and disks instead of a single Sersic profile, as well as the redshift evolution of the Sersic index. A very promising path to a more physically motivated modelling is the use of stellar population synthesis (SPS) models instead of spectral templates to model the galaxies' SEDs. This entails sampling a stellar mass function rather than two luminosity functions and constructing SEDs directly from physical properties of galaxies (such as star formation rates, metallicities and gas properties). This has become feasible in terms of computing time through the emulation of SPS models [121, 212–214]. Emulators can also be used to speed up the UFIG simulations, by mapping the transfer function between catalogs obtained from the galaxy population model and realistic detections. This requires a more detailed understanding of selection effects and the impact of blending. Faster simulations would allow us to test extensions of the model more extensively. Finally, we have discussed how sample variance can impact redshift distributions, when the area considered is limited. In order to obtain realistic sample variance in our simulations, we need to distribute galaxies following the underlying large-scale structure. This can be achieved with the required computational speed by using Subhalo Abundance Matching (SHAM) and approximate simulations as described in [215].

3.A Appendix of chapter 3

3.A.1 Galaxy population model priors

We use a similar model as [27, 120, 186], with a number of modifications, described in Section 3.3. We summarize the model with a description of model parameters, prior distributions and allowed ranges in Table 3.A.1. The model has 46 parameters, but 4 of them are redundant: modes of the template coefficients $\bar{\alpha}_i$ are always forced to sum to $\sum \bar{\alpha}_i = 1$. The luminosity function parameters use the same prior as [186], with standard deviation scaled by a factor of ×3. The prior column in Table 3.A.1 shows the distribution (Normal, Uniform, or Dirichlet) of the prior, as well as the additional bounds applied. For all variables using the Uniform distribution, a joint Sobol sequence was used to generate the prior. A suitable prior for the template coefficients $\bar{\alpha}_i$ is obtained through a catalog level ABC using the COSMOS2015 catalog [86]. We describe this procedure in the next subsection. Table 3.A.1 also lists the mean and standard deviation of each parameter's 1D posterior distribution.

Template coefficient priors from COSMOS2015 catalog

As mentioned in Section 3.3.1, we do not rely on the weights derived in [120] using the New York University Value-Added Galaxy Catalog to differentiate the SED between red and blue galaxies, but impose different priors on the Dirichlet $\bar{\alpha}_i$ parameters for the two galaxy types. This is motivated by the changes to the spectral energy distribution

	Parameter	Meaning	Prior	Posterior
Luminosity functions (10)	$M^*_{\text{b,slope}}$	Slope of the redshift evolution of the parameter M^* in the Schechter LF for blue galaxies, see Equation 3.2	Prior from [186] ×3, $\in [-6, 1.5]$	-4.0 ± 0.3
	$M^*_{\text{b,intcpt}}$	Intercept of the redshift evolution of the parameter M^* in the Schechter LF for blue galaxies, see Equation 3.2	Prior from [186] ×3, $\in [-23, -16]$	-19.9 ± 0.1
	$M^*_{\text{r,slope}}$	Slope of the redshift evolution of the parameter M^* in the Schechter LF for red galaxies, see Equation 3.2	Prior from [186] ×3, $\in [-4, 3]$	-0.3 ± 0.3
	$M^*_{\text{r,intcpt}}$	Intercept of the redshift evolution of the parameter M^* in the Schechter LF for red galaxies, see Equation 3.2	Prior from [186] ×3, $\in [-23, -17]$	-21.0 ± 0.1
	$\phi^*_{\text{b,exp}}$	Exponent of the redshift evolution of the parameter ϕ^* in the Schechter LF for blue galaxies, see Equation 3.3	Prior from [186] ×3, $\in [-2, 1.5]$	-0.31 ± 0.09
	$\phi^*_{\text{b,amp}}$	Amplitude of the redshift evolution of the parameter ϕ^* in the Schechter LF for blue galaxies, see Equation 3.3	Prior from [186] ×3, $\in [1.1{-}5, 1.2{-}2]$	0.0044 ± 0.0004
	$\phi^*_{\text{r,exp}}$	Exponent of the redshift evolution of the parameter ϕ^* in the Schechter LF for red galaxies, see Equation 3.3	Prior from [186] ×3, $\in [-11, 7]$	-1.7 ± 0.2
	$\phi^*_{\text{r,amp}}$	Amplitude of the redshift evolution of the parameter ϕ^* in the Schechter LF for red galaxies, see Equation 3.3	Prior from [186] ×3, $\in [2{-}8, 2.5{-}2]$	0.009 ± 0.001
	α_{blue}	Steepness of the faint-end slope in the Schechter LF for blue galaxies, see Equation 3.1	$\mathcal{U}[-1.5, -1.1]$	-1.29 ± 0.02
	α_{red}	Steepness of the faint-end slope in the Schechter LF for red galaxies, see Equation 3.1	$\mathcal{U}[-0.7, -0.1]$	-0.36 ± 0.05
Galaxy morphology (12)	$\log r50^{\text{blue/red}}_{\text{slope}}$	Slope of the evolution of the average intrinsic physical size of galaxies with absolute magnitude	$\mathcal{U}[-0.4, -0.1]$	b: -0.15 ± 0.01 r: -0.21 ± 0.02
	$\log r50^{\text{blue/red}}_{\text{intcpt}}$	Intercept of the evolution of the average intrinsic physical size of galaxies with absolute magnitude	$\mathcal{U}[0, 2]$	b: 0.84 ± 0.02 r: 0.81 ± 0.05
	$\log r50^{\text{blue/red}}_{\text{std}}$	Standard deviation of the normal distribution we use to sample intrinsic physical galaxy sizes	$\mathcal{U}[0.4, 0.75]$	b: 0.56 ± 0.03 r: 0.44 ± 0.02
	n_s^{blue}	Mode of the Sersic index distribution of blue galaxies	$\mathcal{U}[0.2, 2]$	1.0 ± 0.2
	n_s^{red}	Mode of the Sersic index distribution of red galaxies	$\mathcal{U}[1, 4]$	2.0 ± 0.4
	$e^{\text{blue/red}}_{\text{mode}}$	Ellipticity distribution mode for blue/red galaxies	$\mathcal{U}[0.01, 0.99]$	b: 0.83 ± 0.04 r: 0.69 ± 0.08
	$e^{\text{blue/red}}_{\text{spread}}$	Ellipticity distribution spread for blue/red galaxies	$\mathcal{U}[2, 4]$	b: 2.7 ± 0.1 r: 2.14 ± 0.06
SED coeff. (24)	$\bar{\alpha}^{\text{blue/red}}_{i,0/3}$	Normalized Dirichlet concentration parameters at z=0/3 from which the template coefficients for blue/red galaxies are sampled, i=0,…,4, $\sum_i \bar{\alpha}_i = 1$	5-dimensional Dirichlet, see 3.A.1 for details	See Table 3.A.3
	$\alpha^{\text{blue/red}}_{\text{std},0/3}$	Standard deviation of the normalized Dirichlet concentration parameters at z=0/3 from which the template coefficients for blue/red galaxies are sampled	$\mathcal{U}[1e{-}4, 0.16]$	See Table 3.A.3

Table 3.A.1: Table with galaxy population model parameters, priors and resulting 1D posteriors. Luminosity function is shortened as LF. The details of the prior ranges of the Dirichlet distributions for the template coefficients are explained in subsection 3.A.1.

modelling described in Section 3.3.1, that make the model more interpretable. To capture redshift-colour dependencies of higher redshift galaxies, we derive these priors from a comparison with the COSMOS2015 catalog [86].

First, we select galaxies from this catalog using the following cuts: $z \in [0.3, 4]$, $\mathrm{mag}_{\mathrm{ip}} \in [10, 24.5]$, `TYPE` == 0. The comparison between the simulated and observed galaxies is performed using their redshift and colours, defined with respect to the reference band. We use colour and redshift to avoid constraining the luminosity function, and exploit just the colour-redshift information. We use the reference band ip and compute the colours as a difference with bands NUV, u, B, V, r, zpp, Y, J, H, Ks. For the simulated galaxies, these magnitudes are calculated using UFIG up to catalog generation, with the use of filters provided in the COSMOS2015 dataset. The comparison is performed for red and blue galaxies separately. In the real data we use the `CLASS` provided by COSMOS2015 to separate between star-forming and quiescent galaxies (classified using NUV–r /r – J diagram). We use the nearest-neighbour estimator of Universal Divergence [216] as a distance metric between simulations and real data.

We do not intend to create a posterior on the template coefficient values, but rather find the upper and lower limits on the Dirichlet modes of template coefficient values $\bar{\alpha}_i$. We then perform an iterative procedure of progressively narrowing down the ranges for all coefficients. Starting with a uniform range for $\bar{\alpha}_i \in [0, 1]$, we generate 10000 samples from the luminosity function prior described in Section 3.3.1. We then calculate Universal Divergence between the redshifts and colours from the simulated and COSMOS galaxies. The columns are scaled before the comparison, using a robust scaler from the `scikit-learn` package. We select 2000 best points and calculate the lower and upper limits on $\bar{\alpha}_i$. We then input these new limits and generate another 10000 samples with them. We repeat this process 20 times. The ranges for the coefficients obtained from this procedure define the prior in the main ABC run and are shown in Table 3.A.2, with coefficients rounded roughly to 0.1. The obtained ranges agree with expectations for both blue and red galaxies, with $\bar{\alpha}_2$ and $\bar{\alpha}_3$ dominating for the blue and red galaxies, respectively. We report the means and standard deviations of the obtained posterior of the SED template coefficients in Table 3.A.3.

parameter	$\bar{\alpha}^{\text{blue}}_{0,z}$	$\bar{\alpha}^{\text{blue}}_{1,z}$	$\bar{\alpha}^{\text{blue}}_{2,z}$	$\bar{\alpha}^{\text{blue}}_{3,z}$	$\bar{\alpha}^{\text{blue}}_{4,z}$	$\bar{\alpha}^{\text{red}}_{0,z}$	$\bar{\alpha}^{\text{red}}_{1,z}$	$\bar{\alpha}^{\text{red}}_{2,z}$	$\bar{\alpha}^{\text{red}}_{3,z}$	$\bar{\alpha}^{\text{red}}_{4,z}$
lower limit	0	0	0	0	0	0	0	0	0.6	0
upper limit	0.25	0.25	1	0.7	1	0.3	0.1	0.3	1	0.3

Table 3.A.2: Upper and lower limits on the modes of the Dirichlet coefficients $\bar{\alpha}_i$ for blue and red galaxies derived from the catalog level ABC. The same boundaries are imposed at redshifts $z = 0$ and $z = 3$. These are the upper and lower limits of the Dirichlet priors of the subsequent ABC run on HSC DUD data.

type, z	$\bar{\alpha}_0$	$\bar{\alpha}_1$	$\bar{\alpha}_2$	$\bar{\alpha}_3$	$\bar{\alpha}_4$	α_{std}
blue, $z=0$	0.021 ± 0.005	0.09 ± 0.01	0.50 ± 0.03	0.14 ± 0.02	0.25 ± 0.03	0.099 ± 0.005
blue, $z=3$	0.07 ± 0.02	0.06 ± 0.01	0.55 ± 0.03	0.04 ± 0.02	0.28 ± 0.04	0.075 ± 0.009
red, $z=0$	0.08 ± 0.02	0.07 ± 0.01	0.006 ± 0.004	0.80 ± 0.02	0.04 ± 0.02	0.08 ± 0.02
red, $z=3$	0.13 ± 0.02	$1e-5 \pm 2e-5$	0.10 ± 0.03	0.76 ± 0.03	0.016 ± 0.008	0.04 ± 0.02

Table 3.A.3: ABC posterior means and standard deviations of the SED coefficients for red and blue galaxies at redshift $z = 0$ and $z = 3$.

3.A.2 Details of the ABC runs

In this Appendix, we describe the details of our ABC analysis that were omitted in Section 3.3.5. The ABC iteration engine is similar to the one presented in [186] and depends on a sequence of prior-to-posterior iterations.

Iterations We start by sampling 10000 parameter configurations from the prior defined in Table 3.A.1, with limits in the Dirichlet coefficients $\bar{\alpha}_i$ from Table 3.A.2. We discard samples when the simulation fulfills one of the rejection criteria: (i) having more than 1 million blue or red galaxies (ii) having less than 300 or more than 20000 objects below magnitude 24 in the i band. These are considered extreme conditions, that no simulation that is similar to the real data would fulfil and help us restrict to more likely parts of parameter space, without an excessive use of computing time. For each of the 10000 configurations, we simulate 10 HSC patches in the first iteration, as described in Section 3.3.5. In the following iterations, we increase the number of simulated patches by 1, whereas the number of parameter configurations is fixed to 10000. Table 3.A.4 shows the number of patches simulated per iteration and the corresponding sky area in deg^2.

iteration	$N_{p,\mathrm{sim}}$	sky area
1	10	0.38 deg^2
2	11	0.42 deg^2
3	12	0.46 deg^2
4	13	0.5 deg^2
5	14	0.54 deg^2
6	15	0.58 deg^2
7	16	0.61 deg^2
8	17	0.65 deg^2
9	18	0.69 deg^2
10	19	0.73 deg^2
11	20	0.77 deg^2
12	21	0.8 deg^2
13	22	0.85 deg^2
14	23	0.88 deg^2
15	24	0.92 deg^2
16	25	0.96 deg^2
17	26	1 deg^2
18	27	1.04 deg^2
19	28	1.08 deg^2
20	29	1.11 deg^2
21	30	1.15 deg^2
22	31	1.19 deg^2
23	32	1.23 deg^2

Table 3.A.4: Number of patches used in each iteration and corresponding sky area covered by the HSC DUD patches.

Sample selection The distance measures described in Section 3.3.5 are computed using the SExtractor catalogs, created in all $grizy$ bands, based on the detection in the i band. We perform the PSF estimation using a Convolutional Neural Network [204], in the same way as in [127]. We run SExtractor on the HSC data first, and then on the simulated images during the ABC iterations. From the catalogs, we select galaxies with the following set of cuts applied in all bands with strict `and` conditions:

$$\texttt{FLAGS} < 4, \quad 14 < \texttt{MAG\_APER3} < 30, \quad \texttt{MAG\_AUTO} < 99, \quad 0.1 < \texttt{FLUX\_RADIUS} < 100,$$
$$-3 < \log_{10}(\texttt{SNR}) \equiv \log_{10}\left(\frac{\texttt{FLUX\_AUTO}}{\texttt{FLUXERR\_AUTO}}\right) < 4, \quad 0 < \texttt{ELL} < 1,$$

$$\texttt{N\_EXPOSURES} > 0, \quad 0.5 < r_{50}/\texttt{PSF\_FWHM}, \quad \texttt{CLASS\_STAR} < 0.95$$

where ELL is the absolute ellipticity calculated from windowed moments `**_WIN_IMAGE`, `N_EXPOSURES` is the number of exposures in the coadd at the position of the object, r_{50} is the object size defined as $r_{50} = 2 \cdot \ln(2) \cdot (\texttt{X2\_WIN\_IMAGE} + \texttt{Y2\_WIN\_IMAGE})^{1/2}$, as in [127]. Both the PSF size cut and the `CLASS_STAR` cut are applied to create a pure galaxy sample. In addition, we impose $\texttt{MAG\_APER3} < 25$ in the i band. Note that the SEXTRACTOR detections are matched in the simulations to the true properties of the injected galaxies, as explained in Section 3.3.2, so that a further criterion for simulated objects to be selected for the MMD distances is that the detection has been matched to a true simulated object. We additionally require that the objects do not lie on the image mask. When computing $d_{\mathrm{MMD},23}$ we additionally impose $\texttt{MAG\_APER3\_i} < 23$.

Optimization of the kernel radius parameter To obtain the most sensitive MMD distance, it is common to optimize the parameters of the kernel used to compute it [207]. We use a Gaussian kernel with a single parameter σ, corresponding to the correlation scale. We compute σ for the different MMD distances as the median distance between samples drawn from the same probability distribution (the real data) [120, 207].

Modelling of posterior distributions We create the posterior distribution at each iteration by setting the 20th percentile as a threshold, thus selecting the 2000 out of 10000 samples with the smallest combined distance. We then create a model of this posterior using a Gaussian Mixture Model (GMM), from the `scikit-learn` implementation[6]. We use 20 Gaussians to fit the distribution. Before fitting, we transform the model parameter samples to a gaussianized space. This is done to make the GMM more suited for fitting the distribution, especially for parameters with uniform priors. First we rescale the parameters to lie between $\in [1{-}8, 1 - 1{-}8]$, and then apply a Gaussian inverse-CDF transform. We draw 10000 new samples in the

---

[6] scikit-learn.org/stable/modules/generated/sklearn.mixture.GaussianMixture.html

SExtractor parameter	Value
CATALOG_TYPE	FITS_1.0
DETECT_TYPE	CCD
DETECT_MINAREA	5
THRESH_TYPE	RELATIVE
DETECT_THRESH	1.5
ANALYSIS_THRESH	1.5
FILTER	Y
FILTER_NAME	gauss_3.0_5x5.conv
DEBLEND_NTHRESH	32
DEBLEND_MINCONT	0.00001
CLEAN	Y
CLEAN_PARAM	1.0
MASK_TYPE	CORRECT
MAG_ZEROPOINT	27
PIXEL_SCALE	0.168
STARNNW_NAME	default.nnw
BACK_SIZE	128
BACK_FILTERSIZE	3
BACKPHOTO_TYPE	LOCAL
BACKPHOTO_THICK	24
WEIGHT_TYPE	NONE

Table 3.A.5: SExtractor configuration used in this work both on real images and simulations. The missing parameters change per patch and band and are described in the text.

gaussianized space using Sobol sampling and invert them back to the original space. We verify that the GMM model in the transformed space is a good representation of the posterior by comparing the 2D marginal projections of all parameter combinations for the original 2000 samples and 10000 new GMM samples. The GMM samples from the model posterior are passed as priors to the next iteration of the ABC algorithm.

3.A.3 SExtractor settings

We report in Table 3.A.5 the SExtractor configuration.

3.B Contributions

For Chapter 3, I was the main contributor. I devised the project, extended the codebase, ran the Approximate Bayesian Computation and the final simulations. I created the figures, interpreted the results and wrote the majority of the manuscript.

The early stages of the project were carried out by Dominic Grimm during his Master thesis, which I supervised. He contributed mostly to the design of the forward model of HSC deep images and the creation of a first version of the systematic maps. Tomasz Kacprzak also contributed to the project by designing and running the catalog level ABC on COSMOS2015, reworking parts of the code and participating to writing of the manuscript. Silvan Fischbacher helped with the optimization of the code and discussion of the results. Further contributors are Alexandre Refregier and Luca Tortorelli, who contributed mostly to the interpretation of the results and early code developments. The mentioned contributors are also the co-authors of the submitted work (Moser et al., 2024 [32]) upon which this chapter is based.

CHAPTER 4

REDSHIFT CALIBRATION: A COMPARISON BETWEEN SOMPZ AND MCCL FOR DES Y3

> *All you really need to know for the moment is that the universe is a lot more complicated than you might think, even if you start from a position of thinking it's pretty damn complicated in the first place.*
>
> — DOUGLAS ADAMS, Mostly Harmless

4.1 Introduction

Current and future large-scale cosmological imaging surveys such as the Dark Energy Survey[1] (DES; [18]), the Kilo-Degree Survey[2] (KiDS; [19]), the Hyper Suprime-Cam Subaru Strategic Program[3] (HSC; [20]) , the Vera C. Rubin Observatory Legacy Survey of Space and Time[4] (LSST; [21]) and Euclid[5] [217] are imaging a growing area of the

---

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

[2] http://kids.strw.leidenuniv.nl/

[3] https://hsc.mtk.nao.ac.jp/ssp/survey/

[4] https://www.lsst.org/

[5] https://www.euclid-ec.org/

sky with unprecedented depth. These surveys combine different probes in order to deliver stringent constraints on the cosmological models. Cosmic shear, the weak deformation of galaxies due to the large-scale structure of the Universe intervening between the source and the observer, is currently one of the most promising probes of the content of the Universe and growth of structure, with the potential to shed light on the nature of dark energy [79]. The combination of cosmic shear, galaxy clustering and galaxy-galaxy lensing (known as 3×2-point analysis) is reaching a constraining power close to that of the CMB temperature anisotropies and E mode polarization in the $S_8 - \Omega_m$ plane ([170, 171]).

Weak lensing, despite its simple theoretical framework, is a demanding probe in terms of control of systematics. Among the main sources of systematic uncertainty are the errors in the measurement of the redshift distribution of galaxies on noisy photometry [218]. Ideally, the measurement of spectra of all galaxies in a sample would yield an accurate redshift determination but the large number of galaxies in wide field surveys and their faintness make this infeasible. Wide field surveys thus need to rely on photometric redshift estimates derived from a small number of broad-band filters [173, 219]. Since cosmological parameter estimation can be biased by systematic effects in the redshift distribution (e.g. [175]), it is crucial to develop and compare different redshift calibration methodologies. Many of the methods rely on per object photo-z estimates, obtained either by template-fitting (e.g. [94–96, 220]) or machine learning methods (e.g. [100, 105, 221]), which are then stacked to obtain a distribution. Direct calibration methods estimate the overall redshift distribution of a sample of galaxies by reweighting the distribution of a subsample with good quality spectroscopic or many-band photometric redshifts [107, 179, 222]. One recent approach to reweighting makes use of Self-Organizing Maps (SOMs) [223], an unsupervised dimension reduction algorithm, that discretizes the magnitude-colour space of a galaxy sample in two dimensions and creates an empirical relation between the desired wide field redshift distribution and a sample with well characterized redshifts. This method was proposed for redshift calibration by [109] and is currently used in KiDS [113, 183] and DES [110, 112]. We refer to this method as Self-Organizing Map p(z) (SOMPZ). The methodology can be used in combination with clustering redshifts, the exploitation of spatial cross-correlations between a target sample and a sample with known redshifts [183, 224, 225]. A different approach to the calibration of the overall redshift distribution of a sample of galaxies was proposed by [120] in the context of the Monte Carlo Control Loops (MCCL) framework. This approach derives the redshift distribution from realistic simulations, which are tuned via simulation-based inference to agree statistically with survey images.

In this chapter, we perform a blinded comparison between the tomographic redshift distributions obtained by applying the SOMPZ and MCCL methodologies to the same target sample of galaxies from DES Year 3 (DES Y3) data. This allows us to cross compare the two methodologies and assess the agreement of the retrieved redshift distributions, which leverage different sources of information. A bias in the mean redshift, which contains most of the necessary information for an integrated probe such as weak lensing, could lead to distortions in the retrieved cosmological parameters and is considered one of the possible drivers of the discrepancies found between different surveys [111].

The chapter is organised as follows. In Section 4.2 we present the DES Y3 data on which we perform the analysis and the simulated data employed by the two pipelines. Section 4.3 describes the two methodologies and compares their features. We discuss the implementation details of the two methods in Section 4.4, which includes some intermediate results, and the blinding strategy in Section 4.5. We show the unblinded results of the analysis in Section 4.6, where we gain deeper insight into the two methodologies with detailed comparisons. In Section 4.7 we rerun the MCCL simulations using the ABC posterior from Chapter 3, constrained using HSC deep data and photo-zs from COSMOS2020. We show the global redshift distributions in Section 4.8. We present our conclusions in Section 4.9.

4.2 Data

The two different methodologies are applied to DES Y3 wide field data, with the same selection applied. Furthermore, each methodology uses additional data: SOMPZ relies on the deep and redshift samples, complemented with the `Balrog` artificial data to calibrate the redhsift distribution. On the other hand, MCCL does not use additional calibration data, but relies on complete simulations of the DES Y3 images.

4.2.1 Wide field data

We compare the two redshift calibration methodologies on Dark Energy Survey Year 3 data [226]. The wide field images cover ~ 5000 deg^2 of area, representing a large increase in area compared to the ~ 1500 deg^2 of DES Year 1 [30] with a similar depth. After masking foregrounds and bad regions, the effective area is ~ 4143 deg^2 [225]. In the MCCL methodology, we produce a SExtractor [189] catalog by running, for each coadd in each band, forced photometry on a riz `CHI-MEAN` detection image. This image is created by a weighted linear combination of the r, i, z images following the procedure of [226]. We rerun the detection image creation in our pipeline in order to be able to perform exactly the same procedure for survey images and simulations. Additionally, we use information about the single exposures, to create noise maps and coaddition patterns that are used as systematic maps in the PSF estimation procedure, similarly to [127]. SOMPZ instead is run directly on DES Y3 Gold, the photometric dataset assembled for DES Y3 cosmological analyses [227]. We apply to both catalogs the selection presented in the next subsection. We deploy data taken in the $grizY$ bands for MCCL, even though the g and Y band data are excluded while training the self-organizing maps for SOMPZ. The g band images present difficulties in adequately modelling the point spread function with PIFF [228]. Since we do not rely on METACALIBRATION photometry [229] in this analysis, it is not necessary to exclude this band, but we still exclude it from the SOM in order to match the analysis choices of [112]. This choice allows us to avoid repeating the uncertainty estimation for the SOMPZ methodology. Furthermore, as will be shown in section 4.6, the g band is the most discrepant in the comparison between the simulations used in MCCL and survey data. The Y band is noisier than other bands, as the exposure time for a single exposure in this band is 45 s compared to 90 s for other bands. For this reason, the Y band is only used in the tuning of the simulation parameters for MCCL via Approximate Bayesian Computation (see Sections 4.3.1 and 4.4.3), but only appears as a diagnostic in the rest of the analysis.

Target sample

In order to fairly compare the SOMPZ and MCCL redshift distributions, we apply both methods to the same sample of galaxies, which we call the *target sample*. We design the target sample to resemble the DES Y3 weak lensing source catalog [119], but without relying on METACALIBRATION photometry which is not available in the MCCL pipeline. The redshift distribution of the weak lensing source catalog is calibrated in [112]. We base the simplified selection for the target sample on SEXTRACTOR photometric quantities, which are easily accessible to both pipelines. Note that in the case of SOMPZ, the cuts are applied to the DES Y3 Gold catalog [230] with

```
FLAGS_FOOTPRINT == 1 and FLAGS_FOREGROUND == 0 and
     FLAGS_BADREGIONS < 2 and FLAGS_GOLD < 8,
```

to which we refer as *unmasked Y3 Gold*. In MCCL we first run SEXTRACTOR on both real and simulated images and then mock the unmasked Y3 Gold selection. This implies selecting the survey footprint using the systematic maps of single exposures, removing the tiles near the Large Magellanic Clouds, the objects matched with the Sky2000 catalog of bright stars (below 8.0 mag) [231] and imposing `FLAGS_g,r,i,z` $\leq$ 3. The selection for the target sample is then defined as follows:

```
2.5 < FLUX_RADIUS_i < 7 and 0 < FLUX_RADIUS_g,r,z < 10 and
  17 < MAG_AUTO_i < 23.5 and 17 < MAG_AUTO_g < 99 and
         17 < MAG_AUTO_r,z < 25.5 and
     SNR = FLUX_AUTO / FLUXERR_AUTO > 7 in r, i or z band.
```

The conditions on `MAG_AUTO` and `FLUX_RADIUS` are imposed to mock the magnitude-radius distribution of the weak lensing catalog. The last selection cut in signal-to-noise ratio is designed for removing false detections of faint galaxies. Many of the stars are excluded from the sample due to the cuts in `FLUX_RADIUS_i` and we remove most of the remaining stars using a random forest trained on UFIG simulations, as explained in Section 4.4.1. Target sample will thus refer to the two samples with simplified cuts and star-galaxy separation applied, one for SOMPZ and one for MCCL.

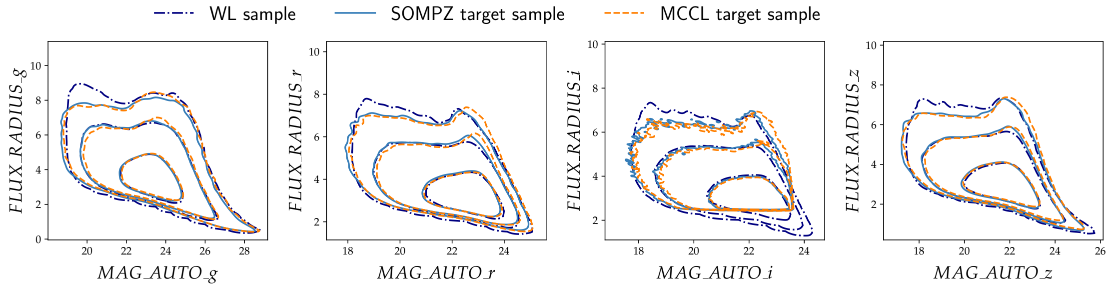


Figure 4.1: We show the 68%, 95% and 99% contours of the `MAG_AUTO` - `FLUX_RADIUS` distribution of galaxies in the g, r, i, z bands with different selections: the WL source galaxies from DES Y3 (dark blue dash-dotted line), the target sample of SOMPZ (light blue solid line) and the target sample of MCCL (dashed orange line).

`Y3 Gold`	WL source catalog	target sample SOMPZ	target sample MCCL
319,296,026	100,208,944	104,491,927	103,822,956

Table 4.2.1: Number of galaxies in the whole Y3 Gold, the weak lensing source catalog and the target samples of this analysis for SOMPZ and MCCL.

The target samples distribution in the g, r, i, z magnitude-radius plane is displayed in Figure 4.1. We observe that the i band, where the more restrictive selection is defined, shows sharp cuts in magnitude and radius, while the magnitude-radius distributions in the other bands resemble the weak lensing source catalog closely. In Table 4.2.1 we report the number of galaxies in the weak lensing source catalog compared to our target samples and the total number of galaxies in the unmasked Y3 Gold catalog.

4.2.2 Deep and redshift samples for SOMPZ

The SOMPZ methodology relies on observations from deep fields, which allow us to break degeneracies between redshift and colours due to the larger number of filter bands and are more reliable in terms of density estimation due to their depth. The deep data is available for the DES deep fields which have measured photometry in the $ugrizJHK_s$ bands. The DES deep fields cover areas of 3.32, 3.29, 1.94, and 1.38 deg^2, for a total overlap of 5.2 deg^2 after masking the bad regions. A detailed description of the data products is provided in [232] and [233]. SOMPZ also requires samples with accurate redshift estimation, such as spectroscopic samples or precise many-band photometric samples. In this analysis we use both public and private spectra from the spectroscopic catalogs zCOSMOS [234], C3R2 [235], VVDS [236], and VIPERS [237]. We additionally use photo-z estimates from two multi-band catalogs: COSMOS2015 30-band catalog [86] that includes optical, UV and mid-infrared data from the 2 deg^2 COSMOS field and PAUS+COSMOS, which is a 66-band catalog that combines 26 COSMOS bands (excluding the infrared) and 40 PAU narrow bands [238]. Our fiducial redshift sample for the SOMPZ analysis in this work is the SPC sample from [112] which ranks the spectroscopic catalogs first, followed by PAUS+COSMOS and COSMOS2015 many band photo-zs. We also use three other redshift samples to check for robustness: PC (photometric only, ranks PAUS+COSMOS first and then COSMOS2015), SC (ranks spectroscopic redshifts first and combines them with COSMOS2015) and C (only COSMOS2015). Due to the different selection in the wide fields, the contribution from the different redshift samples in SPC changes slighty from [112]: we have 40.5% redshifts from spectra, 44% from `PAUS+COSMOS` and 15.5% from COSMOS2015.

4.2.3 Artificial wide data for SOMPZ: `Balrog`

SOMPZ uses `Balrog` [233, 239] to link the deep sample and wide field target sample by characterizing the property modulation and selection effects of deep sources in DES wide field imaging. The `Balrog` sofware is used to inject simulated galaxies, which are model fits of randomly sampled DES Y3 deep galaxies, in the wide field single-epoch images multiple times at different positions on a grid. By processing these artificial images with injected sources through the photometric pipeline, one obtains a catalog that contains different wide realizations of the same deep galaxy. The injection procedure and the processing through the coaddition and measurement pipeline allows the `Balrog` catalog to inherit systematic effects that are either too difficult to forward model, positionally-dependent on the sky or unknown. This catalog is used to reconstruct the transfer function that relates deep photometry with wide field observations and is a key ingredient of the SOMPZ redshift calibration for DES.

4.2.4 Artificial wide data for MCCL: UFIG simulations

The forward modelling approach of MCCL requires a reliable fast image simulator. The Ultra Fast Image Generator (UFIG) is a software for simulating realistic survey images based on a parametric description of the galaxy population and the observational and instrumental effects of the specific survey. The galaxy population model is introduced in [188] and further explained and extended in [32, 120, 186]. Red and blue galaxies are drawn from two Schechter luminosity functions that evolve with redshift and are assigned a Spectral Energy Distribution (SED), which is a linear combination of the five KCORRECT template spectra [208] built using the Bruzual and Charlot stellar population synthesis models [240]. The coefficients are sampled from two Dirichlet distributions of order five, one for blue and one for red galaxies. The parameters of the Dirichlet distribution evolve with redshift to account for the evolution in galaxy colours. The galaxies are distributed randomly in the image, as clustering is currently not present. The simulated galaxies are then assigned a flux in the desired filter bands, as well as a Sersic light profile, with index n (different for red and blue galaxies) and half-light radius r_{50}. These radii come from a log-normal distribution with fixed standard deviation and absolute magnitude dependent mean. Galaxies are also assigned ellipticity components (e_1, e_2), based on a Beta distribution $p(|e|)$ with parameters e_{mode} and e_{spread}. The bright stars are assigned positions from Gaia Data Release 2 [241] and magnitude from the Besançon model of the Milky Way [199]. Fainter stars have random positions and come from the Besançon model. The bright stars (`MAG_AUTO` between 17 and 22) are used to estimate the PSF patterns using Convolutional Neural Networks as described in [204], which are then rendered on simulated images at pixel level. We produce DES coadds but simulate the coaddition process using the information about single exposures and CCD boundaries as described in section III.C of [127]. Other instrumental and observational effects such as gain of the detector, read-out noise and reddening are also rendered in the simulated images, with the aim of making them approximately comparable to real telescope images while maintaining a fast runtime. The UFIG simulations are heavily employed in the MCCL approach.

4.3 Methods

We apply the MCCL and SOMPZ methods to the described target sample in order to compare them and assess their agreement when constraining the tomographic redshift distributions of the same sample of galaxies. In the following subsections we outline the two methodologies and highlight some common features and differences.

4.3.1 MCCL

The MCCL method was first proposed by [27] as a forward modelling approach for calibrating cosmic shear measurements. The method was successfully applied for shear calibration in [191], determining the redshift distribution of a galaxy sample in [120] and for an end-to-end cosmic shear analysis in [127]. At the heart of the MCCL methodology lies the possibility of performing fast and realistic simulations of the survey of interest, that include instrumental and observational effects. This makes use of simulated survey images produced with UFIG [188, 190]. The simulator implements a parametric model of the galaxy population, summarised in Section 4.2.4 and described more extensively in Section 3.3.1 of Chapter 3. The key requirement for this simulation software is computational speed since the whole survey footprint is simulated multiple times throughout the analysis. The tuning of the parameters of the model is performed through an Approximate Bayesian Computation (ABC) analysis since the likelihood functions are intractable [242, 243]. We start from a conservative prior on the parameters and use UFIG to simulate several DES coadds, sampling the parameters from the prior. By evaluating the difference between the real images and simulations in terms of distance measures and selecting the parameters that minimize them, we are able to iteratively constrain the model. The ABC scheme is similar to that used in [186] and in the previous chapter. We underline the implementation details in Section 4.4.3. Once our ABC posterior is converged, we perform 40 simulations of the full survey of interest with parameters sampled from the posterior. We call the first simulation *fiducial* since it uses the posterior point that minimizes the combined distance measure. The simulations are analysed in the same way as the real images.

We extract magnitudes, sizes, ellipticities and colour distributions using SEXTRACTOR, which is run with the same settings on real and simulated images. We make sure that the simulated distributions span the same space of the DES real images. This step is crucial because it enables us to trust the fidelity of the UFIG simulations, identify the sources of discrepancies and possible extensions of the model. The redshift distributions are constructed from the true redshifts in the UFIG simulations matched with SEXTRACTOR detections, which depend on the ABC posterior parameters of the luminosity function, on the selection applied and on the colours for tomographic bin assignment. We summarize the steps of the MCCL redshift calibration:

1. use parametric models to describe the galaxy properties, observational and instrumental effects and generate realistic simulated images;
2. starting from an observationally motivated prior, constrain the galaxy population model parameters via simulation-based inference;
3. run several simulations of the full DES area with parameters sampled from the ABC posterior and check statistical agreement with real images;
4. obtain redshift distribution for each simulation, accounting for uncertainty.

4.3.2 SOMPZ

SOMPZ is a photometric approach to redshift calibration: it constructs a mapping between redshift and observed magnitudes and colours in a limited number of broadband filters. The method is characterized by the use of self-organizing maps to facilitate redshift calibration, by reducing the dimensionality in magnitude-colour space. In particular, we make use of the DES deep fields that overlap with near-infrared surveys to break degeneracies in the statistical colour-redshift relation [110, 112, 232]. This method is used with three main galaxy samples: the wide field DES Y3 target sample consisting of riz photometry for ~100 million galaxies, the DES deep fields consisting of $ugrizJHK_s$ photometry for ~3 million galaxies [232], and a redshift sample consisting of the intersection of the DES deep fields and spectroscopic and many-band photometric redshift surveys. SOMPZ uses the deep fields to determine the abundance of galaxies in $ugrizJHK_s$-band colour space and the redshift distributions in those regions of colour space to avoid selection biases introduced when reweighting redshifts according to only their abundance in $(g)riz$ colour space (see, e.g. [244]).

In practice, this inference method is facilitated by the use of two SOMs which classify the galaxies in the deep and wide samples into discrete classes, called *cells*, of colour and colour-magnitude space. The discrete classes are sometimes referred to as galaxy phenotypes ([110, 124]). The redshift distribution of a given galaxy assigned to wide SOM cell $\hat{c}$ is then given by:

$$p(z|\hat{c},\hat{s},\theta) = \sum_{c} p(z|c)\, p_{\texttt{Balrog}}(c|\hat{c},\hat{s},\theta)\, p(\hat{c}|\hat{s},\theta) \tag{4.1}$$

where z is redshift, c represents deep SOM cell indices, $\hat{c}$ represents wide SOM cell indices, $\hat{s}$ is the sample selection function, and θ represents any additional conditions such as position on the sky. By assigning wide SOM cells to tomographic bins, the redshift distribution of the tomographic bin can be computed by marginalizing (i.e summing) over the $n(z)$ of its constituent cells:

$$\begin{aligned} p(z|\hat{b},\hat{s}) &= \sum_{\hat{c}\in\hat{b}} p(z|\hat{c},\hat{s})\, p(\hat{c}|\hat{s},\hat{b}) && (4.2)\\ &= \sum_{\hat{c}\in\hat{b}} \sum_{c} p(z|c,\hat{c},\hat{s})\, p_{\texttt{Balrog}}(c|\hat{c},\hat{s})\, p(\hat{c}|\hat{s},\hat{b}). && (4.3) \end{aligned}$$

In this inference method, the `Balrog` image simulation software [233] plays the key role of determining the likelihood of a given deep many-band colour to be observed at a given region of noisier wide colour-magnitude space. The middle factor $p_{\texttt{Balrog}}(c|\hat{c},\hat{s},\theta)$, called the *transfer function*, encapsulates this information using multiple wide-field realizations of deep-field galaxies across the DES footprint to express the likelihood of a deep colour to be observed at a certain region of wide colour-magnitude space. This transfer function serves to correctly weight the well-constrained redshift distribution $p(z|c)$ of each deep SOM cell according to the probability of detecting those galaxies.

In addition to breaking degeneracies in the colour-redshift relation, this scheme facilitates avoiding otherwise prohibitive selection biases resulting from the use of spectroscopic redshifts for weak lensing redshift calibration (see, e.g. [244]) because it uses spectroscopic redshifts only of galaxies for which 8 bands of DES deep field photometry provide relatively well-constrained $p(z)$. In this work we use the pipeline from the application of SOMPZ to the DES Y3 source catalogue [112], with minimal changes presented in section 4.4.2.

4.3.3 Complementarity

The two methods are complementary, since they use different sources of information to calibrate the redshift distribution. A common aspect is the fact that both methods are aimed at measuring the redshift distribution of a sample of galaxies, rather than the redshift of a single galaxy.

MCCL relies on wide field images and on the possibility to simulate populations of galaxies that statistically agree with the ones in survey images. The parametric functions that are used in UFIG to model galaxy properties, such as magnitudes, sizes, spectral energy distributions and radial profiles, and render the objects on the image, are based on simple assumptions. The key modelling steps for redshift calibration are the luminosity functions and the construction of SEDs, parametrized in UFIG as a linear combination of 5 KCORRECT templates, with coefficients sampled from a Dirichlet distribution, with parameters evolving with redshift. This model has been validated against PAUS many band photometry [194] and spectroscopic data. [196] presented USPEC, a spectra generator which, following the same working principle as UFIG, successfully simulated a sample of luminous red galaxies comparable to SDSS/CMASS. Furthermore, [197] showed that the same galaxy population model can also reproduce galaxy properties (spectroscopic redshift distributions and stellar population properties) of bluer galaxies such as those in SDSS/CMASS Sparse. As the method uses the entire wide field imaging for constraining the ABC posterior, it is not sensitive to cosmic variance, as it is common for methods that use deep but small areas as redshift calibration samples.

Many cosmological and astrophysical phenomena are not yet accounted for, for instance the clustering of galaxies, which is currently being implemented [215]. The simple galaxy population model does not include galaxy morphological features, such as bulges and spiral arms, gas physics and time-varying objects such as AGNs. The simplifying assumptions could lead to biases, and one needs to be careful in checking the agreement with real images and whether it is necessary to extend the model. Furthermore, the method is designed to be fast and optimized for high performance computing but computationally expensive, since it requires multiple realistic simulations of the whole survey footprint. An advantage is that the methodology naturally produces a family of redshift posterior distributions.

SOMPZ leverages information from the deep fields, where deep photometry is available in more bands, in order to relate accurate redshifts to noisy wide field galaxies statistically. The galaxies in deep fields may be affected by cosmic variance, and this effect is quantified [112, 245]. A crucial improvement over other methods is the possibility to build a *transfer function* using `Balrog` that accounts for selection effects in the wide fields. `Balrog` injects model fits of deep galaxies in the wide field single-epoch images multiple times, and then performs coaddition and processing steps similarly to the DES Data Management team (DESDM) [246]. In this way, `Balrog` inherits systematics effects that are difficult to model and that could vary across the survey footprint. Due to the computational cost of injecting into single-epoch images, the `Balrog` run used for DES Y3 covers ~20% of the survey footprint and thus does not sample the full Y3 observing conditions. UFIG and `Balrog` have different approaches and serve different purposes: UFIG focuses on simulation speed and allows to capture systematic effects globally, limited to our understanding of the survey properties. The aim of the simulations is to forward model the sample and match its global properties. `Balrog`, on the other hand, captures local and unknown effects in the images, at a large computational cost and is aimed at characterizing measurement biases and selection effects.

A simplifying approximation required by the SOMPZ method is $p(z|c, \hat{c}, \hat{s}) \simeq p(z|c, \hat{b}, \hat{s})$, since very few galaxies belong to both c and $\hat{c}$, given the large number of possible cell combinations [112]. This causes a systematic offset in the estimated mean redshift in each bin, when validated against BUZZARD simulations ([112, 247]), which could be mitigated by injecting more deep galaxies in a wider footprint of the survey. The method could also benefit from targeted spectroscopic follow-ups, since some deep SOM cells remain empty or scarcely populated, or more narrow-band imaging which avoids selection biases. The characterization of the uncertainty requires modelling several contributions that affect mainly the deep and redshift samples and is highly non trivial [112].

4.4 Implementation

This section summarizes the changes to the two redshift calibration pipelines implemented in this work. We start by introducing the procedure to exclude most of the stars that remain in our target sample after the selection cuts. We proceed by describing the minor modifications to SOMPZ with respect to [112] and the changes to MCCL since [127] and [186], which include extensions of the galaxy population model in UFIG and improvements of the ABC scheme similar to Chapter 3. We show the resulting ABC posterior and the comparison between photometric properties of real and simulated data. Furthermore, we present the tomographic bin assignment strategy which is the same as [112] but is applied also to the MCCL pipeline, in order to assign the same galaxies to the same bins in the two methodologies to allow the comparison. We conclude with sections briefly discussing the $n(z)$ computation and uncertainty estimation.

4.4.1 Star galaxy separation

The selection used to define the target sample described in Section 4.2 excludes many of the stars, especially due to the `FLUX_RADIUS` cut in the i band. Nevertheless, we need to define a framework for star-galaxy separation which is applicable to both methodologies in order to reduce the stellar contamination of the target sample. In order to classify objects as stars or galaxies, we train a random forest classifier from `scikit-learn` [248] on the UFIG simulations, where we know which objects are simulated stars (`star_gal` flag). The classifier is trained using g, r, i and z band `MAG_AUTO`, `FLUX_RADIUS` and colours of 6 million objects balanced so that 50% of the objects are stars and 50% galaxies from a simulation of 1000 randomly selected tiles. In the training we do not apply any cuts. We validate the classifier on 3 million objects from the UFIG target sample (with SEXTRACTOR cuts). Purity and completeness of the galaxy sample are both ~ 99.3%. When we evaluate the classifier on real data from the unmasked Y3 Gold catalog, after selection and star-galaxy separation are applied, we observe that the fraction of objects with `EXTENDED_CLASS_MASH_SOF == 0` (high confidence stars and QSOs, see [230] for definition) is 0.3% whereas ~ 99% of the objects have `EXTENDED_CLASS_MASH_SOF == 3`. We thus obtain a stellar comtamination comparable to that of the weak lensing source catalog.

4.4.2 SOMPZ modifications

The SOMPZ methodology is used in the same way as in [112], with few exceptions. The first difference is the selection of the target sample, which requires updating the `Balrog` sample and implied transfer function with the same cuts. Furthermore, we use `FLUX_AUTO` from SEXTRACTOR instead of METACALIBRATION fluxes to compute the luptitudes (see [249] for definition and [110] for the usage in SOMPZ) used to train the wide SOM. This is motivated by the fact that the cuts are designed in SEXTRACTOR quantities, and the selected objects do not necessarily have well measured METACALIBRATION photometry. Furthermore, the usage of SEXTRACTOR photometry allows us to assign galaxies from UFIG simulations to the SOM for tomographic bin assignment and for cross-comparisons. The usage of SEXTRACTOR photometry also implies that we do not apply response and lensing weights, described in section 4.2 of [112], but only weigh by number of injections for the `Balrog` sample. We visualize the wide and deep self-organizing maps in Appendix 4.A.3.

4.4.3 MCCL modifications

The galaxy population model and ABC inference engine are very similar to those used in the HSC deep field analysis presented in Chapter 3. The updates to the morphology sector and to the parameters of the spectral energy distribution are implemented in this work, whereas the luminosity function parametrization follows [120, 127, 186]. We use the ABC procedure from Chapter 3, with slight modifications in the distance measures and MMD kernel width optimization. Changes that are unique to this work are the addition of a new parameter capturing the uncertainty in the number of faint stars and the use of imputation to fill in missing columns in galaxy catalogs during the ABC.

We describe each of these updates in greater detail below and show the resulting ABC posterior. After constraining the galaxy population model via ABC, we simulate the whole survey footprint 40 times, sampling the parameters from the ABC posterior. We compare the resulting photometric properties of the simulated galaxy target samples to DES real data in order to assess the agreement. The comparison is displayed at the end of Section 4.4.3.

Updated galaxy population model

The galaxy population model consists of two populations of galaxies: red and blue. Each of these types is modeled by dedicated set of parameters for luminosity, size, Sersic index, and ellipticity. The galaxies are sampled from two separate Schechter luminosity functions where the parameters $M^*(z)$ and $\phi^*(z)$ evolve with redshift with the functional forms

$$M^*(z) = M*_{\text{slope}}\, z + M^*_{\text{intcpt}} \tag{4.4}$$

$$\phi^*(z) = \phi^*_{\mathrm{amp}} \exp(\phi^*_{\mathrm{exp}} z). \tag{4.5}$$

The steepness of the faint-end slopes of the Schechter luminosity function are fixed to the default values $\alpha_{\mathrm{red}} = -0.5$ and $\alpha_{\mathrm{blue}} = -1.3$. The morphology of galaxies is described by three parameters $\log r_{50}^{\mathrm{intcpt}}$, $\log r_{50}^{\mathrm{slope}}$ and $\log r_{50}^{\mathrm{std}}$ separately for blue and red galaxies, as in the previous chapter. The Beta distribution describing galaxy ellipticities and the Betaprime distribution of Sersic indices is also unchanged. We draw the coefficients of the SED templates from a five-dimensional Dirichlet distribution separately for blue and red galaxies. The redshift evolution of the template coefficients is encoded in two sets of parameters at redshift $z = 0$ and $z = 1$ that we interpolate between. We use the new redundant parametrization of the SED coefficients presented in the previous chapter and separate the mean and variance of the Dirichlet distribution, with some minor modifications. We use parameters $\bar{\alpha}_i,\ i = 0,\ldots,4$ describing the mean (instead of the mode) of the Dirichlet distributions and α_{std}. This parameter corresponds to the standard deviation of the 5-dimensional Dirichlet coefficients with equal concentrations. The $\bar{\alpha}_i$ are normalized $\sum_i \bar{\alpha}_i = 1$, and the final Dirichlet coefficients are calculated as $\alpha_i = \bar{\alpha}_i \cdot 0.16\alpha_{\mathrm{std}}^{-2} - 1$. We keep the weights derived from [120] using the New York University Value-Added Galaxy Catalog, instead of using the prior derived in the ABC catalog-level preprocessing using the COSMOS2015 catalog presented in Section 3.A.1. We expect these choices to have small impact in a wide field, where most galaxies are relatively low redshift.

A parameter $\delta\sigma_{\mathrm{bkg}}$ scales the background noise level added to the simulated images, which is different in each pixel and derived from single exposures as in [127]. A new parameter $\delta N\star$ controls the total number of stars. It simply multiplies the number of stars that are being drawn from the Besançon model by $\delta N\star$. The magnitudes of stars are sampled with replacement, and the positions are assigned randomly within the HEALPIX pixel (`nside=8`, see [120, 127] for more details).

The priors were modified to save computing time by starting the inference in more plausible regions of parameters space. The summary of the model and priors is shown in Table 4.A.1 in Appendix 4.A.1. The final model has 46 parameters (compared to the previous chapter we fix α_{red} and α_{blue} and add $\delta\sigma_{bkg}$ and $\delta N\star$), out of which 4 are redundant.

Factorised ABC inference

We perform an Approximate Bayesian Computation in order to constrain the parameters of the galaxy population model. Our primary distance measure is the Maximum Mean Discrepancy (MMD) [207], which uses mean embeddings to measure distances between multi-dimensional probability distributions. The input vector of the MMD consists of magnitudes `MAG_AUTO`, radii `FLUX_RADIUS`, ellipticities computed using SExtractor windowed moments and flux fractions, which encode colours. The MMD distance is similar to the previous chapter but uses the total magnitude `MAG_AUTO`, includes radii and ellipticities in all bands $grizY$ and no redshift information, thus making the MMD input vector 20-dimensional. We complement the MMD with the fractional difference between the number of objects detected in the real data and simulations. Additional information is provided in Appendix 4.A.2, which also contains the description of the optimization procedure for the MMD kernel size.

We run prior-to-posterior iterations of the ABC algorithm, each on a unique subset of images. We divide all tile images into N sets d_i. These sets consist of increasing number of tiles: the first set has $|d_1| = 200$ tiles, and the last $|d_{17}| = 1204$ tiles. During an iteration i, for each parameter set θ_j that we sample, we randomly draw (without replacement) a small random subset s_i^j of tiles from d_i, $s_i^j \sim d_i$, and we simulate them. For the first iteration, we chose $|s_1| = 4$ tiles, and the last $|s_N| = 48$. We compute distances between the simulated tile and the corresponding DES tile (same tile identifiers). Then, the distance for a given sample is the average of distances of individual tiles. Due to this fact, the final distance can be considered a draw from a noisy version of the distance computed using all tiles in d_i, with noise reducing as $|s_i| \to |d_i|$. This way, we randomize both in the model parameter space and in the data space.

We ran 17 iterations of the algorithm with increasing number of tiles simulated. The details of these interations are shown in Table 4.A.2 in Appendix 4.A.2. In each iteration, we simulated 10000 samples from the prior and kept the 1000 samples with the lowest distance as the posterior. We then resampled the posterior using a Gaussian Mixture

Model (GMM) with 20 components and drew the 10000 samples that serve as prior for the next iteration. Before fitting the GMM, we performed a transformation to make the features more Gaussian.

We use a sub-selection of objects from the SEXTRACTOR catalogs to calculate the distances in order to remove stars and measurement failures. This selection is done imposing conditions in all bands $grizY$ on the `FLAG` field, the magnitudes, sizes, and shapes. With this number of strict conditions, the remaining selected sample is small. To address this, we require the object to be selected in at least 3 bands and use an *imputation* [250] algorithm[6] to fill in missing columns. This procedure increases the number of objects used in distance measure calculation by ≈ 30%. See Appendix 4.A.2 for more details on the configuration of ABC runs, sample selection, imputation and GMM density estimation of the posterior.

Finally, we removed a number of tiles due to two reasons: general SEXTRACTOR failures giving extremely high number of detections, and tiles in the area of the Large Magellanic Cloud. We were left with 10041/10338 tiles. Moreover, there were some tiles which always gave very high distances due to unknown effects. To account for this, we used the median as an averaging method for combining distances from multiple tiles. This way, our combined distances are more robust to outliers.

ABC posterior

We present the obtained ABC posterior in Figures 4.2, 4.3 and 4.4. We performed 17 iterations of the ABC algorithm (see Section 4.4.3 and Appendix 4.A.2). The model has 46 parameters, which we separate into: luminosity function parameters, SED template coefficients and galaxy morphology parameters.

Figure 4.2 shows the luminosity functions constraints for red and blue galaxies separately. The gray and coloured contours show the prior and the posterior, respectively. The posterior is generally fully contained in the prior, which indicates that our choice of prior should not affect the results significantly. The luminosity function parameters

---

[6] github.com/epsilon-machine/missingpy

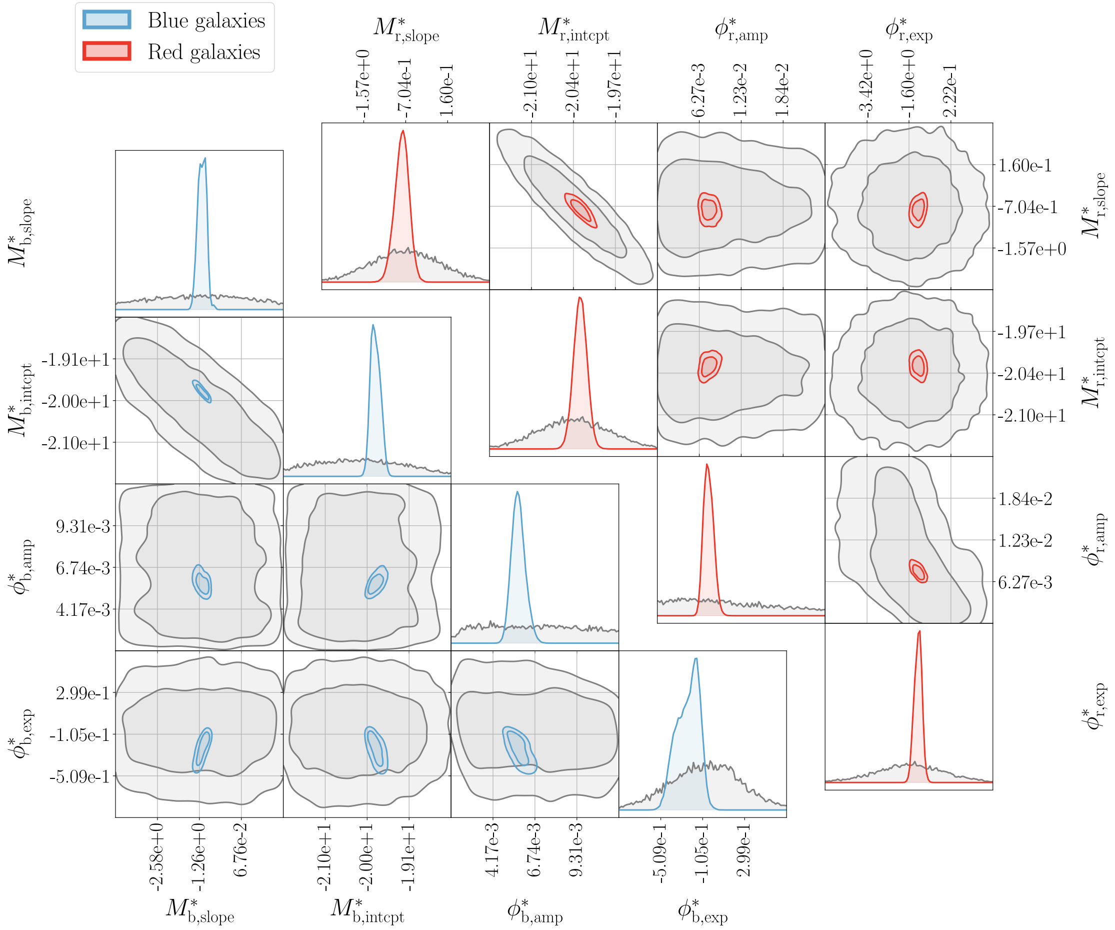


Figure 4.2: Luminosity function parameters from the prior (grey) and posterior (coloured) distributions. In the bottom triangle parameters for blue galaxies are shown in light blue, parameters for red galaxies are shown in red in the top triangle.

exhibit correlations between parameters $M*_{\text{slope}}$ and $M*_{\text{intcpt}}$, and $\phi*_{\text{amp}}$ and $\phi*_{\text{exp}}$, both for red and blue galaxies. We cannot directly compare the constraints to the previous chapter since the parametrization of the luminosity function parameters is different.

Figure 4.3 shows the SED template coefficients distribution, split into red and blue galaxies. The darker and lighter colours show the distribution at redshift $z = 0$ and $z = 1$, respectively. The parameters correspond to the following templates from KCORRECT: $\bar{\alpha}_0$ to the overall passive galaxy spectrum with contribution from young stars, $\bar{\alpha}_1$ to the emission line galaxy with strong Star Formation Rate (SFR), $\bar{\alpha}_2$ to the emission line galaxy, $\bar{\alpha}_3$ to the typical passive galaxy and $\bar{\alpha}_4$ to the post-starburst galaxy as indicated on the top of Figure 4.3. The five KCORRECT templates are shown in Chapter 3 in Figure 3.1. We note that for the blue galaxies, the weight of the two emission line galaxy templates $\bar{\alpha}_1$ and $\bar{\alpha}_2$ are larger than the prior mean and the post-starburst $\bar{\alpha}_4$ is also well represented. For red galaxies, the posterior is well within the prior range, with the strength of the passive galaxy template $\bar{\alpha}_3$ being increased. The evolution of the template coefficients is notable for the blue galaxies, especially for the dusty red template ($\bar{\alpha}_0$). That is expected from the galaxy evolution models, given the higher rate of star formation at higher redshift. We find very little evolution of the red galaxies with redshift, consistently with [251]. The parameters responsible for the variance of the template coefficients α_{std} differ with redshift for blue galaxies and is relatively stable for red galaxies. The distribution of coefficients at $z = 0$ is qualitatively similar to that obtained in the previous chapter but less constrained and in some cases bimodal.

Figure 4.4 shows the morphology parameters for red and blue galaxies. Additional panels for the parameters controlling the background noise level $\delta\sigma_{bkg}$ and number of stars $\delta N\star$ are duplicated in both corners. These two panels show additional secondary modes in the distribution. The intercept of the size-magnitude relation $\log r_{50}^{\text{intcpt}}$ is similar for the two populations, but we find a large difference in the slope $\log r_{50}^{\text{slope}}$. This is not unexpected, since red and blue galaxies undergo different processes in their formation. The Sersic index for blue galaxies peaks at $n_s^{\text{blue}} = 0.7$ and for red galaxies at $n_s^{\text{red}} = 3.5$, which is consistent with the literature. Ellipticity distribution mode e_{mode} is higher for blue galaxies, and the spread e_{spread} is slightly lower. The distributions

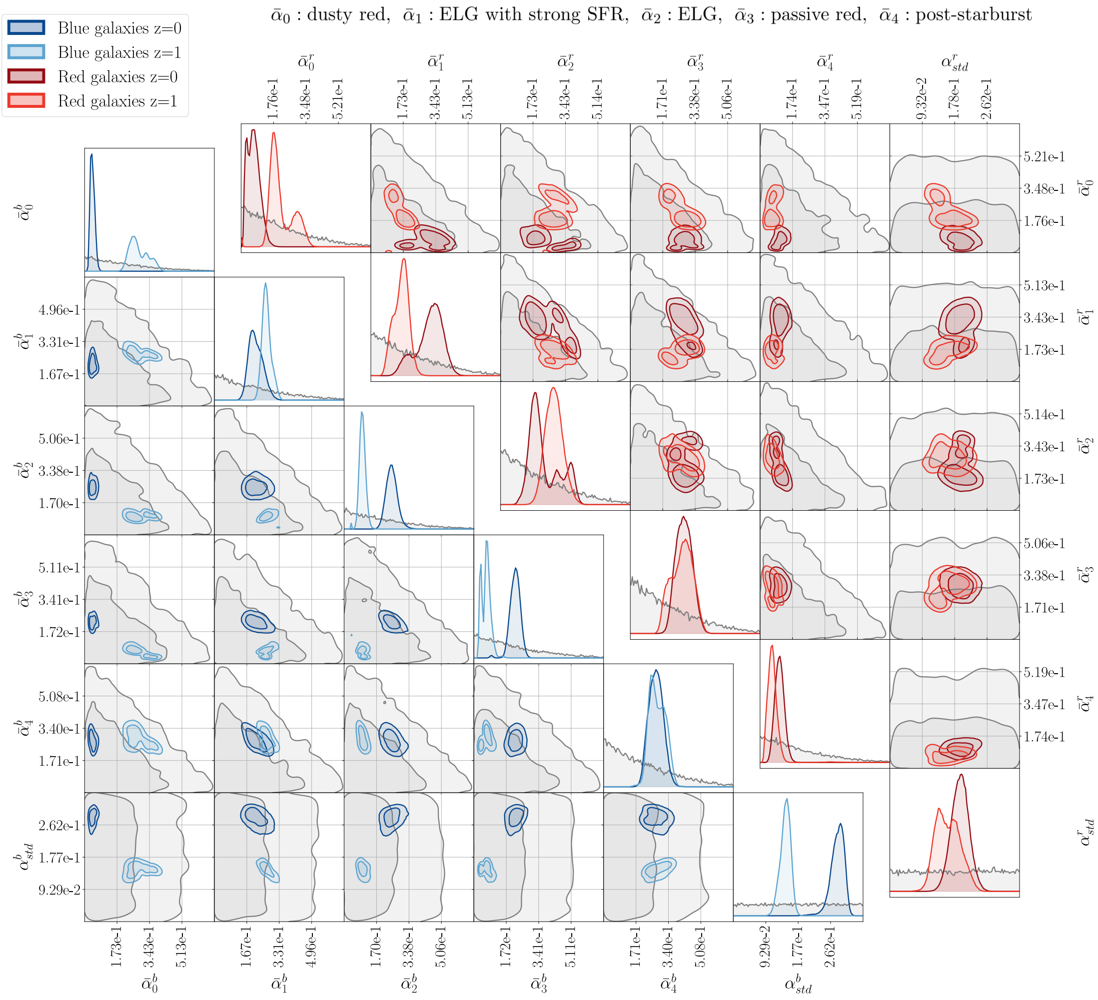


Figure 4.3: SED template coefficients from the prior (grey) and posterior (coloured) distributions. The bottom and top triangles show parameters for the blue and red galaxies, respectively. The dark colour shows the values of the z=0 parameters, while the lighter colour those for the z=1 parameters. Other redshifts use a linear interpolation/extrapolation between these two. Parameters $\bar{\alpha}_i$ are responsible for the means of the Dirichlet distributions, while α_{std} for the variance.

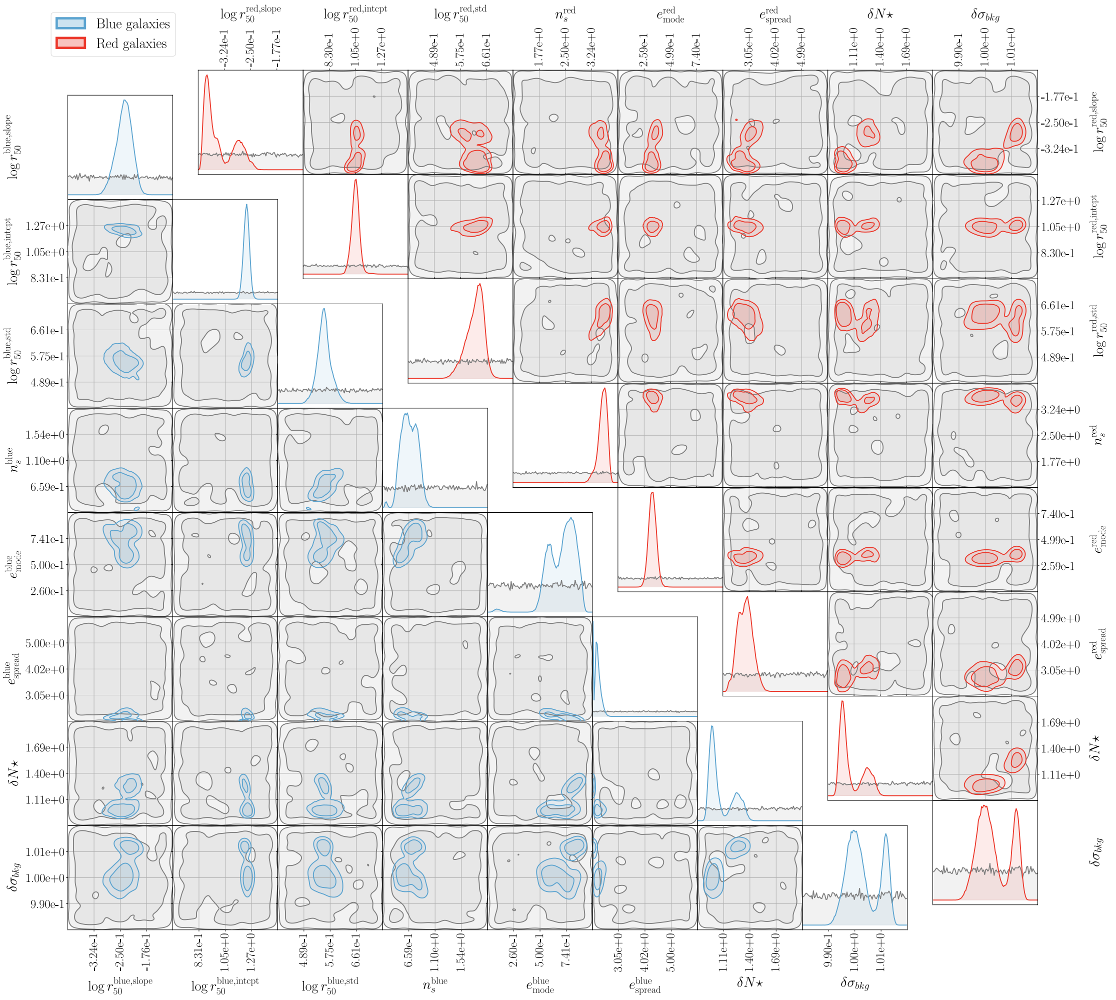


Figure 4.4: Galaxy morphology parameters from the prior (grey) and posterior (coloured) distributions. Additional parameters $\delta\sigma_{bkg}$ and $\delta N\star$ are duplicated in both corners: $\delta\sigma_{bkg}$ captures uncertainty in the level of pixel noise on the images, while $\delta N\star$ captures the uncertainty in the number of stars. Colours as in Figure 4.2.

for blue galaxies are broadly consistent with what we found in the previous chapter, whereas we find a difference in terms of $\log r_{50}^{\text{slope}}$ and $\log r_{50}^{\text{std}}$ for red galaxies. The Sersic index n_s^{red} is also considerably larger here than in the ABC posterior constrained on HSC deep fields. The differences are not unexpected since the sample used in the previous chapter is considerably deeper and constraining galaxy morphology is challenging.

Comparison between UFIG simulations and real survey data

We verify that the ABC posterior is able to produce a galaxy sample that statistically agrees with the one extracted from real images. The distributions of photometric properties that we compare are obtained by running SEXTRACTOR in dual-image mode using an riz detection image both for the real images and for 40 UFIG simulated surveys obtained by sampling the parameters from the ABC posterior. We show a tomographic comparison of magnitudes, radii and colour distributions in Figure 4.5. The assignment of galaxies to tomographic bins is outlined in Section 4.4.4. Under each histogram we display the percent residual in order to quantify the agreement. We only display a subset of the bands here for the tomographic setting and report the full $grizY$ corner plot in a non tomographic setting in appendix 4.A.4. We notice that the magnitudes agree in the global distribution to a better than 10% degree in most of the interval and the DES real data distribution lies in the space of the simulations. The agreement is prominent in the first bin and we observe some discrepancies in the higher redshift bins, mostly in the tails of the distributions, for a small percent of the galaxies. The distribution of `FLUX_RADIUS` in the r band, but the same applies to the other bands which are not included in Figure 4.5, is more discrepant: the agreement is limited to the peak of the distribution and degrades quickly in the tails where the real DES data is offset from the space of the simulations. The same trend was observed in the previous chapter, where the size distribution of simulated galaxies showed a tail of large objects that is not present in the data. The limited agreement in the tails is acceptable since the bulk of the distribution is consistent with simulations and sizes are a second order effect in the determination of the redshift distribution: their only impact on the $n(z)$ comes from the cuts determining which galaxies enter the target

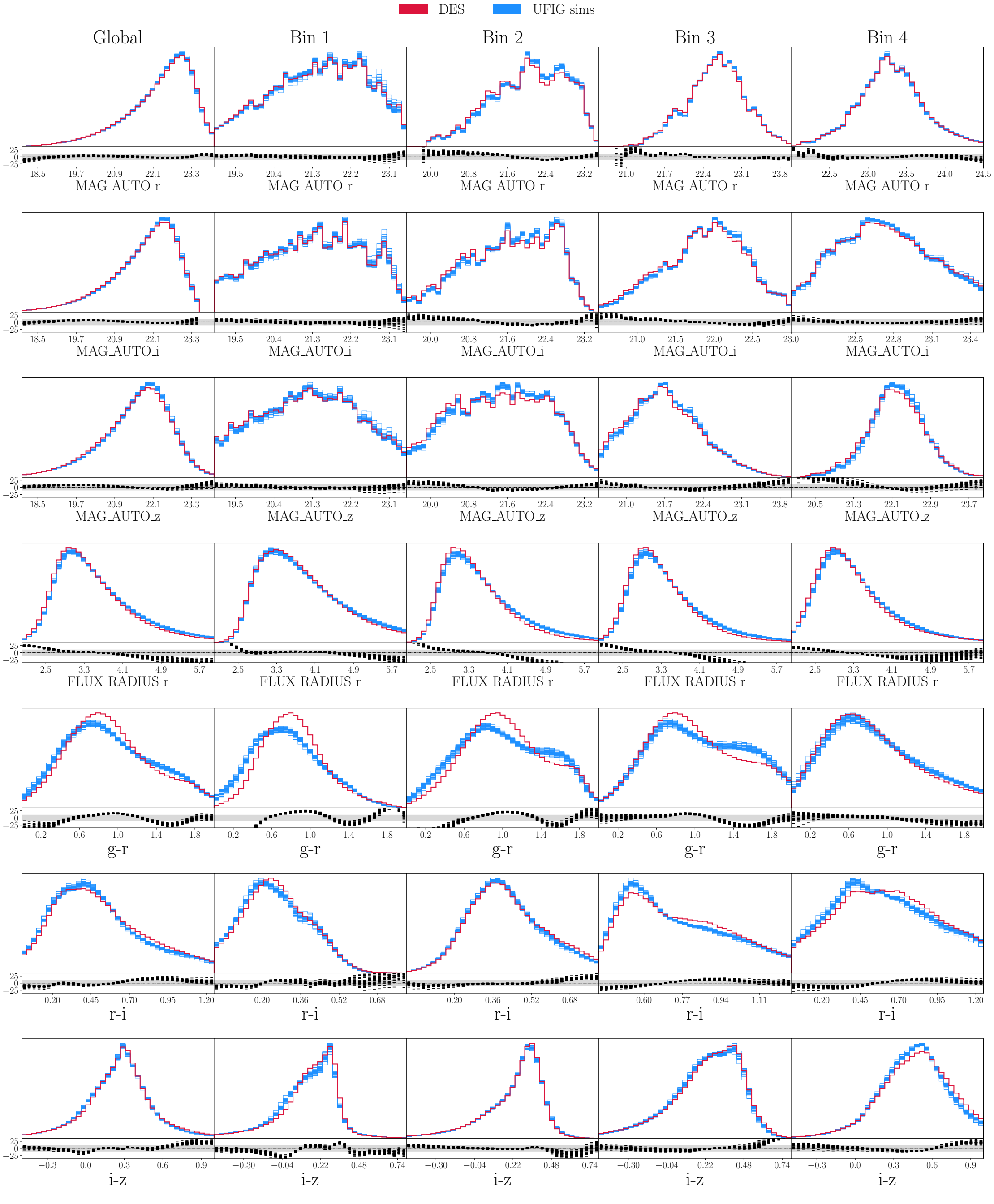


Figure 4.5: Tomographic comparison between the 1D distributions of photometric properties of 40 samples of simulated UFIG galaxies from the ABC posterior (in blue) and the real DES target sample (in red). In the rows we show normalized distributions of `MAG_AUTO` in r, i and z band, the `FLUX_RADIUS` in the r band followed by the $g-r$, $r-i$ and $i-z$ colours. The columns comprise of the global distributions and the 4 tomographic bins. Below each plot we show the percent discrepancy between each simulation and real DES data. The shaded grey area goes from −10% to +10%.

sample. We verify that the discrepancy in the size distribution has a negligible impact on mean redshifts in section 4.A.4, as this is part of our unblinding conditions. In future work, we plan to extend the model for galaxy sizes to include size evolution with redshift at fixed absolute magnitude. Figure 4.5 also shows a tomographic comparison of $g-r$, $r-i$ and $i-z$ colours. The $r-i$ and $i-z$ colour distributions agree sufficiently between real images and UFIG simulations, with a better than 10% agreement in most of the colour interval, with the exception of the $i-z$ colour at low redshift. The $g-r$ colour distribution shows considerable inconsistencies, especially at low redshift. Since the g band is not used for tomographic bin assignment and the g band selection is very broad, this is acceptable for the current work but requires further investigation in the future. It might point to the need to reassess the template spectra at the low wavelength end. We performed the tomographic comparison in all 5 bands ($grizY$) for magnitudes, sizes, colours and ellipticities. In Figure 4.6 we show the 68% and 95% contours of the global 2 dimensional magnitude-colour distributions. As expected from the 1D comparison, the main discrepancies involve the g - r colour whereas the other bands are in statistical agreement between real images and UFIG simulations. Given the overall agreement, this allows us to rely on the UFIG simulations from the ABC posterior to calibrate the redshift distribution.

4.4.4 Tomographic binning

Current cosmological surveys split both the source and the lens sample into tomographic bins. This improves the precision of the measurement of cosmological parameters by retrieving part of the three dimensional information. In order to compare the two redshift calibration methodologies in a tomographic setting, we need to assign the same galaxies to the same tomographic bins. We use 4 tomographic bins as in the fiducial DES Y3 source sample redshift calibration [112]. This is achieved by using the SOMPZ binning methodology in MCCL. We fix 5 bin edges, sort the galaxies from the redshift sample by their spectroscopic redshift or the median of their $p(z)$ (in case of many band photo-z) and split them according to the bin edges. We use the `Balrog` catalog matched with the redshift sample which contains wide and deep photometry and fiducial redshifts in order to assign wide cells to tomographic bins. In this matched

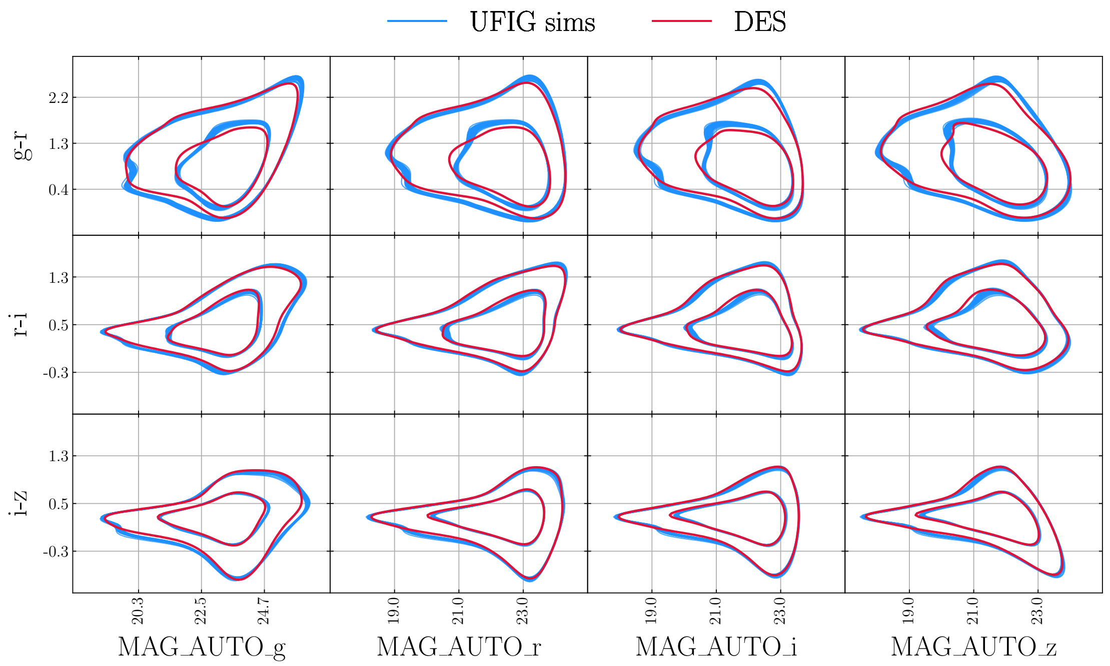


Figure 4.6: Two dimensional comparison of real survey data and 40 UFIG simulations. We show the 68% and 95% level contours of the global magnitude-colour distributions: g, r, i and z `MAG_AUTO` are on the x axis and $g-r$, $r-i$ and $i-z$ colours on the y axis.

catalog, each galaxy is assigned both to a wide cell $\hat{c}$ (which can differ for different wide realizations of the same deep galaxy) and to a deep cell c. We assign a wide cell $\hat{c}$ to the redshift bin to which most of its galaxies' redshifts belong. The bin edges are manually adjusted and the procedure is iterated until the 4 bins have roughly 25% of the wide-field galaxies each. The final boundaries of the 4 tomographic bins are [0.0, 0.37, 0.625, 0.86, 2.0].

Once we have trained the SOMs and we have determined which wide cells belong to each tomographic bin, we can import this information into MCCL. We do this by assigning both the simulated UFIG galaxies and the real galaxies from the target sample to the wide SOM cells and to bins according to $\hat{c}$. This tomographic bin assignment strategy differs to the fiducial MCCL method for assigning galaxies to bins, which is based on a random forest classifier.

4.4.5 Redshift distribution estimation

After defining the SOM cell assignment to redshift bins and their boundaries, we compute the *n(z)* distributions. For SOMPZ this means making use of the approximated equation

$$p(z|\hat{b},\hat{s}) \approx \sum_{\hat{c}\in\hat{b}} \sum_{c} p(z|c,\hat{b},\hat{s}) p_{\texttt{Balrog}}(c|\hat{c},\hat{s}) p(\hat{c}|\hat{s},\hat{b}). \tag{4.6}$$

$p(\hat{c}|\hat{s},\hat{b})$ is estimated by the wide field cell occupation, $p_{\texttt{Balrog}}(c|\hat{c},\hat{s})$ is the transfer function, which is computed using the `Balrog` sample, and $p(z|c,\hat{b},\hat{s})$ uses the overlap between the `Balrog` and redshift samples, where reliable redshifts, deep and wide photometry are available. The approximation takes the name of *bin conditionalization* because $p(z|c,\hat{b},\hat{s})$ is conditioned on $\hat{b}$, instead of $\hat{c}$, due to the large number of $c,\hat{c}$ pairs and thus scarcity of galaxies belonging to them. For MCCL, we simply use the true redshifts of the selected simulated galaxies that fall into a given wide SOM cell, assigned to bin b.

4.4.6 Redshift uncertainty estimation

The way to account for uncertainties in the redshift distribution is different in the two methodologies. MCCL reconstructs the *n(z)* posterior by using the redshifts obtained from UFIG simulations, which are sampled from the ABC luminosity functions for red and blue galaxies. By taking different samples of parameters from the ABC posterior, we produce a set of simulations and thus a family of possible redshift distributions. This naturally incorporates the uncertainty and recovers smooth large-scale redshift distributions, since clustering is not accounted for in UFIG simulations.

By contrast, SOMPZ produces both a fiducial and an ensemble of redshift distributions whose variation encodes the uncertainty due to biases in the redshift samples used, shot noise and sample variance in the deep fields, photometric calibration uncertainty, uncertainties due to the `Balrog` image simulation procedure, and the inherent limitations of the SOMPZ method [112]. Due to the significant differences in the origins of

Uncertainty	Bin 1	Bin 2	Bin 3	Bin 4
Combined Uncertainty: SOMPZ (from 3sDir)	0.012	0.008	0.006	0.009
MCCL: uncertainty from ABC posterior	0.029	0.018	0.011	0.022

Table 4.4.2: In the first row, we report the combined uncertainty (comprising of shot noise, sample variance, redshift sample uncertainty, `Balrog` uncertainty, photometric calibration uncertainty and method uncertainty) on the mean redshift in each tomographic bin from table 2 of [112] for SOMPZ. In the second row, the uncertainty in mean redshift obtained by running 40 simulations in MCCL. The edges of the tomographic bins are [0.0, 0.37, 0.625, 0.86, 2.0].

these uncertainties and the way they are constrained, each of them is incorporated using a different methodology. We defer the reader to [112] for details on how these sources of uncertainty are estimated and propagated.

Since our wide sample is designed to match the weak lensing source catalog closely and we use the same deep and redshift samples with the exact same deep photometry, we expect the uncertainties on mean redshift obtained in [112] to be a good estimate for our SOMPZ uncertainties. The use of these uncertainties is appropriate because the change from METACALIBRATION photometry to SEXTRACTOR photometry only affects the wide field fluxes, which are not a predominant contribution to the error budget. The only differences that could impact the uncertainty estimation is the wide field selection which alters the `Balrog` transfer function and the minor modification of the bin edges. We report the SOMPZ combined uncertainty from [112] in Table 4.4.2, together with the MCCL uncertainties.

To assess the fact that changing the redshift sample we calibrate on has a limited impact on mean redshift, we reproduce Figure 8 from [112] in Figure 4.7. We display the shift in mean redshift of our target sample in the four bins when calibrated on the SPC, PC, SC and C redshift samples relative to the mean of SPC, PC and SC. The shift in each bin is close to the result displayed in [112], with a less pronounced redshift dependence of Δ_z.

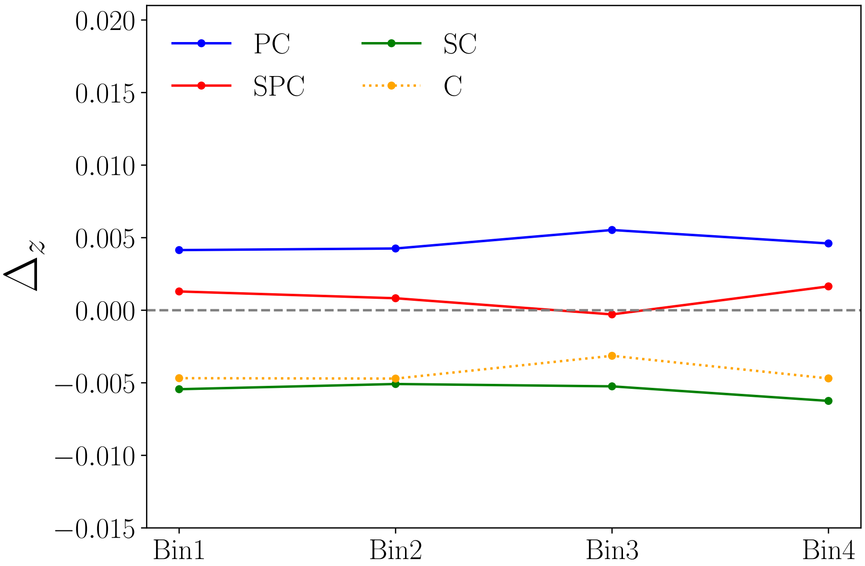


Figure 4.7: Mean redshift changes in the wide field redshift distribution in each tomographic bin, when calibrating on different redshift samples (SPC, PC, SC and C), relative to the average mean redshift of SPC, PC and SC.

4.5 Blinding

In order to avoid confirmation bias, which would make us favor analysis choices that lead to an agreement between redshift distributions from SOMPZ and MCCL, we performed a blind analysis. The blinding scheme we introduced was a simple shift of the redshift distribution in each tomographic bin by a random number uniformly distributed in $[-0.1, 0.1]$. This prevented us from comparing the mean redshifts but did not modify the overall shapes of the distributions. We defined a set of conditions that needed to be met by the two redshift calibration techniques before we could proceed to unblinding. The satisfaction of these conditions is verified in Appendix 4.A.4, where we explain the conditions more rigorously.

1. MCCL:
   a) the distances between simulations from ABC posterior and the DES data are consistent with the distances between ABC posterior simulations themselves;

   b) the distribution of photometric quantities of the real images lies in the space of the UFIG simulations;
   c) the residual discrepancies between the photometric distributions of real data and simulations have a negligible effect on mean redshift in each tomographic bin.
2. SOMPZ:
   a) the change in mean redshift, when switching between redshift samples is for each bin smaller than 0.015;
   b) the fraction of wide galaxies that come from deep cells without redshift is below 2%.

4.6 Unblinded results

The analysis was performed blinded and unblinded only when the unblinding criteria, qualitatively outlined in Section 4.5 and explained and verified in Appendix 4.A.4, were met. In this section we present the unblinded tomographic redshift distributions of the target sample obtained by running the SOMPZ and MCCL pipelines on a target sample from DES Y3 data. We show the retrieved $n(z)$ distributions, assess their consistency and then perform cross-comparisons of the two methodologies.

4.6.1 Tomographic redshift distributions

The tomographic redshift distributions of the target sample obtained with the SOMPZ and MCCL methodologies are displayed in Figure 4.8. We report the mean redshifts, the uncertainties of the two methodologies and the combined uncertainty in Table 4.6.3. We observe that the error bars of MCCL are roughly twice the error bars from SOMPZ. The estimation of uncertainties has already been explained in Section 4.4.6. A noticeable feature of Figure 4.8 is the fact that, while the MCCL redshift distributions are smooth, the shape of the SOMPZ $n(z)$ shows spikes due to the sample variance of

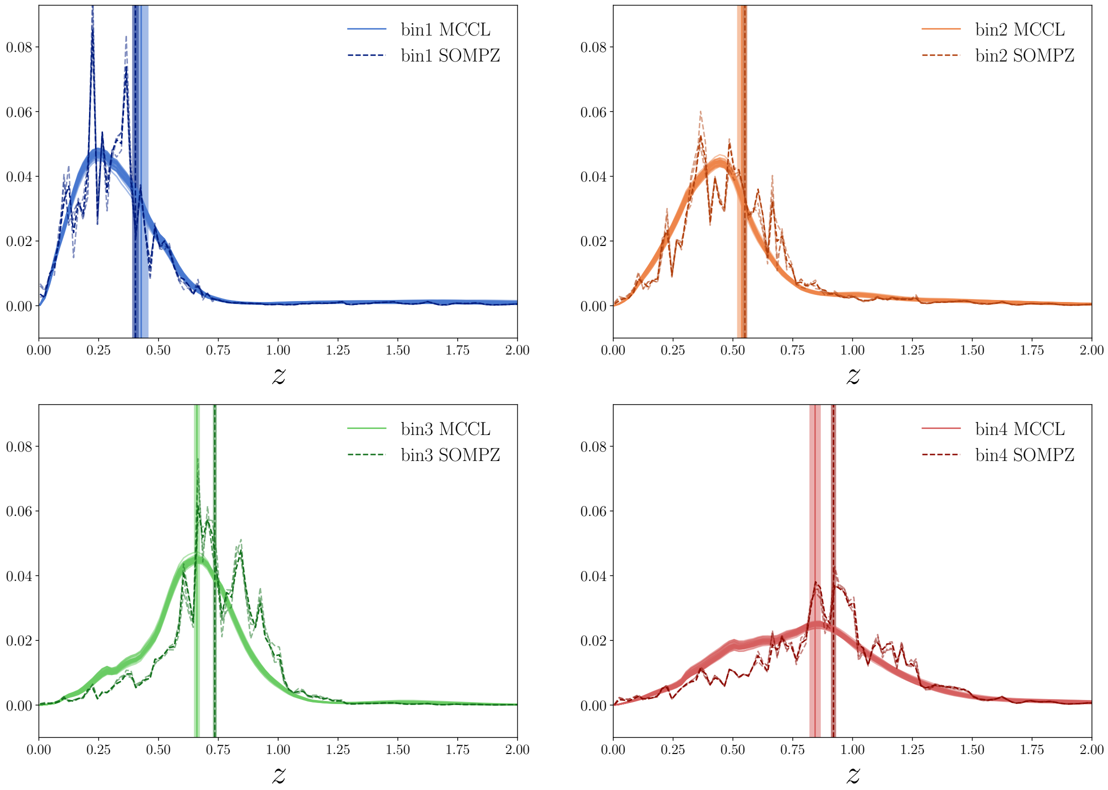


Figure 4.8: Redshift distributions in the four tomographic bins obtained with MCCL (solid line) and SOMPZ (dashed line). The MCCL lines show the posterior $n(z)$ from 40 simulations. For SOMPZ, the most visible dashed line is calibrated on the SPC redshift sample and the fading lines on SC, PC and C.

the redshift and deep samples. The redshift distributions are in good agreement at low redshift in the first two tomographic bins. For these two bins the means agree at the $\approx 1\sigma$ level. Tomographic bins 3 and 4 show a mean offset and also a difference in the shapes. MCCL predicts consistently lower mean redshifts in the higher redshift bins than SOMPZ, leading to a disagreement between the two methods. The third bin in particular shows a $\approx 7\sigma$ disagreement in the mean. Such a mean redshift offset has the potential to strongly impact the cosmological constraints from weak lensing, especially since most of the cosmological information comes from faint high redshift objects. We show and discuss the global redshift distributions in Section 4.8.

Tomographic bin	$\bar{z}_{\mathrm{SOMPZ}}$	$\bar{z}_{\mathrm{MCCL}}$	$\sigma_{\mathrm{combined}}$
Bin 1	0.404 ± 0.012	0.428 ± 0.029	0.031
Bin 2	0.550 ± 0.008	0.537 ± 0.018	0.019
Bin 3	0.735 ± 0.006	0.661 ± 0.011	0.012
Bin 4	0.920 ± 0.009	0.843 ± 0.022	0.024

Table 4.6.3: Means of the redshift distributions from SOMPZ and MCCL are reported per tomographic bin, together with their uncertainty. The combined error $\sigma_{\mathrm{combined}}$ is reported in the last column.

4.6.2 Detailed comparison

We perform cross-comparisons between the two methods in order to gain deeper understanding and shed light on the disagreement at high redshifts. We use the wide self-organizing map trained by SOMPZ as described in Section 4.3.2, which is displayed in Appendix 4.A.3, and assign UFIG simulated galaxies to it. In this way, we can use the SOM to explore the colour distribution of wide field galaxies and the colour-redshift relation implied by the two methodologies.

Figure 4.9 shows the mean and standard deviation of the redshift distributions in each wide SOM cell (first and second rows), as well as the wide cell occupation as a percentage of the total number of galaxies (third row). The left column of Figure 4.9 is populated with the DES Gold target sample and shows redshifts calibrated by SOMPZ, whereas the central column is populated with the simulated target sample of galaxies from the fiducial UFIG simulation and thus calibrated by MCCL. The last column on the right shows the difference between the first two. The cells therefore contain the same galaxy phenotypes, since the DES and UFIG target samples are simply assigned to the trained wide SOM, but the redshift determination is performed using the two different techniques. We observe that the mean redshift has a very similar structure in both SOMs, with maxima and minima in approximately the same locations. The maximum redshift difference in a cell is $\Delta z \approx 0.15$, and the negative $\Delta_z = \langle z|\hat{c}\rangle_{\mathrm{UFIG}} - \langle z|\hat{c}\rangle_{\mathrm{SOMPZ}}$ are more abundant and have higher values. Furthermore, the standard deviation is similar between SOMPZ and MCCL, with a prevalence of higher variance cells in SOMPZ. The cell occupation is in good agreement overall but shows some local differences. In order to assess the impact of the

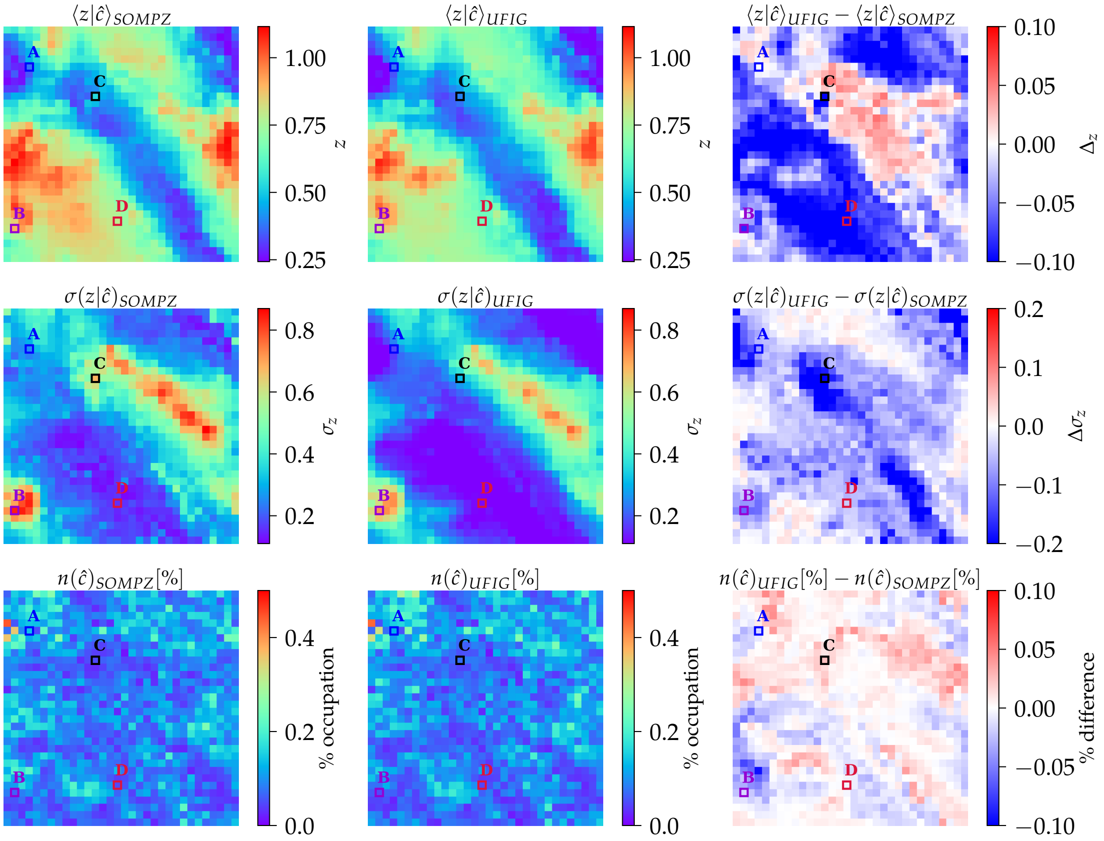


Figure 4.9: Mean redshift, standard deviation of the redshift distribution and percent cell occupation in each wide SOM cell obtained by SOMPZ and from the target sample from the fiducial UFIG simulation assigned to the same SOM. The third column displays the difference between the two approaches. We highlight cells of which we show the redshift distribution in Figure 4.11.

differences in the distribution of the abundance of galaxies in three-band colour-magnitude space between UFIG simulations and real data, we use the UFIG populated wide SOM to reweigh the SOMPZ $n(z)s$ in the four tomographic bins. We repeat this procedure for the 40 UFIG simulations and obtain mean shifts from $\bar{z}_{\mathrm{SOMPZ}}$ of $[0.018 \pm 0.005, 0.004 \pm 0.001, -0.002 \pm 0.001, -0.022 \pm 0.001]$ in the four redshift bins. These shifts are only significant in bins 1 and 4 and, while they ameliorate the agreement between SOMPZ and MCCL, they only partly explain the mean offset in bin 4 and do not impact the disagreement in bin 3. Differences in wide field colour-magnitude distribution alone cannot explain the offset between the two methods at high redshifts.

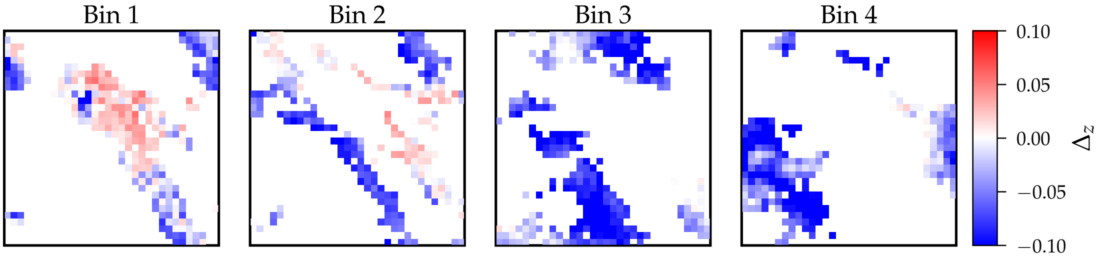


Figure 4.10: The differences in mean redshift between SOMPZ and MCCL are displayed in each SOM cell. Each heatmap shows the cells assigned to one of the tomographic bins.

We show the Δ_z separately for different tomographic bins in Figure 4.10. We see that, while the first two bins contain both positive and negative Δ_z, bins 3 and 4 are dominated by negative Δ_z. Furthermore, the area where the standard deviation of redshift in MCCL is very smooth (around label **D** in Figure 4.9) is almost fully contained in bin 3, which is the most constrained redshift bin in MCCL and the most discrepant when compared to SOMPZ. We only showed a comparison between real data and the fiducial UFIG simulation but very similar conclusions can be drawn for the other simulations from the ABC posterior.

We now look at the redshift distributions in wide SOM cells highlighted in Figure 4.9:

- **A**: with consistent mean redshift and standard deviation estimation;
- **B**: where the redshift estimate is different but the standard deviation is very similar;
- **C**: where both the mean redshift and standard deviation disagree;
- **D**: which is contained in the low standard deviation area of MCCL.

In Figure 4.11 we show for each of the four cells the calibrated SOMPZ $n(z)$s (dark green line), and the redshift distributions of UFIG galaxies from the target sample assigned to the same cell in two different UFIG simulations: the fiducial simulation (orange) and a second simulation (sim 1, purple). Overall, we observe very similar distributions for cells **A**, **B** and **C**, including the bimodalities. Cell **D** has both a mean offset and a

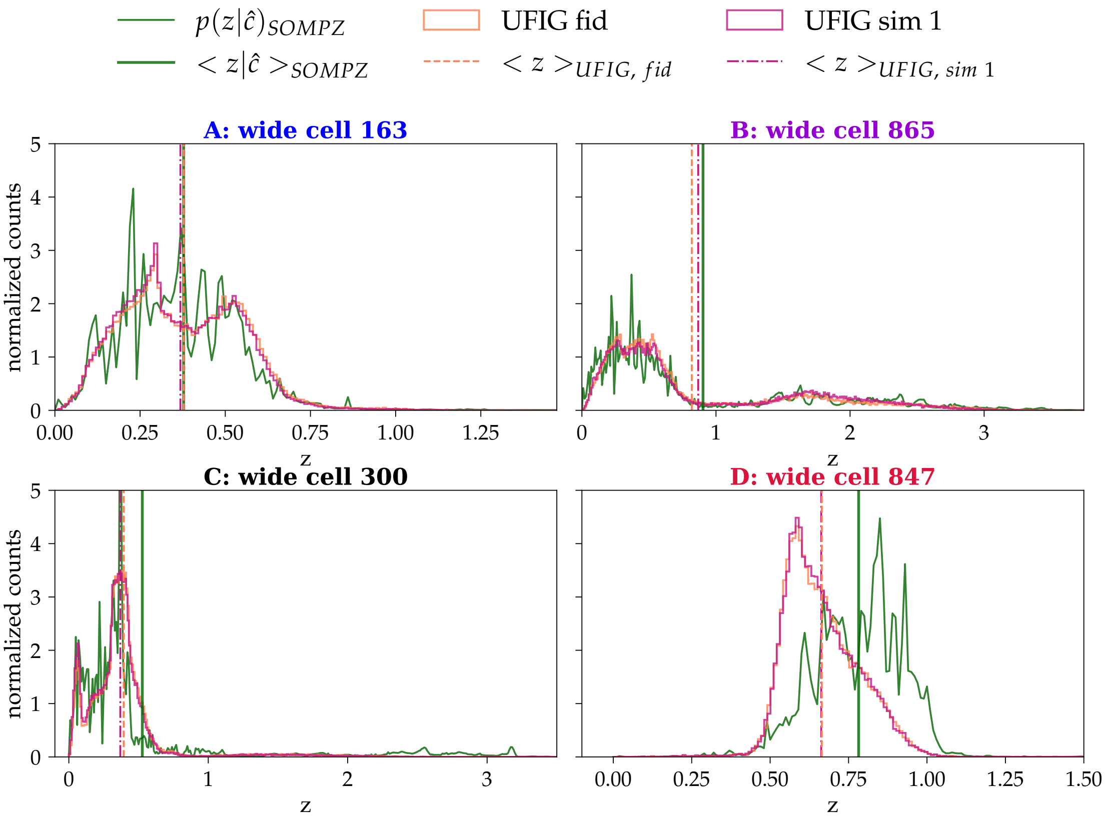


Figure 4.11: Redshift distributions in the four wide SOM cells highlighted in Figure 4.9 calibrated by SOMPZ (dark green line) and from the galaxies assigned to the cell by two UFIG simulations (fiducial, solid orange line - simulation 1, purple). The mean redshift in the cell is displayed as a orange dashed line (UFIG fiducial), a dashddotted purple line (UFIG simulation 1) and a green solid line (SOMPZ).

discrepant shape. We notice that the discrepancy in mean redshift in cell **B** is driven by the bimodality of the $n(z)$ in the cell, which contains a bulk of galaxies at low redshift and a high redshift tail. Since the number of galaxies in a single wide cell is limited, ~ 10000 objects per wide cell, the mixtures of the two modes can differ, causing the observed mean redshift discrepancies. We also observe a variability between different UFIG simulations in the composition of the two modes of the distribution, as can be seen by comparing the fiducial simulation and simulation 1 in cell **B**. Sample variance, especially in the deep sample, can have an impact on the redshift calibration. In cell **C**, we observe that the $n(z)$ estimated by SOMPZ and MCCL are roughly in agreement, but the difference in mean redshift is driven by a small but very high redshift tail in the

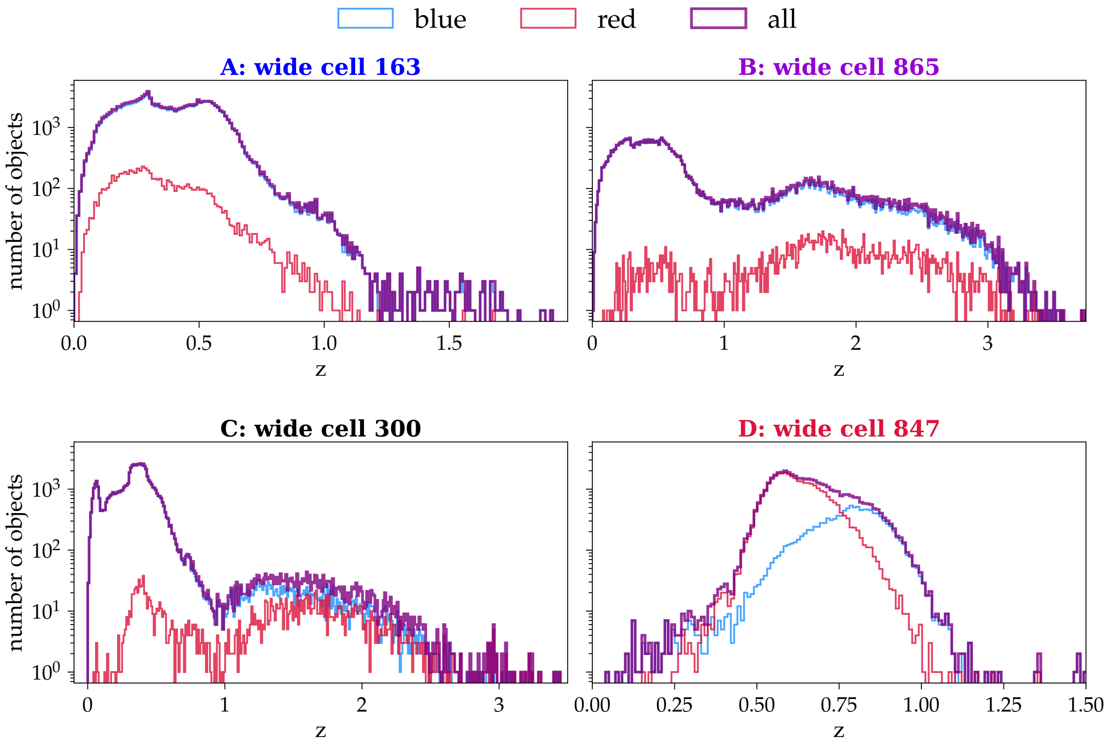


Figure 4.12: Histograms of galaxies' redshifts from the fiducial UFIG simulation (solid purple line) in the four wide SOM cells highlighted in Figure 4.9. The redshift distributions of blue and red galaxies are shown in blue and red respectively.

SOMPZ $n(z)$. Cell **D** displays discrepant $n(z)$s, which we investigate in the remaining part of the section.

In Figure 4.12 we show the histograms of UFIG redshifts from the fiducial simulation in the same cells as Figure 4.11 but additionally separate galaxies coming from the red and blue luminosity functions. We use a logarithmic scaling on the y-axis to make small populations of galaxies visible and only show the fiducial simulation to avoid overcrowding the plot (the conclusions for other simulations are similar with minimal composition changes). Cells **A**, **B** and **C** are dominated by blue galaxies. Cell **D**, on the other hand, contains both blue and red galaxies, with a majority of reds. The bulk of blue galaxies is at redshift $z \approx 0.8$ thus aligning with the SOMPZ calibration, whereas the red galaxies peak at $z \approx 0.5$. We only show this for cell **D** but several cells dominated by red galaxies display the same behaviour. This points to the red galaxy population as the driver of the discrepancy between SOMPZ and MCCL.

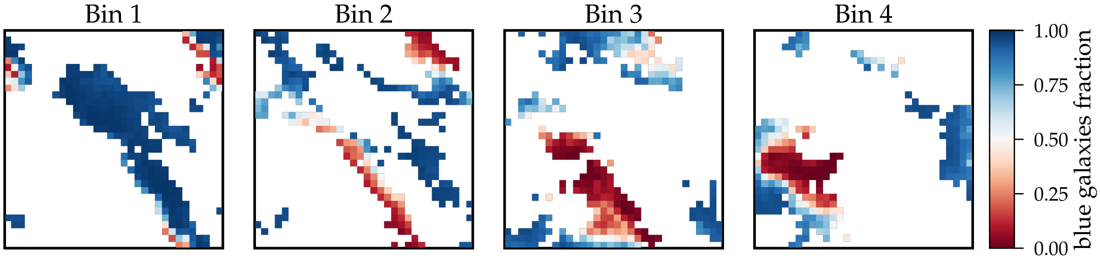


Figure 4.13: Fraction of blue galaxies in the fiducial UFIG simulation in each wide SOM cell, separated into tomographic bins. The colour goes from dark red (100% red galaxies) to dark blue (100% blue galaxies).

In order to verify this hypothesis, we show in Figure 4.13 the fraction of blue galaxies per wide cell in each tomographic redshift bin from the fiducial UFIG simulation. We observe that, while most objects in bin 1 are blue, the red galaxies are dominating some parts of the wide SOM in bins 2, 3 and 4. In particular, the low standard deviation area of MCCL around cell **D** consists of red galaxies and is mostly assigned to bin 3 and, for a smaller part, to bin 4. By looking at Figures 4.10 and 4.13, we notice that cells that are predominantly populated by red galaxies have large negative Δ_z, especially in bins 2 and 3. In bin 4 the cells with most negative Δ_z contain a mixture of red and blue galaxies. The colour-redshift relation of red galaxies is thus identified as a main driver of the discrepancy between MCCL and SOMPZ at high redshifts, in combination with the wide-field abundance of objects in colour-magnitude space in the case of bin 4.

4.7 Results with HSC deep fields ABC posterior

Following the developments from Chapter 3, we run 40 simulations of DES Y3 from the ABC posterior constrained using HSC deep fields data and COSMOS2020 photometric redshifts (from here on *HSC ABC posterior*) [32]. In order to save computing time, we only run on 10% of the DES tiles chosen at random. We verify that this is sufficient by running the fiducial simulation both for the full area and 10% of the tiles and comparing the results. We process the data and extract the target sample. The tomographic comparison of photometric properties of the target sample in 1D is shown in Figure 4.14 and the global 2D contours in Figure 4.15. The level of agreement with real

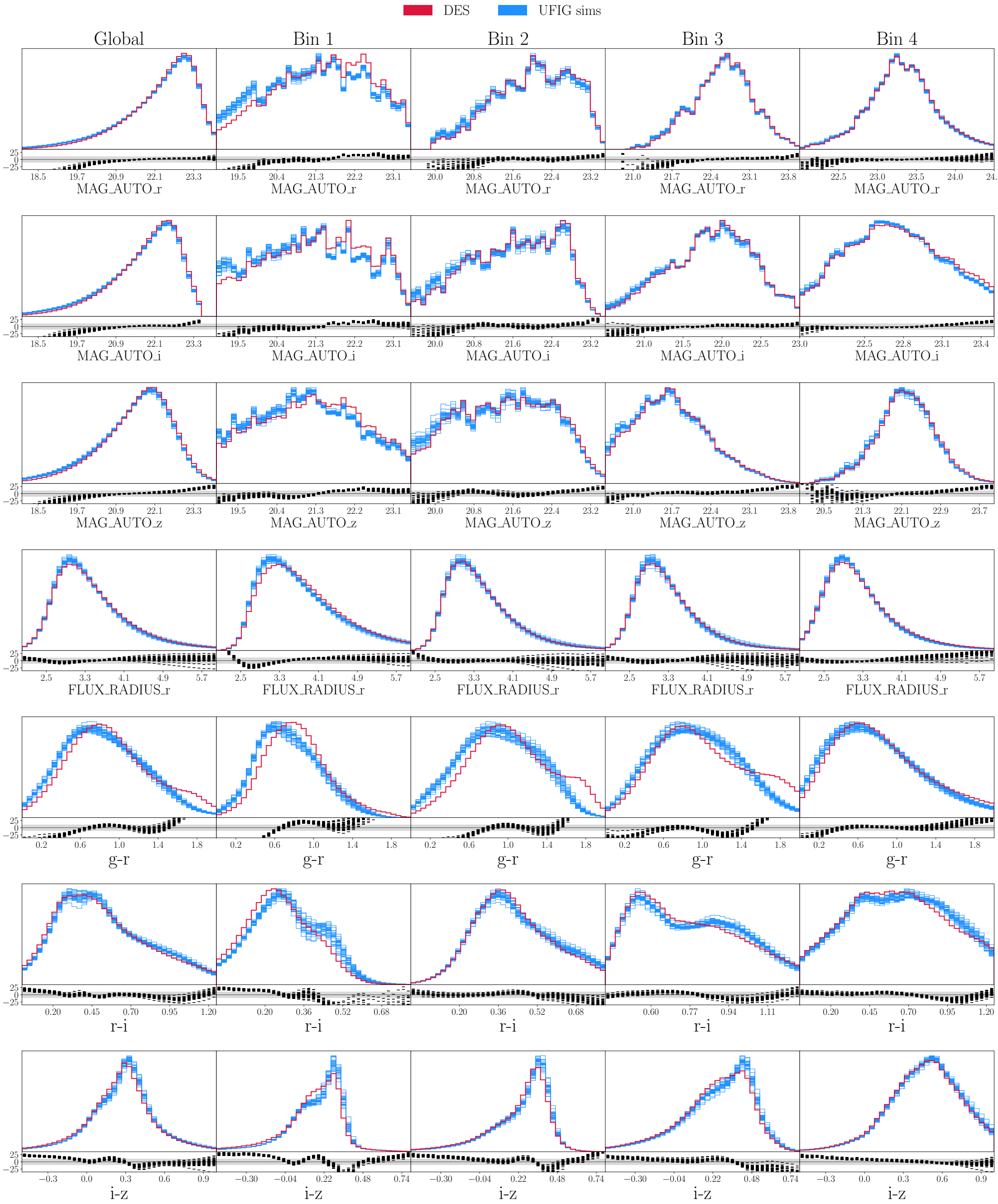


Figure 4.14: Tomographic comparison between the 1D distributions of photometric properties of 40 samples of simulated UFIG galaxies from the *HSC ABC posterior* (in blue) and the real DES target sample (in red). In the rows we show normalized distributions of `MAG_AUTO` in r, i and z band, the `FLUX_RADIUS` in the r band followed by the $g-r$, $r-i$ and $i-z$ colours, the columns comprise of the global distributions and 4 tomographic bins. Below each plot we show the percent discrepancy between each simulation and real DES data. The shaded grey area goes from −10% to +10%.

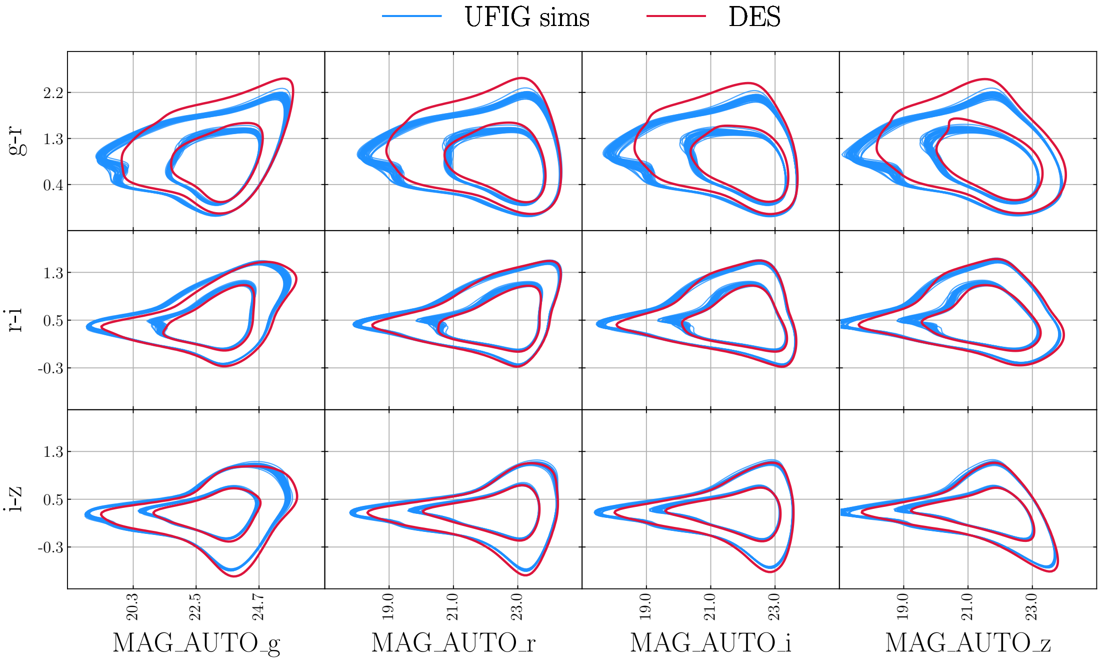


Figure 4.15: Two dimensional comparison of real survey data (in red) and 40 UFIG simulations from the *HSC ABC posterior* (in blue). We show the 68% and 95% level contours of the global magnitude-colour distributions: g, r, i and z `MAG_AUTO` are on the x axis and $g-r$, $r-i$ and $i-z$ colours on the y axis.

data is similar to what we obtained in Section 4.4.3 using the ABC posterior calibrated on DES wide field data, with magnitudes in the r and i band agreeing marginally worse and magnitudes in the z band moderately better. The agreement between the radii in the higher redshift bins is also improved, possibly due to the smaller impact of the PSF in HSC deep fields allowing better constraints on the size model. The colours maintain a similar level of agreement to Section 4.4.3, with $g-r$ showing the largest discrepancies. Overall, the simulations are in quite good agreement with real data in terms of photometry, indicating that the galaxy population model as constrained in the *HSC ABC posterior* can be used to simulate realistic DES Y3 wide field data.

We now look at the tomographic redshift distributions obtained using the *HSC ABC posterior*. We show the four $n(z)$ distributions of the target sample from SOMPZ and the new MCCL run in Figure 4.16. We notice that the overall trends and shapes of the MCCL redshift distributions do not differ substantially from those obtained from the ABC posterior constrained using DES Y3 data shown in Figure 4.8. There

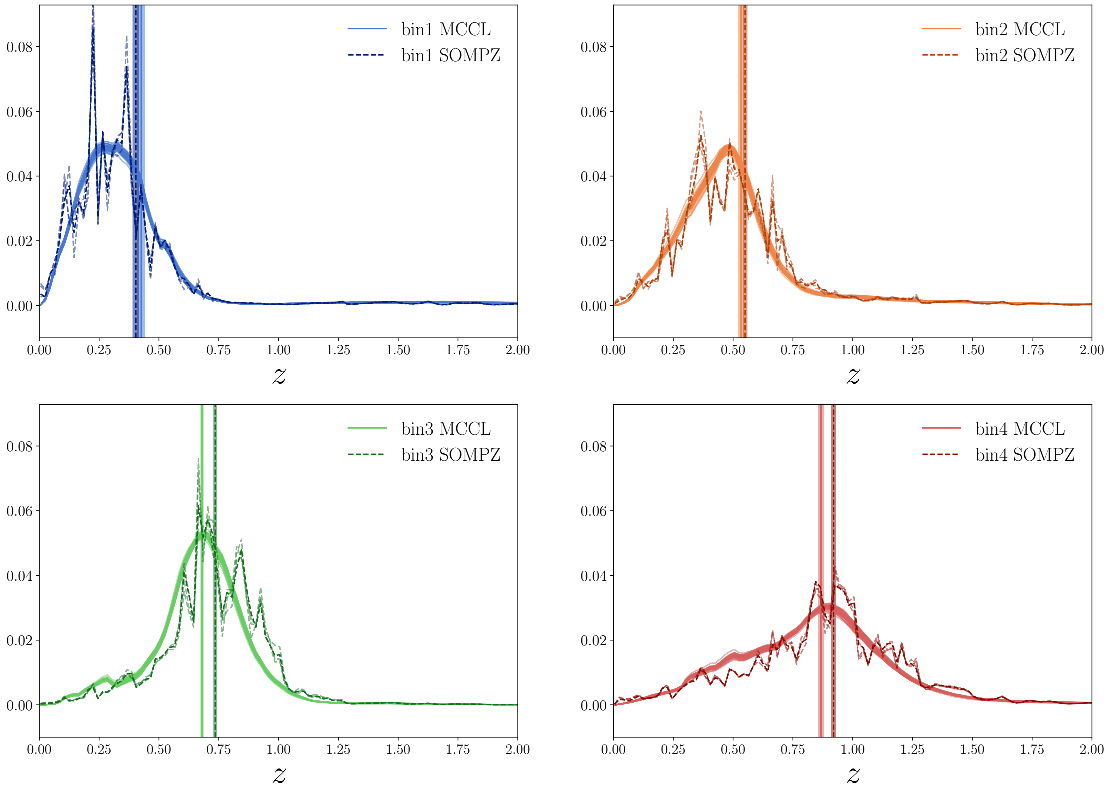


Figure 4.16: Redshift distributions in the four tomographic bins obtained with MCCL using the *HSC ABC posterior* (solid line) and SOMPZ (dashed line). The MCCL lines show the posterior $n(z)$ from 40 simulations. For SOMPZ, the most visible dashed line is calibrated on the SPC redshift sample and the fading lines on SC, PC and C.

are improvements in the third and fourth redshift bins concerning the shape of the distributions: the low redshift bumps (respectively at $z \approx 0.25$ in bin 3 and $z \approx 0.5$ in bin 4) are now less prominent. We report the resulting mean redshifts and error

Tomographic bin	$\bar{z}_{\mathrm{SOMPZ}}$	$\bar{z}_{\mathrm{MCCL}}$	$\sigma_{\mathrm{combined}}$
Bin 1	0.404 ± 0.012	0.426 ± 0.015	0.019
Bin 2	0.550 ± 0.008	0.533 ± 0.011	0.013
Bin 3	0.735 ± 0.006	0.680 ± 0.005	0.008
Bin 4	0.920 ± 0.009	0.867 ± 0.010	0.014

Table 4.7.4: The means of the redshift distributions from SOMPZ and MCCL are reported per tomographic bin, together with their uncertainty. The MCCL run consists of 40 simulations using the ABC posterior calibrated on HSC deep data and COSMOS2020 photo-zs. The combined error $\sigma_{\mathrm{combined}}$ is reported in the last column.

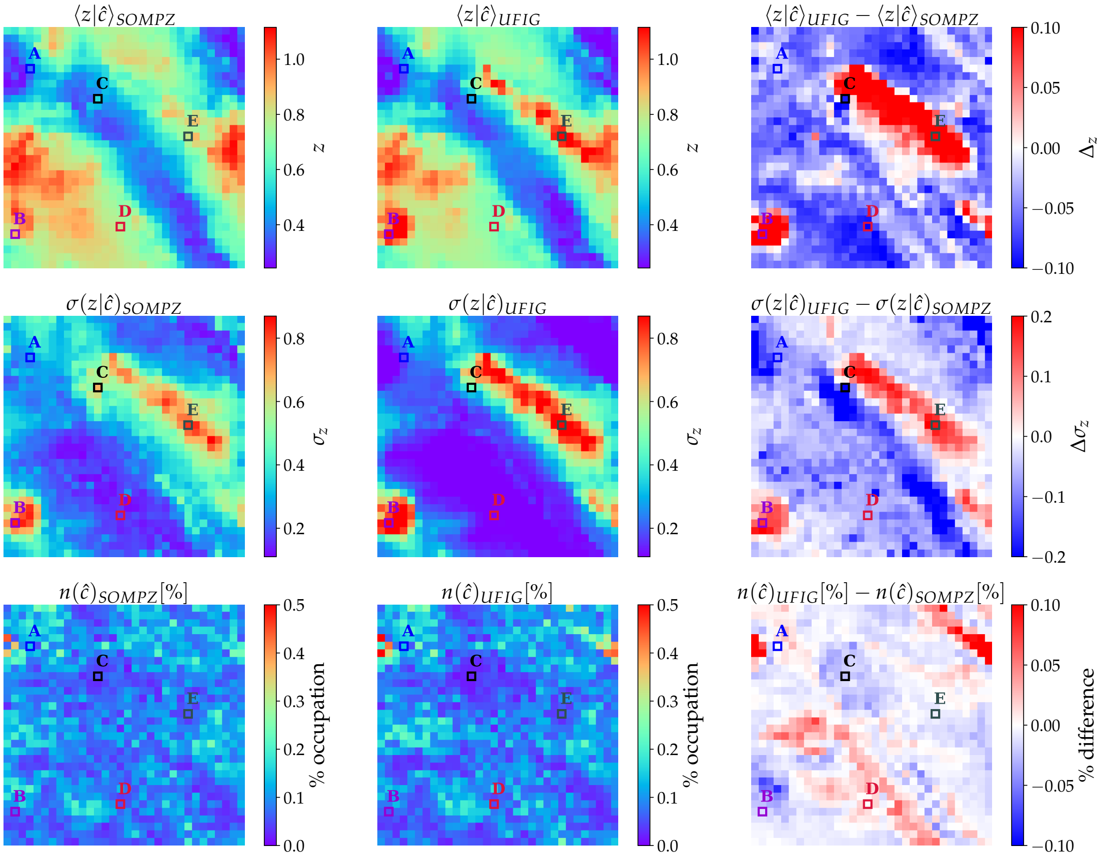
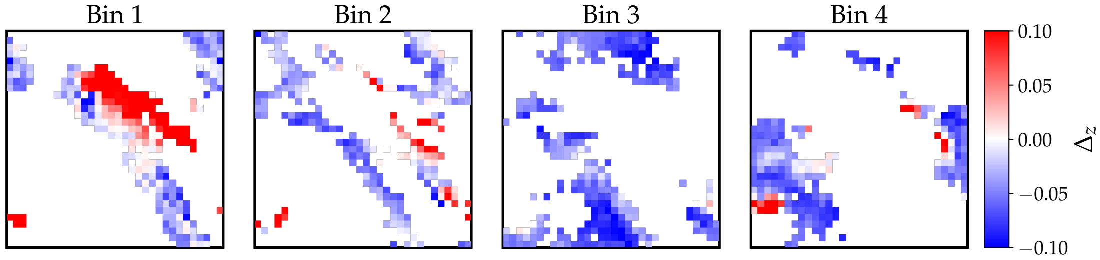


Figure 4.17: Mean redshift, standard deviation of the redshift distribution and percent cell occupation in each wide SOM cell obtained by SOMPZ and from the target sample from the fiducial UFIG simulation from the *HSC ABC posterior* assigned to the same SOM cell. The third column displays the difference between the two approaches.

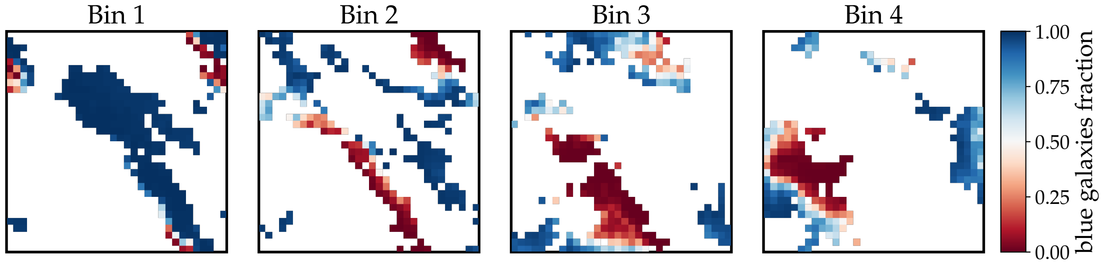
bars in Table 4.7.4, together with the unchanged SOMPZ results. The change in mean redshift between the two runs is most significant in bin 3. While the mean redshifts per tomographic bin of MCCL and SOMPZ are now closer, the decreased uncertainty in the MCCL mean redshift estimates makes the disagreement similarly significant.

We present the SOM-level comparisons between SOMPZ and the MCCL fiducial simulation from the *HSC ABC posterior* in Figures 4.17 and 4.18. We highlight cells for which we show the redshift distributions in Figures 4.A.6 and 4.A.7 in Appendix 4.A.5 and in Figure 4.19 for cell **E**.

We report the mean redshifts and standard deviations of redshifts in each wide SOM



(a) The differences in mean redshift between SOMPZ and MCCL are displayed in each SOM cell. Each heatmap shows the cells assigned to one of the tomographic bins.



(b) Fraction of red and blue galaxies in the fiducial UFIG simulation in each wide SOM cell, separated into tomographic bins.

Figure 4.18: We show the SOM-level comparisons between SOMPZ and the fiducial UFIG simulation from the *HSC ABC posterior*.

cell from both the SOMPZ direct calibration and the fiducial UFIG simulation using the *HSC ABC posterior* in the first two rows of Figure 4.17. The third row displays the wide cell occupation as a percentage of the total number of galaxies from the DES target sample and the target sample from the fiducial UFIG simulation. The last column displays the difference between the first two columns per wide SOM cell. We repeat our test of the wide-field three-band colour-magnitude abundance by reweighting the SOMPZ tomographic redshift distributions according to the cell occupation found in the 40 UFIG simulations. We obtain mean shifts from $\bar{z}_{\mathrm{SOMPZ}}$ of $[0.002 \pm 0.002, 0.001 \pm 0.001, 0.001 \pm 0.001, -0.016 \pm 0.001]$ in the four redshift bins. In this case the shift is only significant in bin 4, where it ameliorates the mean redshift agreement but does not completely explain the discrepancy between the two methods. In Figure 4.18 (a) we show the differences in mean redshift in each cell Δ_z separetely in the four tomographic bins. In Figure 4.18 (b), we display the fraction of blue galaxies per wide SOM cell and also separate the four tomographic bins. We notice that the positive Δ_z between SOMPZ and MCCL are now larger (up to $\Delta_z \approx 0.45$), especially in

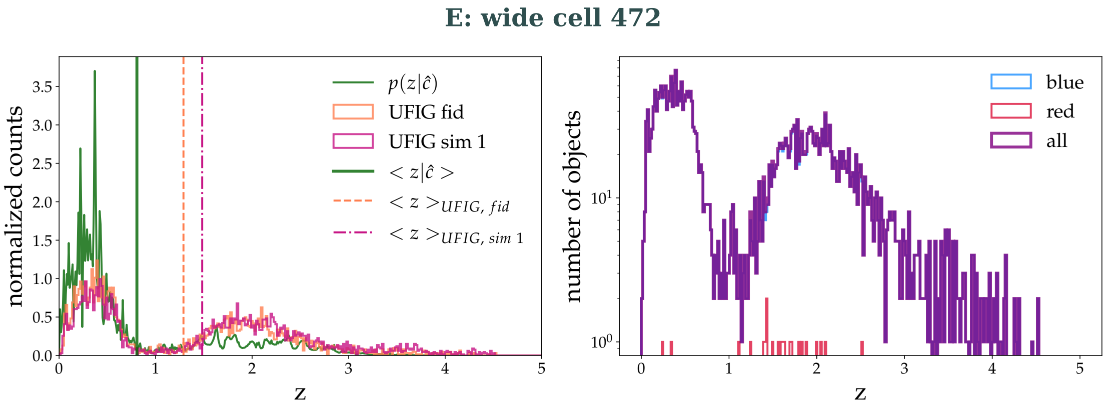


Figure 4.19: On the left, the redshift distributions in wide SOM cell **E** highlighted in Figure 4.17 from the SOMPZ calibration (dark green) and two UFIG simulations (solid orange and purple lines) from the *HSC ABC posterior*. The redshift distributions from the blue and red UFIG galaxies are shown in blue and red respectively in the figure on the right, together with the overall distribution in purple.

the SOM cells assigned to the first tomographic bin. This discrepancy in the mean does not affect the overall agreement of the first tomographic redshift bin, probably due to the small number of galaxies assigned to these cells or because they are compensated by other cells with negative Δ_z. We look at the $n(z)$ in one of these cells (highlighted with the new label **E**) in Figure 4.19. On the left hand side we display the redshift calibration from SOMPZ and two UFIG simulations, and on the right hand side we split red and blue galaxies from UFIG. We see that the cell is populated by blue galaxies in UFIG and presents a very strong bimodality. UFIG produces more galaxies in the high redshift mode compared to the weight assigned to it by SOMPZ. Since the galaxies extend to redshift $z \approx 4$, the mean of the distribution in this wide SOM cell is strongly affected. We noticed an excess of high redshift galaxies in the tail of the distributions in Chapter 3, which will be investigated in future work.

4.8 Global redshift distributions

To highlight the changes between the two MCCL runs and assess the agreement between MCCL and SOMPZ regardless of tomographic bin assignment, we show in

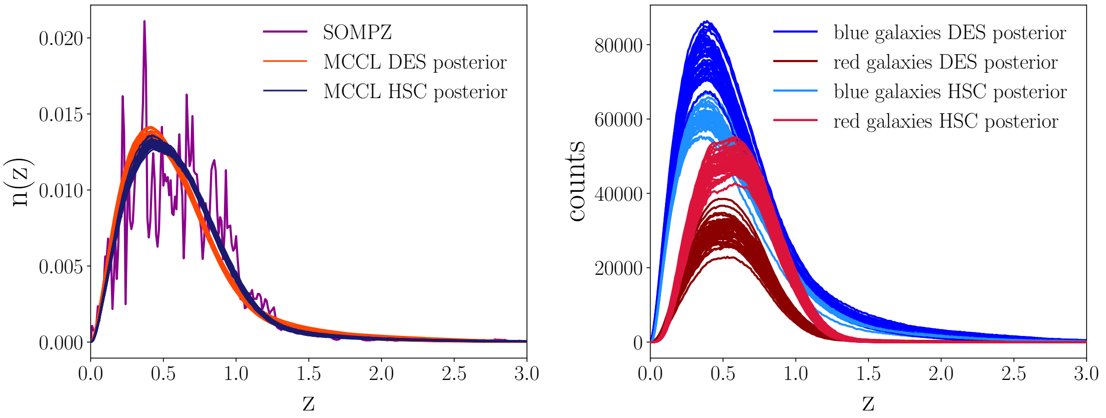


Figure 4.20: The global redshift distribution of the target sample obtained from SOMPZ (in purple) and MCCL ran using both ABC posteriors calibrated on DES (orange) and HSC+COSMOS2020 (dark blue) is shown on the left. On the right hand side, we display the histograms of red and blue galaxies from MCCL from 40 simulations sampled from the two posteriors, normalized so that the total number of galaxies is set to 10 millions.

Figure 4.20 the calibrated global redshift distributions of the target sample. In the left panel of Figure 4.20, we display in purple the global $n(z)$ obtained from SOMPZ calibrated on the SPC redshift sample and in orange and blue the MCCL $n(z)$s obtained from the two different ABC posteriors, constrained by DES wide field photometry and HSC deep field data and COSMOS2020 photo-zs. We estimate the error in mean redshift from SOMPZ approximately using a number weighted quadrature sum of the errors in the four tomographic bins and report means and uncertainties in Table 4.8.5. We notice that the change in the MCCL mean redshift resulting from changing to the *HSC ABC posterior* is $\approx 1\sigma$. As already observed for the tomographic redshift distributions, the agreement between the means of the global redshift distributions obtained by MCCL and SOMPZ is roughly unchanged when changing ABC posterior ($\approx 2\sigma$), but the shape of the global $n(z)$ from the *HSC ABC posterior* resembles the

$\bar{z}_{\mathrm{SOMPZ}}$	$\bar{z}_{\mathrm{MCCL,DES}}$	$\bar{z}_{\mathrm{MCCL,HSC}}$
0.652 ± 0.018	0.610 ± 0.018	0.625 ± 0.009

Table 4.8.5: Mean and uncertainty of the global redshift distribution of the DES Y3 target sample, as obtained by SOMPZ and MCCL. In the case of MCCL we report the results obtained using the two different ABC posteriors, calibrated on DES wide data and HSC deep data complemented with COSMOS2020 photometric redshifts.

SOMPZ redshift distribution to a better degree.

We also plot on the right hand side of Figure 4.20, the global redshift distributions for red and blue galaxies separately from both ABC posteriors. We observe a change in the red population, which has a higher mean redshift and larger redshift variance when we calibrate the galaxy population model on HSC deep fields and COSMOS2020. The blue population varies less in terms of shape of the $n(z)$, but the ratio of blue to red galaxies is different, with a larger abundance of red galaxies.

4.9 Conclusions

We presented a blind comparison between the phenotypic SOMPZ redshift calibration and the MCCL simulation-based inference method on DES Year 3 data. We compared the two methodologies to highlight similarities and differences and applied both to the same target sample. In the case of SOMPZ, we applied the method presented in [112] with minor changes determined by the selection cuts and the use of SEXTRACTOR photometry. In the case of MCCL, we introduced some modifications of the galaxy population model and ABC inference. After fulfilling the unblinding conditions, we compared the redshift distributions from the two methodologies in four tomographic bins, obtained by assigning wide SOM cells to each bin and populating the wide SOM with UFIG galaxies within the MCCL methodology. We found good agreement in the means and shapes of the distributions in the two lower redshift bins, whereas we found a discrepancy in mean and shape in tomographic bins 3 and 4. We further investigated the disagreement and found that the red galaxy population seems to drive it, in combination with differences in the three-band colour-magnitude distribution of galaxies. Finally, we repeated our analysis with the ABC posterior calibrated with HSC deep data and COSMOS2020 photometric redshifts from Chapter 3. We observed that the photometric properties of the DES Y3 target sample are reproduced reasonably well with this posterior. The tomographic redshift distributions from MCCL are shifted to higher redshifts but the tension between the two methodologies is not relieved due to the increase in precision. The agreement for the mean of the global redshift distribution remains at the 2σ level. The discrepancy is thus not significant for the global redshift distribution but is exacerbated by tomographic bin assignment for the higher redshift bins. This indicates a difference in the colour-redshift relation implied by SOMPZ in the DES deep fields and UFIG simulations.

One way to further investigate the origin of the discrepancy between the two methods would be to run SOMPZ on UFIG simulations. As opposed to Buzzard [251], UFIG generates full image simulations, so that the same SOMPZ procedure could be applied to real data and simulations. The `Balrog` injection could be mocked by producing the same images with different exposure times in UFIG to reduce computational costs. Deep and redshift samples could also be created with UFIG. This would require the extension of UFIG to infrared bands, which is currently under development for the simulation of KiDS-VIKING data [252].

As described in the conclusion of the previous chapter, there are several possible improvements of the MCCL methodology. One promising extension consists in the use of Stellar Population Synthesis (SPS) instead of empirical template spectra to generate galaxy SEDs (e.g. [212, 213]). This can affect the $n(z)$s and particularly have an impact on tomographic bin assignment, since it modifies the redshift-colour relation. We found that red galaxies in the third redshift bin have a limited variance in redshift; this could be caused by a lack of flexibility of the spectral templates, which were derived at low redshift from SDSS spectroscopic data, and their redshift evolution. An advantage of SPS is the possibility to compare the inferred constraints to recent measurements of physical properties of galaxies, such as the stellar mass function up to high redshifts (e.g. [253]). This approach has been applied successfully at catalog level by [121, 254] for the calibration of redshift distributions and other galaxy evolution observables. Furthermore, the addition of infrared bands in UFIG can be powerful for breaking the colour-redshift degeneracies via simulation-based inference of many-band data. One possibility would be to use self-organizing maps as distance measures between data and simulations.

The SOMPZ direct calibration will be enriched in the near future through the collection of targeted spectroscopic data to span the deep SOM introduced by [255]. This effort is already underway with multiple instruments, including DEIMOS, LRIS, and MOSFIRE on the Keck telescopes [256], the VLT optical and near-infrared multi-object spectrographs, FORS2 and KMOS [257], and DESI [258]. The 4MOST Consortium[7] though the complementary WAVES [259] and 4C3R2 [260] surveys will further constrain the colour-redshift relation by populating the deep SOM with high multiplicity. The addition of redshifts above $z \approx 1.6$ will be increasingly important, in view of LSST and Euclid. Other crucial aspects include the understanding and optimization of the selection functions of different surveys and the availability of deep field photometric data with equivalent depth in the infrared wavelengths.

4.A Appendix of chapter 4

4.A.1 Galaxy population model

We use a similar model as in [27, 120, 186] and the previous chapter, with some differences, described in Section 4.4.3. Here, we summarize the meaning of the different model parameters, report priors and allowed ranges in Table 4.A.1.

4.A.2 Details of the ABC runs

In this appendix, we describe the details of the ABC analysis. This consists of prior-to-posterior iterations, similarly to [194] and the previous chapter [32].

Optimization We start with the prior defined in Table 4.A.1. Then, we simulate a number of random tiles s_1^j for each of the samples j in the prior. The tiles are chosen from a unique set d_i for each iteration i, $s_i^j \sim d_i$, as described in Section 4.4.3. In each iteration, we vary the number of simulated tiles $|s_i|$, the size of the tile set $|d_i|$, the composition of the combined distance, and the random seeds. Table 4.A.2 shows the details of this process. Each iteration consisted of 10000 simulated samples, for which we calculated the distance measures (see Section 4.4.3). We selected the posterior by

---

[7] https://www.4most.eu/cms/home/

	Par.	Meaning	Prior
Luminosity functions (8)	$M^*_{\mathrm{b,slope}}$	Slope of the redshift evolution of the parameter M^* in the Schechter luminosity function for blue galaxies	Prior from [186] ×3, $\in [-4, 1.5]$
	$M^*_{\mathrm{b,intcpt}}$	Intercept of the redshift evolution of the parameter M^* in the Schechter luminosity function for blue galaxies	Prior from [186] ×3, $\in [-22, -18]$
	$M^*_{\mathrm{r,slope}}$	Slope of the redshift evolution of the parameter M^* in the Schechter luminosity function for red galaxies	Prior from [186], $\in [-4, 3]$
	$M^*_{\mathrm{r,intcpt}}$	Intercept of the redshift evolution of the parameter M^* in the Schechter luminosity function for red galaxies	Prior from [186], $\in [-23, 17]$
	$\phi^*_{\mathrm{b,exp}}$	Decay constant of the redshift evolution of the parameter ϕ^* in the Schechter luminosity function for blue galaxies	Prior from [186], $\in [-2, 1.5]$
	$\phi^*_{\mathrm{b,amp}}$	Amplitude of the redshift evolution of the parameter ϕ^* in the Schechter luminosity function for blue galaxies	Prior from [186], $\in [1.1{-}5, 1.2{-}2]$
	$\phi^*_{\mathrm{r,exp}}$	Decay constant of the redshift evolution of the parameter ϕ_* in the Schechter luminosity function for red galaxies	Prior from [186], $\in [-11, -7]$
	$\phi^*_{\mathrm{r,amp}}$	Amplitude of the redshift evolution of the parameter ϕ^* in the Schechter luminosity function for red galaxies	Prior from [186], $\in [2{-}8, 3.5{-}2]$
Galaxy morphology (12)	$\log r_{50}{}^{\mathrm{blue/red}}_{\mathrm{slope}}$	Slope of the evolution of the average intrinsic physical size of galaxies with absolute magnitude	$\mathcal{U}[-0.4, -0.1]$
	$\log r_{50}{}^{\mathrm{blue/red}}_{\mathrm{intcpt}}$	Intercept of the evolution of the average intrinsic physical size of galaxies with absolute magnitude	$\mathcal{U}[0.6, 1.5]$
	$\log r_{50}{}^{\mathrm{blue/red}}_{\mathrm{std}}$	Standard deviation of the normal distribution we use to sample intrinsic physical galaxy sizes	$\mathcal{U}[0.55, 0.6]$
	n_s^{blue}	Mode of the Sérsic index distribution of blue galaxies	$\mathcal{U}[0.2, 2]$
	n_s^{red}	Mode of the Sérsic index distribution of red galaxies	$\mathcal{U}[1, 4]$
	$e^{\mathrm{blue/red}}_{\mathrm{mode}}$	Ellipcitiy distribution mode for blue/red galaxies	$\mathcal{U}[0.01, 0.99]$
	$e^{\mathrm{blue/red}}_{\mathrm{spread}}$	Ellipcitiy distribution spread for blue/red galaxies	$\mathcal{U}[2.05, 10.0]$
SED coeff. (24)	$\bar{\alpha}^{\mathrm{blue/red}}_{i,0/1}$	Normalized Dirichlet concentration parameters at z=0/1 from which the template coefficients for blue/red galaxies are sampled, i=1,...,5, $\sum_i \alpha_i$=1	$\mathcal{D}[1, 1, 1, 1, 1]$
	$\alpha^{\mathrm{blue/red}}_{\mathrm{std},0/1}$	Standard deviation of the normalized Dirichlet concentration parameters at z=0/1 from which the template coefficients for blue/red galaxies are sampled	$\mathcal{U}[0.005, 0.35]$
Other (2)	σ_{bkg}	Scale factor for the background noise level in the simulations	$\mathcal{U}[0.8, 2.0]$
	$\delta N\star$	Multiplicative scaling of the number of stars	$\mathcal{U}[0.8, 1.2]$

Table 4.A.1: Table with galaxy population model parameters. The model has 46 parameters, but 4 of them are redundant: template coefficients $\bar{\alpha}_i$ are always forced to sum to $\sum \bar{\alpha}_i = 1$. The luminosity function parameters use same prior as [186], with standard deviation scaled by a factor of ×3 for blue LF M^* parameters. The last column shows the distribution (Normal, Uniform, or Dirichlet) of the prior, as well as the additional bounds applied. For all variables using the Uniform distribution, a joint Sobol sequence was used to generate the prior.

iteration	$N_{t,\text{sim}}$	$N_{t,\text{reserved}}$	sky area	distances
1 - 4	4	200	100 deg^2	$d_{ng}{\cdot}0.1 + d_{\text{MMD}}{\cdot}0.9$
5 - 6	8	401	200 deg^2	$d_{ng}{\cdot}0.1 + d_{\text{MMD}}{\cdot}0.9$
7 - 8	12	502	251 deg^2	$d_{ng}{\cdot}0.1 + d_{\text{MMD}}{\cdot}0.9$
9 - 10	16	602	301 deg^2	$d_{ng}{\cdot}0.1 + d_{\text{MMD}}{\cdot}0.9$
11 - 12	20	803	401 deg^2	$d_{ng}{\cdot}0.1 + d_{\text{MMD}}{\cdot}0.9$
13 - 14	24	803	401 deg^2	$d_{ng}{\cdot}0.2 + d_{\text{MMD}}{\cdot}0.8$
15 - 16	32	903	451 deg^2	$d_{ng}{\cdot}0.2 + d_{\text{MMD}}{\cdot}0.8$
17	48	1204	602 deg^2	$d_{ng}{\cdot}0.1 + d_{\text{MMD}}{\cdot}0.9$

Table 4.A.2: Iterations of the ABC algorithm, including number of reserved and simulated tiles, sky area and combination of distance measures.

choosing the 1000 samples with the lower distances. We used these samples to train a Gaussian Mixture Model of the posterior, and then used it to draw new 10000 samples. This resampled posterior is passed to the next iteration as the prior, and the process is repeated.

Sample selection The distance measures were calculated using the SEXTRACTOR catalogs, created in all $grizY$ bands, based on the riz detection image. This process is described in Section 4.2.1. We then ran the PSF estimation using the Convolutional Neural Networks [204], in the same way as in [127] and [32]. We ran SEXTRACTOR for the DES data first, and then for the simulated images during the ABC iterations. From the catalogs, we selected galaxies with a special set of cuts. These cuts are different from the ones defining the target sample presented in Section 4.2.1. The ABC distance selection cuts were as follows:

$$\begin{aligned} &\texttt{FLAGS} < 4, \quad 15 < \texttt{MAG\_AUTO} < 30, \quad 0.1 < \texttt{FLUX\_RADIUS} < 10, \\ &0 < \texttt{ELL} < 1, \quad \texttt{N\_EXPOSURES} > 0, \quad 0.65 < r_{50}/\texttt{PSF\_FWHM} < 2, \end{aligned} \tag{7}$$

where `ELL` is the absolute ellipticity calculated from windowed moments `**_WIN_IMAGE`, and `N_EXPOSURES` is the number of exposures in the coadd at the position of the object, r_{50} is the object size defined as $r_{50} = 2\cdot\log(2)\cdot(\texttt{X2\_WIN\_IMAGE}+\texttt{Y2\_WIN\_IMAGE})^{1/2}$, as in [127]. The last cut is applied to create a pure galaxy sample by selecting objects larger than the PSF. Note that we use forced photometry from a detection image, so for some bands the coadd image may be empty for a galaxy detected in the stack. We additionally require that the galaxy image did not lie on an image boundary in the coadd, setting flags: `FLAG_COADD_BOUNDARY` = 1 and `FLAG_IMAGE_BOUNDARY` = 1. These fields were calculated the same way as in [127].

Dealing with missing columns These criteria have to be satisfied in at least three out of the five bands, otherwise the object is removed from the catalog. For objects that have missing data in some bands, we use an *imputation* method from [250], using the K-Nearest Neighbours algorithm from the MISSINGPY[8] package. The neighbours are calculated in a transformed space, where the columns were scaled using `scikit-Learn` Robust Scaler[9]. The imputation process gives ≈ 30% more objects than a strict cut in all bands.

Distance measures The distances were calculated similarly to [32, 120, 186]. We use the following distances, with d_{comb} being the combined distance used to select the samples for the posterior:

- $d_{\mathrm{ng}} = (N_{\mathrm{SIM}} - N_{\mathrm{DES}})/N_{DES}$ is the fractional distance between number of objects in the DES data N_{DES} and simulations N_{SIM} that pass the cuts in Equation 7;
- d_{MMD} is the MMD distance between the simulated catalog and DES catalog. This distance was calculated using an array containing `MAG_AUTO`, `FLUX_RADIUS`, `ELL`, and flux fraction f_i (see Section 4.4.3), in each band. That gives a 20-dimensional vector for each object. Before calculating the MMD distance, each column was scaled, so that its mean and standard deviations are close to 1. We use the same scaling throughout the analysis for both DES and simulated data. The kernel size for this distance was optimized, see paragraphs below;
- $d_{\mathrm{comb}} = w_1 \cdot d_{\mathrm{ng}} + w_2 \cdot d_{\mathrm{MMD}}$, the final combined distance, is a combination of d_{ng} and d_{MMD}. To combine distances, we first rescale them to have the minimum

---

[8] https://github.com/epsilon-machine/missingpy

[9] scikit-learn.org/stable/modules/generated/sklearn.preprocessing.RobustScaler.html

equal to zero and median equal to 1. The final distance is a weighted sum with weights specified in Table 4.A.2.

Optimization of the kernel radius parameter To obtain the most sensitive MMD distance, it is common to optimize the parameter of the kernel used to compute it [207]. We use a Radial Basis Function (RBF) kernel with a single parameter σ_{RBF}, corresponding to the correlation scale. To optimize the value of this parameter, we simulate the entire survey footprint of 10041 tiles at a single parameter set, taken as the median of the prior for each parameter. We compute the MMD distance for an array of 50 values in range $\sigma_{\mathrm{RBF}} \in [0.1, 10]$ for each tile. We then select the value of σ_{RBF} that maximizes the signal-to-noise ratio of the MMD distance distribution from all tiles, defined using robust statistics as:

$$SN[\sigma_{\mathrm{RBF}}] = \mathrm{Md}[\mathrm{d}_{\mathrm{MMD}}(\sigma_{\mathrm{RBF}})]/\mathrm{Mad}[\mathrm{d}_{\mathrm{MMD}}(\sigma_{\mathrm{RBF}})], \tag{8}$$

where Md is the median and Mad is the median absolute deviation. We found the value of $\sigma_{\mathrm{RBF}} = 1$ to be optimal and used it throughout the rest of the analysis.

Modeling of posterior distributions The posterior is resampled in the same way as in the HSC analysis in the previous chapter, with the exception of the fact that we select 1000 out of 10000 ABC points in the posterior, rather than 2000. The resampling is a Gaussian Mixture Model with 20 Gaussians and is performed in a gaussianized space. The resampled posterior is then passed as a prior to the following iteration.

4.A.3 Self-organizing maps visualization

We show the mean i band magnitude and colours in all bands in each cell of the wide and deep SOM in Figures 4.A.1 and 4.A.2.

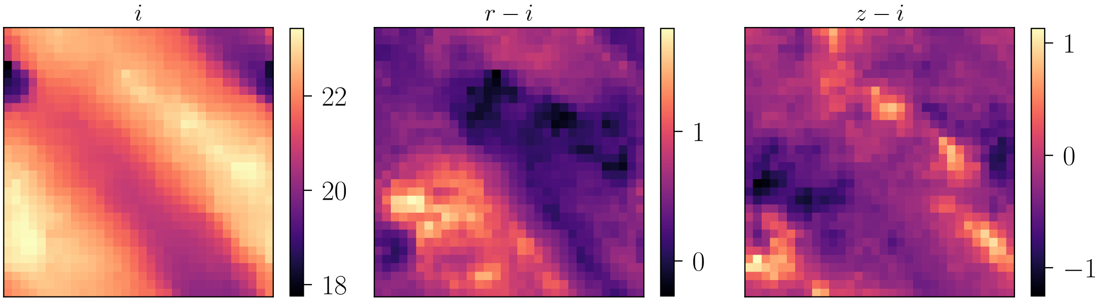


Figure 4.A.1: We visualize the mean of the i band magnitude and of the $r-i$ and $z-i$ colours per wide SOM cell. The map is toroidal so that the left and right edges correspond to the same areas of colour-magnitude space, as do top and bottom.

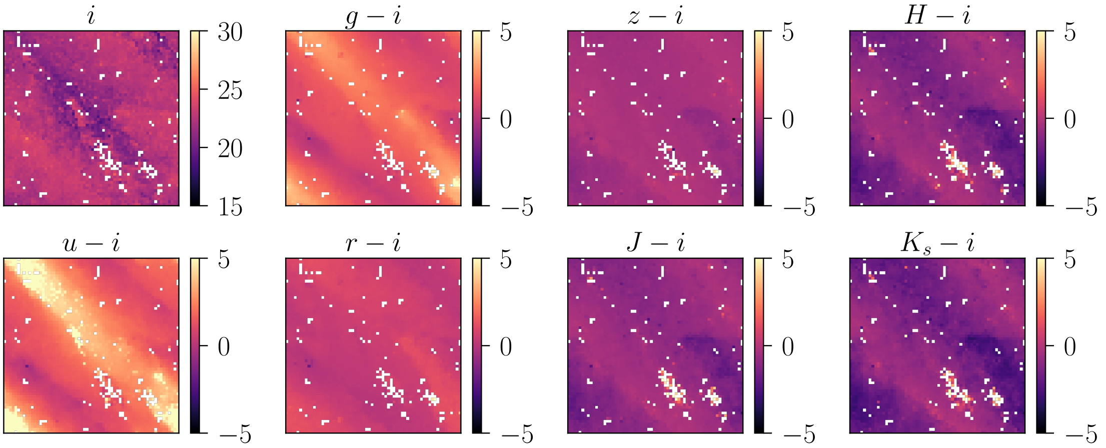


Figure 4.A.2: We visualize the mean of the i band magnitude and of the seven colour combinations in each deep SOM cell. The map is toroidal so that the left and right edges correspond to the same areas of colour-magnitude space, as do top and bottom.

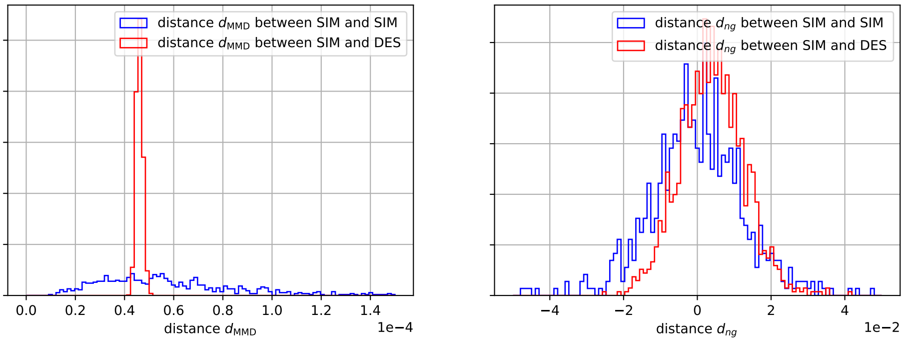


Figure 4.A.3: Distribution of distance measures d_{MMD} and d_{ng} between the simulations themselves (blue) and between the DES data and the simulations (red). Each of the 1000 simulations in the histogram was taken from a different random ABC posterior point. Each distance is an average of 48 tiles (see Appendix 4.A.2 for details on runs and distance calculation).

4.A.4 Unblinding validation

We define the following unblinding criteria and check if they are met:

1.a) First, we verify that the DES data lies somewhere inside the space spanned by the UFIG simulations. In a general way, this can be achieved by comparing distance measures between real and simulated data (DES -vs-SIM), and between SIM-vs-SIM pairs, drawn from the ABC posterior. If the DES-vs-SIM distances lie in a space covered by SIM-vs-SIM distances, then we consider the observed data to lie inside the simulation space according to this distance measure. We calculate the distance measures between simulations themselves $d_{\mathrm{MMD}}(\mathrm{SIM_i}, \mathrm{SIM_j})$ and between the simulations and the DES data $d_{\mathrm{MMD}}(\mathrm{SIM_i}, \mathrm{DES})$. The distances $d_{\mathrm{MMD}}(\mathrm{SIM_i}, \mathrm{SIM_j})$ are calculated at random samples i, j in the posterior. If the SIM-vs-SIM distances were much smaller than the SIM-vs-DES distances, that could be a sign of overfitting. The SIM-vs-DES distances are contained within the SIM-vs-SIM ones, which indicates that our simulations are consistent with the DES data in the distance space. The result of this test for the last iteration is shown in Figure 4.A.3.

1.b) We make sure that there is good agreement between UFIG simulations and real

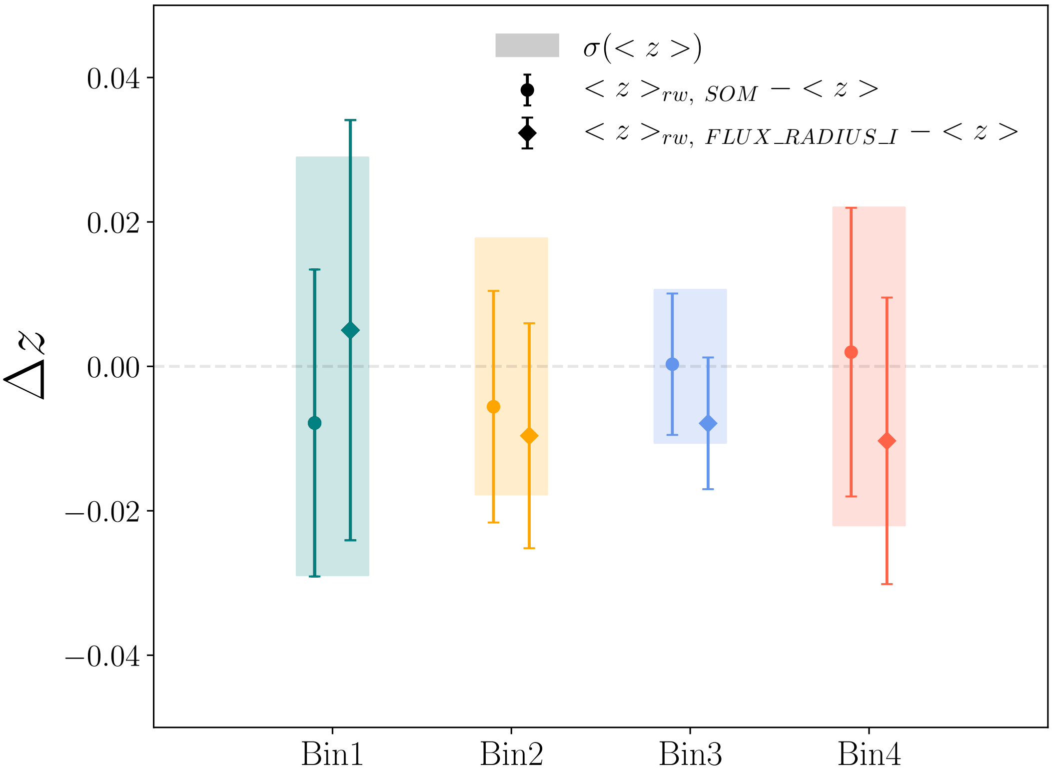


Figure 4.A.4: We display the shift in mean redshift obtained by two different reweightings: the dot indicates the Δz obtained by reweighting riz magnitudes and colours using the SOM and the diamond making i band `FLUX_RADIUS` match between simulations and real images in each tomographic bin. The shaded regions and error bars show the standard deviation of the mean redshift in each of the cases.

images in terms of photometric quantities. We show a tomographic comparison of magnitudes, colours and radii in Figure 4.5, as well as global 2D colour-magnitude distributions in Figure 4.6. In Figure 4.A.5 we display an extended comparison, which includes more global distributions of properties in all $grizY$ bands. We note that the real images lie in the span of the UFIG simulated images for most of the properties. As we already observed in Figure 4.5, the colours involving the g band show some discrepancies similarly to the `FLUX_RADIUS` in all bands, which differs in the tails. We mostly focus on r, i, z magnitudes and $r-i$ and $i-z$ colours and notice that the distributions of real data lie in the space of the simulations, with some small remaining differences. We use 10

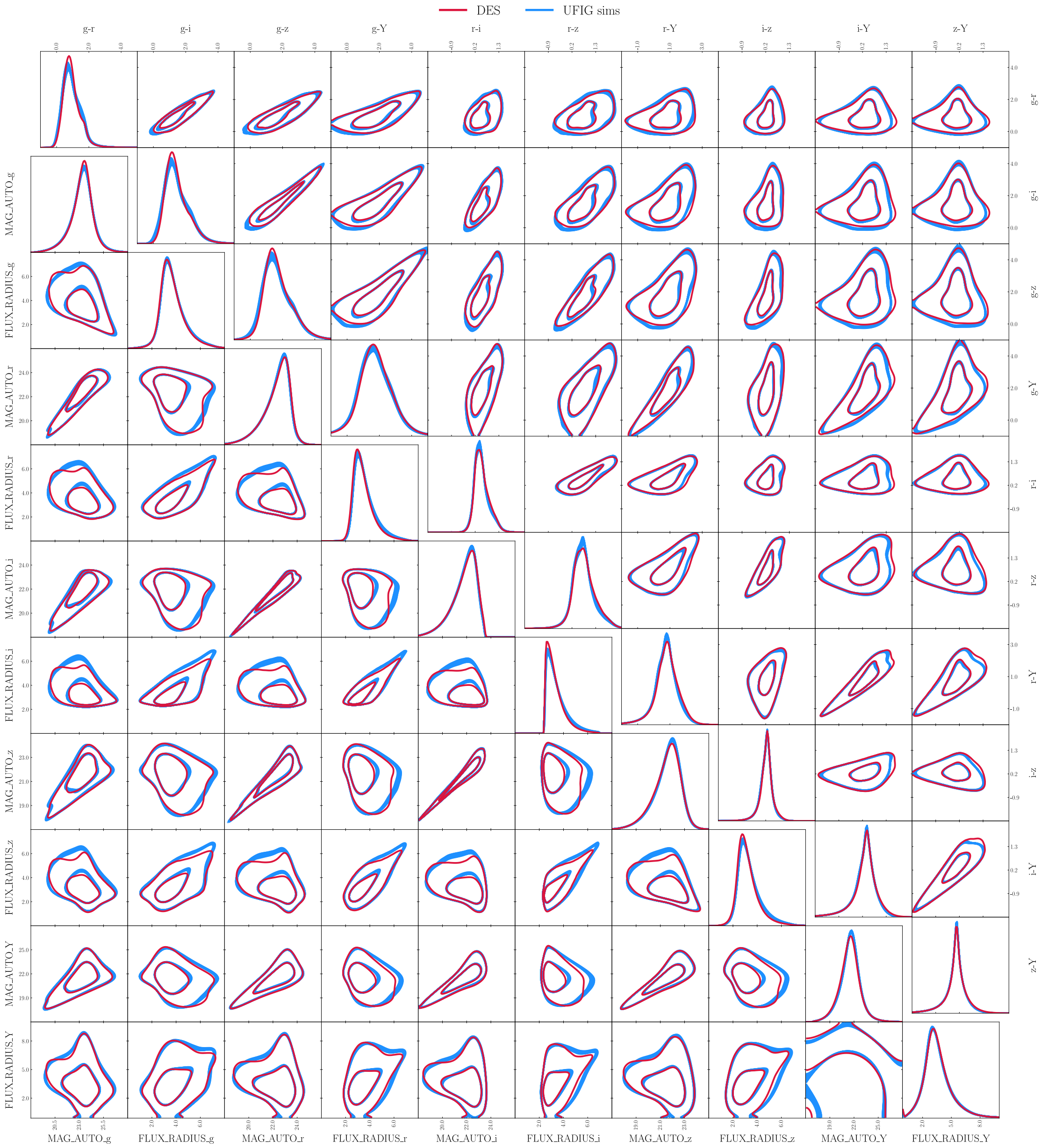


Figure 4.A.5: Corner plot comparing the 68% and 95% level contours of sizes, magnitudes and colour-colour distributions of the target sample from real data (red solid line) and 40 UFIG simulations (blue solid lines).

million galaxies from the target sample in Figure 4.A.5.

1.c) Since our galaxy population model relies on some simplifying assumptions, we expect a small residual dependence of mean redshift on the difference between real images and simulations. Indeed, we notice some of these differences in magnitudes, sizes, and colours, with the most pronounced difference in the $g-r$ colour. We test the significance of this differences in terms of their impact on the final $n(z)$. We do this by reweighting the distributions of photometric properties of simulations to better agree with real images and verify the impact of the reweighting on mean redshift in each tomographic bin. We achieve this in two different ways for magnitudes and galaxy sizes. In the magnitude case, we make use of the wide SOM, since the galaxies from UFIG simulations have been already assigned to it for tomographic bin assignment. By weighting each galaxy with the ratio between the occupation of the wide cell it belongs to in real data and in the UFIG simulation, we obtain a weighting scheme that makes simulations perfectly match real data in riz magnitudes and colours. We can then verify the shift in mean redshift obtained by this reweighting in each bin. Notice that each bin will contain a fraction of the wide SOM cells. In the galaxy size case, we perform a simple reweighting based on the 1D `FLUX_RADIUS` histogram in each band. We require the shift in redshift in each bin to be within the uncertainty in all cases, condition which is met and displayed in Figure 4.A.4 for the SOM reweighting and the `FLUX_RADIUS` leading to the largest shift (i band). The shift is insignificant, even though there are differences in the shapes of the reweighted $n(z)$s. As expected, the SOM reweighting slightly tightens the constraints on mean redshift since we expect the shape of the $n(z)$ to be strongly driven by the magnitude-colour distributions.

2.a) The difference in mean redshift when changing the redshift sample is shown in Figure 4.7 and is of comparable size to [112] (and below 0.015 in each bin).

2.b) The fraction of wide galaxies coming from deep cells without redshift is always under 2%, and typically 1%.

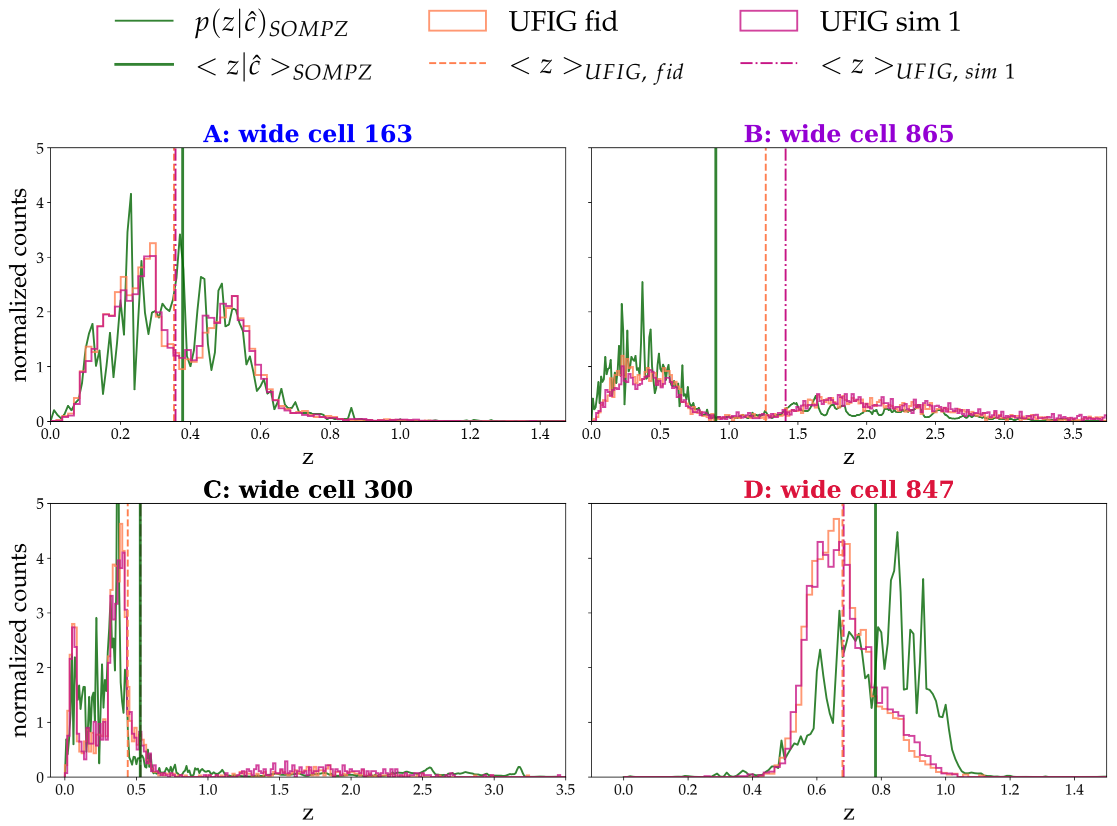


Figure 4.A.6: Redshift distributions in the four wide SOM cells highlighted in Figure 4.17 calibrated by SOMPZ (dark green line) and from the galaxies assigned to the cell by two UFIG simulations (fiducial, solid orange line - simulation 1, purple). The mean redshift in the cell is displayed as a orange dashed line (UFIG fiducial), a dashddotted purple line (UFIG simulation 1) and a green solid line (SOMPZ).

4.A.5 $n(z)$ in selected wide SOM cells from HSC ABC posterior

We present the redshift distributions derived from SOMPZ and MCCL in the wide SOM cells highlighted in Figure 4.17. The MCCL simulations are from the *HSC ABC posterior*. In Figure 4.A.6 we show for each of the four cells the calibrated SOMPZ $n(z)$s (dark green line), and the redshift distributions of UFIG galaxies from the target sample assigned to the same cells in two different UFIG simulations: the fiducial simulation (orange) and a second simulation (sim 1, purple). In Figure 4.A.7, we show the binned

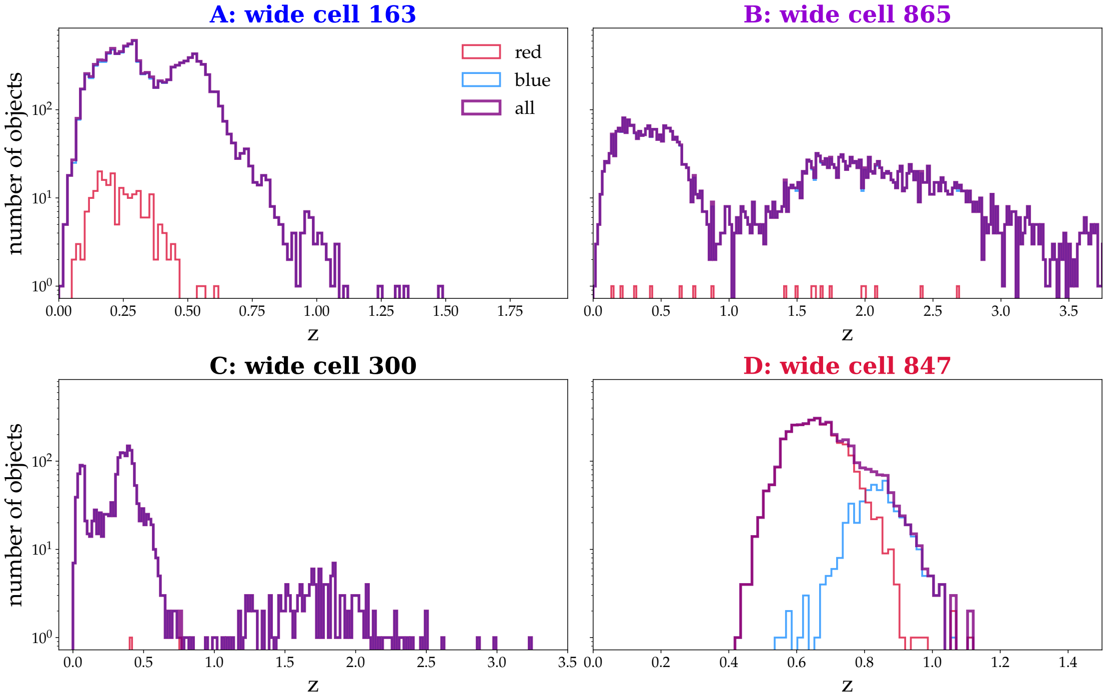


Figure 4.A.7: Redshift distribution in the four wide SOM cells highlighted in Figure 4.17 from the fiducial UFIG simulation (solid purple line). The redshift distributions from the blue and red UFIG luminosity functions is shown in blue and red respectively.

galaxy counts in the UFIG simulations and separate galaxies from the red and blue luminosity functions. The y-axis is logarithmic. When compared to the MCCL run using the ABC posterior calibrated on DES Y3 data, we observe similar trends in cells **A** and **D**. Wide SOM cell **B** shows a bimodal behaviour, with a more pronounced high redshift mode which increases the mean of the redshift distribution (high Δ_z between SOMPZ and MCCL). In cell **C** we see a difference between the fiducial simulation and simulation 1: while the fiducial simulation is very similar to that in Figure 4.11, simulation 1 assigns some high redshift galaxies to the cell, moving the mean towards the SOMPZ prediction. Finally, we observe some changes in the red-to-blue galaxy ratio: cells **A**, **B** and **C** are dominated by blue galaxies, whereas cell **D** by red galaxies. This was already the case in the run using the DES ABC posterior but the red/blue galaxies separation is now more pronounced.

4.B Contributions

For Chapter 4, I was the main contributor. The project was conducted within the redshift and weak lensing working groups of the Dark Energy Survey collaboration. I designed the target sample, modified and ran both the MCCL and SOMPZ pipelines for the comparison. Furthermore, I produced the plots, drew the results, wrote most of the manuscript and presented and discussed the results with the DES collaboration. Tomasz Kacprzak worked on MCCL developments and ran the ABC inference on DES wide field data. He also contributed to the interpretation of the results and the writing of the manuscript. Justin Myles, Daniel Grün, Alexandra Amon and Jamie McCullough helped me with adapting the SOMPZ pipeline for the comparison, providing the needed data products and writing small sections of the text. Alexandre Refregier and Luca Tortorelli contributed to the interpretation and discussion of the results. Michael Troxel extended the DES data products to make this project possible. Will Hartley provided useful insights into SED fitting and deep fields. Alex Alarcon, Spencer Everett and Joe DeRose reviewed a preliminary version of the chapter as internal reviewers within the DES collaboration and approved the unblinding once the conditions were met. Finally, the past and present conveners of the redshift and weak lensing working groups (Carles Sanchez, Alex Alarcon, Huan Lin, Alexandra Amon, Judit Prat, Daniel Grün and Ami Choi) provided valuable comments.

CHAPTER 5

CONCLUSIONS

> *I may not have gone where I intended to go, but I think I have ended up where I needed to be.*
>
> — DOUGLAS ADAMS, The Long Dark Tea-Time of the Soul

The ΛCDM cosmological model is successful at describing the distribution of structure in our Universe, as well as other observables. While this is a remarkable achievement, the nature of dark energy and dark matter is still unknown and tensions have arisen between the model parameters constrained by late-time and early-time cosmological probes. Cosmological galaxy surveys are a well-established tool to investigate the nature of dark energy and the possibility of modifications of the theory of gravity, through the combination of cosmic shear, galaxy clustering and galaxy-galaxy lensing. In the near future, a number of cosmological wide-field surveys will yield an unprecedented amount of imaging and spectroscopic data, which will require improved accuracy and flexibility of the theoretical predictions and careful control of systematics in order to be deployed to its full potential. Approximations will need to be reconsidered and effects that were negligible in previous analyses will need to be included or marginalized over.

In this thesis, we addressed some of the challenges posed by large-scale structure analyses, both in terms of theoretical predictions and observational systematics.

In Chapter 2 we extended the `PyCosmo` framework for the solution of the Einstein-Boltzmann system of ordinary differential equations. We showcased the possibility to easily implement new systems of equations thanks to the symbolic Python implementation, that is automatically converted into optimized C/C++ code. We started from a simple model of dark energy with constant equation of state and then introduced the equations for massive neutrinos, which require numerical integrations and enlarge the ODE system considerably. Finally, we added the radiation streaming approximation, which makes the execution speed remarkably faster with negligible loss in accuracy. The strength of this approach is that it reduces the time between the theoretical development of a new theory and its numerical implementation, allowing to test and constrain state-of-the-art models.

We then turned our focus to observations. Cosmic shear measurements are affected by a number of systematic uncertainties, such as nonlinear corrections in the matter power spectrum, intrinsic alignments of galaxies, imperfect modelling of the PSF and other shape measurement biases. One of such systematics is the determination of the redshift distributions of source and lens galaxies from noisy photometry in a limited number of broad-band filters. The redshift distribution appears in the derivation of large-scale structure observables, so that its imperfect calibration can induce biases in the inferred cosmological parameters. We use forward modelling to calibrate the redshift distribution of a galaxy sample: we constrain an empirical galaxy population model with observations and produce realistic image simulations, process them in the same way as the survey data and retrieve the redshift distribution of interest, naturally including sample selection.

In Chapter 3 we used this approach, part of the Monte Carlo Control Loops (MCCL) framework, to simultaneously constrain the galaxy population model parameters and the redshift distribution of a magnitude limited galaxy sample. We derived a posterior for the model parameters via Approximate Bayesian Computation (ABC) using HSC deep field imaging data, complemented with accurate many-band photometric redshifts from COSMOS2020. We obtained realistic simulations of the HSC deep field images, that reproduce the photometric properties of the real data well, including the colour-redshift relation up to $z \approx 4$. Deep fields are limited in angular size, but have a depth comparable to upcoming surveys. For this reason, they constitute an ideal dataset for testing the applicability of our method for Stage IV surveys. We obtained a good agreement of the redshift distributions up to magnitude 24 in the i band and proposed several model updates to further improve our forward model.

In Chapter 4 we performed a blind tomographic comparison between the redshift distributions obtained from MCCL and Self-Organizing Maps p(z) (SOMPZ), a direct redshift calibration methodology employed by state-of-the-art surveys. We applied both methods to the same target sample of galaxies from DES Y3. The methodologies are complementary. MCCL relies on image simulations that include realistic observational and instrumental effects and are calibrated on wide-field data via simulation-based inference. SOMPZ leverages eight-band photometry from the DES deep fields, image injection to measure the survey transfer function, a sample of galaxies with known redshifts and a dimensionality reduction technique to obtain the wide-field redshift distribution probabilistically. We found a good level of agreement between the means of the first two tomographic redshift bins. The two higher redshift bins show a significant mean offset. We investigated the discrepancy and found indications of a difference in the implied colour-redshift relation, particularly for red galaxies.

Forward modelling presents a promising strategy for redshift calibration, as it naturally incorporates otherwise difficult to treat selection effects. Future developments of the MCCL framework include the implementation of realistic clustering of galaxies in the simulations, which can be a source of systematic uncertainty and will allow the study of effects such as blending, the interplay between blending and photometric redshifts and source clustering. Furthermore, we identified limitations in the size model, which could be extended to include size evolution with redshift at fixed absolute magnitudes, more complex morphologies (such as bulges and disks) or redshift evolving Sersic indices. The inclusion of a physically motivated modelling of the spectral energy distributions using stellar population synthesis is desirable, both to enhance the flexibility of the colour evolution with redshift and to compare our galaxy population model constraints to recent measurements from galaxy evolution studies. Finally, emulators have the potential to speed up the simulation process and facilitate testing of model extensions.

5.A Contributions

I wrote Chapter 5 as a summary of the main findings of this thesis and outlook for future work.

CHAPTER 6

CONTRIBUTIONS

The contributions to the different parts of this thesis are shown below, and they are also reported in the Appendix of each Chapter.

- **Chapter 1**
  I wrote the introductory Chapter 1 based on my review of the cited literature.

- **Chapter 2**
  For Chapter 2, I was the main contributor. I translated the equations to the correct formalism and implemented the massive neutrinos and radiation streaming approximation equations in the `PyCosmo` framework. I also contributed to the documentation of the code for the release on `PyPi`, drawing and interpretation of the results and writing the manuscript. The work on massive neutrinos was initiated as part of my Master's thesis, of which this chapter constitutes a major extension.

  Further contributors: Christiane S. Lorenz, Uwe Schmitt, Alexandre Refregier, Janis Fluri, Federica Tarsitano and Lavinia Heisenberg. They contributed in particular to the interpretation of the results and manuscript preparation. Christiane S. Lorenz derived in the appropriate formalism and implemented the equations for the wcdm model and Uwe Schmitt developed the `sympy2c` framework [141] for symbolic manipulation of equations that is used throughout the chapter. Janis Fluri set up the `PyCosmo`Hub. The mentioned contributors are also the co-authors of the published work [31] upon which this chapter is based.

- **Chapter 3**
  For Chapter 3, I was the main contributor. I devised the project, extended the code-base, ran the Approximate Bayesian Computation and the final simulations. I created the figures, interpreted the results and wrote the majority of the manuscript.

  The early stages of the project were carried out by Dominic Grimm during his Master thesis, which I supervised. He contributed mostly to the design of the forward model of HSC deep images and the creation of a first version of the systematic maps. Tomasz Kacprzak also contributed to the project by designing and running the catalog level ABC on COSMOS2015, reworking parts of the code and participating to writing of the manuscript. Silvan Fischbacher helped with the optimization of the code and discussion of the results. Further contributors are Alexandre Refregier and Luca Tortorelli, who contributed mostly to the interpretation of the results and early code developments. The mentioned contributors are also the co-authors of the submitted work (Moser et al, 2024, [32]) upon which this chapter is based.

- **Chapter 4**
  For Chapter 4, I was the main contributor. The project was conducted within the redshift and weak lensing working groups of the Dark Energy Survey collaboration. I designed the target sample, modified and ran both the MCCL and SOMPZ pipelines for the comparison. Furthermore, I produced the plots, drew the results, wrote most of the manuscript and presented and discussed the results with the DES collaboration. Tomasz Kacprzak worked on MCCL developments and ran the ABC inference on DES wide field data. He also contributed to the interpretation of the results and the writing of the manuscript. Justin Myles, Daniel Grün, Alexandra Amon and Jamie McCullough helped me with adapting the SOMPZ pipeline for the comparison, providing the needed data products and writing small sections of the text. Alexandre Refregier and Luca Tortorelli contributed to the interpretation and discussion of the results. Michael Troxel extended the DES data products to make this project possible. Will Hartley

provided useful insights into SED fitting and deep fields. Alex Alarcon, Spencer Everett and Joe DeRose reviewed a preliminary version of the chapter as internal reviewers within the DES collaboration and approved the unblinding once the conditions were met. Finally, the past and present conveners of the redshift and weak lensing working group (Carles Sanchez, Alex Alarcon, Huan Lin, Alexandra Amon, Judit Prat, Daniel Grün and Ami Choi) provided valuable comments.

- **Chapter 5**
  I wrote Chapter 5 as a summary of the main findings of this thesis and outlook for future work.

BIBLIOGRAPHY

[1] J. A. Frieman, M. S. Turner, and D. Huterer, “Dark energy and the accelerating universe,” *Annual Review of Astronomy and Astrophysics*, vol. 46, p. 385–432, Sept. 2008.

[2] G. Hinshaw, D. Larson, E. Komatsu, *et al.*, “Nine-year wilkinson microwave anisotropy probe (wmap) observations: Cosmological parameter results,” *The Astrophysical Journal Supplement Series*, vol. 208, p. 19, sep 2013.

[3] A. G. Riess, A. V. Filippenko, P. Challis, *et al.*, “Observational evidence from supernovae for an accelerating universe and a cosmological constant,” *The Astronomical Journal*, vol. 116, p. 1009, sep 1998.

[4] S. Perlmutter, G. Aldering, G. Goldhaber, *et al.*, “Measurements of Ω and Λ from 42 High-Redshift Supernovae,” ApJ, vol. 517, pp. 565–586, June 1999.

[5] P. Astier and R. Pain, “Observational evidence of the accelerated expansion of the universe,” *Comptes Rendus Physique*, vol. 13, p. 521–538, July 2012.

[6] D. J. Eisenstein, I. Zehavi, D. W. Hogg, *et al.*, “Detection of the Baryon Acoustic Peak in the Large-Scale Correlation Function of SDSS Luminous Red Galaxies,” ApJ, vol. 633, pp. 560–574, Nov. 2005.

[7] A. G. Sánchez, R. Scoccimarro, M. Crocce, *et al.*, “The clustering of galaxies in the completed SDSS-III Baryon Oscillation Spectroscopic Survey: Cosmological implications of the configuration-space clustering wedges,” MNRAS, vol. 464, pp. 1640–1658, Jan. 2017.

[8] B. Fields and K. Olive, “Big bang nucleosynthesis,” *Nuclear Physics A*, vol. 777, pp. 208–225, Oct. 2006. Funding Information: We thank Richard Cyburt for many productive collaborations. The work of K.A.O. was partially supported by DOE grant DE-FG02-94ER-40823. The work of B.D.F. was supported by the National Science Foundation under grant AST-0092939.

[9] J. L. Bernal, L. Verde, and A. G. Riess, “The trouble with h0,” *Journal of Cosmology and Astroparticle Physics*, vol. 2016, p. 019–019, Oct. 2016.

[10] A. Leauthaud, S. Saito, S. Hilbert, *et al.*, “Lensing is low: cosmology, galaxy formation or new physics?,” *Monthly Notices of the Royal Astronomical Society*, vol. 467, pp. 3024–3047, 02 2017.

[11] C. Ma and E. Bertschinger, “Cosmological perturbation theory in the synchronous and conformal newtonian gauges,” *Astrophys.J.*, no. 455, pp. 7–25, 1995.

[12] A. Refregier, L. Gamper, A. Amara, and L. Heisenberg, “PyCosmo: An integrated cosmological boltzmann solver,” *Astronomy and Computing*, 2018.

[13] M. CHEVALLIER and D. POLARSKI, “Accelerating universes with scaling dark matter,” *International Journal of Modern Physics D*, vol. 10, p. 213–223, Apr. 2001.

[14] E. V. Linder, “Exploring the expansion history of the universe,” *Phys. Rev. Lett.*, vol. 90, p. 091301, Mar 2003.

[15] J. N. Bahcall and R. Davis, “Solar Neutrinos - a Scientific Puzzle,” *Science*, vol. 191, pp. 264–267, 1976.

[16] M. Lattanzi and M. Gerbino, “Status of neutrino properties and future prospects - Cosmological and astrophysical constraints,” *Front. in Phys.*, vol. 5, p. 70, 2018.

[17] D. Blas, J. Lesgourgues, and T. Tram, “The cosmic linear anisotropy solving system (class). part ii: Approximation schemes,” *Journal of Cosmology and Astroparticle Physics*, vol. 2011, p. 034–034, Jul 2011.

[18] D. E. S. Collaboration:, T. Abbott, F. B. Abdalla, *et al.*, “The Dark Energy Survey: more than dark energy – an overview,” *Monthly Notices of the Royal Astronomical Society*, vol. 460, pp. 1270–1299, 03 2016.

[19] J. T. A. de Jong, G. A. Verdoes Kleijn, K. H. Kuijken, and E. A. Valentijn, “The kilo-degree survey,” *Experimental Astronomy*, vol. 35, p. 25–44, Aug 2012.

[20] H. Aihara, N. Arimoto, R. Armstrong, *et al.*, “The Hyper Suprime-Cam SSP Survey: Overview and survey design,” *Publications of the Astronomical Society of Japan*, vol. 70, 09 2017. S4.

[21] Željko Ivezić, S. M. Kahn, J. A. Tyson, *et al.*, “Lsst: From science drivers to reference design and anticipated data products,” *The Astrophysical Journal*, vol. 873, p. 111, mar 2019.

[22] R. Laureijs, J. Amiaux, S. Arduini, *et al.*, “Euclid definition study report,” 2011.

[23] Euclid Collaboration, G. Desprez, S. Paltani, *et al.*, “Euclid preparation. X. The Euclid photometric-redshift challenge,” A&A, vol. 644, p. A31, Dec. 2020.

[24] R. Akeson, L. Armus, E. Bachelet, *et al.*, “The wide field infrared survey telescope: 100 hubbles for the 2020s,” 2019.

[25] A. Amara and A. Réfrégier, “Systematic bias in cosmic shear: extending the Fisher matrix,” MNRAS, vol. 391, pp. 228–236, Nov. 2008.

[26] R. Reischke, “Propagating photo-z uncertainties: A functional derivative approach,” MNRAS, Dec. 2023.

[27] A. Refregier and A. Amara, “A way forward for cosmic shear: Monte-carlo control loops,” *Physics of the Dark Universe*, vol. 3, pp. 1 – 3, 2014.

[28] H. Aihara, Y. AlSayyad, M. Ando, *et al.*, “Third data release of the Hyper Suprime-Cam Subaru Strategic Program,” *Publications of the Astronomical Society of Japan*, vol. 74, pp. 247–272, 02 2022.

[29] J. R. Weaver, O. B. Kauffmann, O. Ilbert, *et al.*, “COSMOS2020: A panchromatic view of the universe to z ∼ 10 from two complementary catalogs,” *The Astrophysical Journal Supplement Series*, vol. 258, p. 11, jan 2022.

[30] A. Drlica-Wagner, I. Sevilla-Noarbe, E. S. Rykoff, *et al.*, “Dark energy survey year 1 results: The photometric data set for cosmology,” *The Astrophysical Journal Supplement Series*, vol. 235, p. 33, Apr 2018.

[31] B. Moser, C. S. Lorenz, U. Schmitt, *et al.*, “Symbolic implementation of extensions of the PyCosmo Boltzmann solver,” *Astronomy and Computing*, vol. 40, p. 100603, July 2022.

[32] B. Moser, T. Kacprzak, S. Fischbacher, *et al.*, “Simulation-based inference of deep fields: galaxy population model and redshift distributions,” *arXiv e-prints*, p. arXiv:2401.06846, Jan. 2024.

[33] D. Baumann, *Cosmology*. Cambridge University Press, 2022.

[34] A. Liddle, *An introduction to modern cosmology; 2nd ed.* Chichester: Wiley, 2003.

[35] S. Dodelson and F. Schmidt, *Modern Cosmology*. 2020.

[36] M. Bartelmann, “The dark universe,” *Rev. Mod. Phys.*, vol. 82, pp. 331–382, Feb 2010.

[37] D. Huterer, *The Hubble Law and Geometry of Space*, p. 14–40. Cambridge University Press, 2023.

[38] U. Seljak, “Lectures on dark matter,” 2000.

[39] S. Weinberg, *Cosmology*. 2008.

[40] A. Friedmann, “Über die Krümmung des Raumes,” *Zeitschrift fur Physik*, vol. 10, pp. 377–386, January 1922.

[41] G. Lemaître, “Un Univers homogène de masse constante et de rayon croissant rendant compte de la vitesse radiale des nébuleuses extra-galactiques,” *Annales de la Société Scientifique de Bruxelles*, vol. 47, pp. 49–59, January 1927.

[42] H. P. Robertson, “Kinematics and World-Structure,” *The Astrophysical Journal*, vol. 82, p. 284, November 1935.

[43] A. G. Walker, “On the formal comparison of milne’s kinematical system with the systems of general relativity,” *Monthly Notices of the Royal Astronomical Society*, vol. 95, pp. 263–269, January 1935.

[44] F. Zwicky, “Die Rotverschiebung von extragalaktischen Nebeln,” *Helvetica Physica Acta*, vol. 6, pp. 110–127, Jan. 1933.

[45] E. Hubble, “A Relation between Distance and Radial Velocity among Extra-Galactic Nebulae,” *Proceedings of the National Academy of Science*, vol. 15, pp. 168–173, Mar. 1929.

[46] E. Lifshitz, “Republication of: On the gravitational stability of the expanding universe,” *General Relativity and Gravitation*, vol. 49, 01 2017.

[47] E. Bertschinger, “Cosmological perturbation theory and structure formation,” 2000.

[48] R. A. Sunyaev and J. Chluba, “Signals from the epoch of cosmological recombination - karl schwarzschild award lecture 2008,” *Astronomische Nachrichten*, vol. 330, pp. 657–674, aug 2009.

[49] J. Chluba and R. M. Thomas, “Towards a complete treatment of the cosmological recombination problem,” *Monthly Notices of the Royal Astronomical Society*, pp. no–no, dec 2010.

[50] J. Chluba, G. Vasil, and L. Dursi, “Recombinations to the rydberg states of hydrogen and their effect during the cosmological recombination epoch,” *Monthly Notices of the Royal Astronomical Society*, vol. 407, no. 1, pp. 599–612, 2010.

[51] A. H. Guth, “Inflationary universe: A possible solution to the horizon and flatness problems,” Phys. Rev. D, vol. 23, pp. 347–356, Jan. 1981.

[52] D. Langlois, “Inflation and cosmological perturbations,” in *Lectures on Cosmology*, pp. 1–57, Springer Berlin Heidelberg, 2010.

[53] K. Dolag, S. Borgani, S. Schindler, *et al.*, "Simulation techniques for cosmological simulations," *Space Science Reviews*, vol. 134, pp. 229–268, feb 2008.

[54] R. E. Smith, J. A. Peacock, A. Jenkins, *et al.*, "Stable clustering, the halo model and non-linear cosmological power spectra," *Monthly Notices of the Royal Astronomical Society*, vol. 341, pp. 1311–1332, 06 2003.

[55] R. Takahashi, M. Sato, T. Nishimichi, *et al.*, "Revising the halofit model for the nonlinear matter power spectrum," *The Astrophysical Journal*, vol. 761, p. 152, dec 2012.

[56] S. Bird, M. Viel, and M. G. Haehnelt, "Massive neutrinos and the non-linear matter power spectrum," *Monthly Notices of the Royal Astronomical Society*, vol. 420, pp. 2551–2561, 02 2012.

[57] W. H. Press and P. Schechter, "Formation of Galaxies and Clusters of Galaxies by Self-Similar Gravitational Condensation," ApJ, vol. 187, pp. 425–438, Feb. 1974.

[58] A. R. ZENTNER, "THE EXCURSION SET THEORY OF HALO MASS FUNCTIONS, HALO CLUSTERING, AND HALO GROWTH," *International Journal of Modern Physics D*, vol. 16, pp. 763–815, may 2007.

[59] A. COORAY and R. SHETH, "Halo models of large scale structure," *Physics Reports*, vol. 372, pp. 1–129, dec 2002.

[60] M. Asgari, A. J. Mead, and C. Heymans, "The halo model for cosmology: a pedagogical review," 2023.

[61] H. Mo, F. C. van den Bosch, and S. White, *Galaxy Formation and Evolution*. 2010.

[62] R. A. Crain and F. van de Voort, "Hydrodynamical simulations of the galaxy population: Enduring successes and outstanding challenges," *Annual Review of Astronomy and Astrophysics*, vol. 61, no. 1, pp. 473–515, 2023.

[63] P. D. Mitchell, C. G. Lacey, C. D. P. Lagos, *et al.*, "Comparing galaxy formation in semi-analytic models and hydrodynamical simulations," *Monthly Notices of the Royal Astronomical Society*, vol. 474, pp. 492–521, 10 2017.

[64] G. Moriondo, C. Giovanardi, and L. K. Hunt, “Near-infrared surface photometry of early-type spiral galaxies. I. Bulge and disk decomposition,” A&AS, vol. 130, pp. 81–108, May 1998.

[65] D. W. Hogg, I. K. Baldry, M. R. Blanton, and D. J. Eisenstein, “The K correction,” 10 2002.

[66] I. Strateva, Ž. Ivezić, G. R. Knapp, *et al.*, “Color Separation of Galaxy Types in the Sloan Digital Sky Survey Imaging Data,” AJ, vol. 122, pp. 1861–1874, Oct. 2001.

[67] R. Johnston, “Shedding Light on the Galaxy Luminosity Function,” *Astron. Astrophys. Rev.*, vol. 19, p. 41, 2011.

[68] V. Desjacques, D. Jeong, and F. Schmidt, “Large-scale galaxy bias,” Phys. Rep., vol. 733, pp. 1–193, Feb. 2018.

[69] D. N. Limber, “The Analysis of Counts of the Extragalactic Nebulae in Terms of a Fluctuating Density Field. II.,” ApJ, vol. 119, p. 655, May 1954.

[70] T. Treu, “Strong Lensing by Galaxies,” ARA&A, vol. 48, pp. 87–125, Sept. 2010.

[71] A. J. Shajib, G. Vernardos, T. E. Collett, *et al.*, “Strong lensing by galaxies,” 2022.

[72] M. Bartelmann and P. Schneider, “Weak gravitational lensing,” Phys. Rep., vol. 340, pp. 291–472, Jan. 2001.

[73] A. Refregier, “Weak gravitational lensing by large-scale structure,” *Annual Review of Astronomy and Astrophysics*, vol. 41, pp. 645–668, sep 2003.

[74] P. Schneider, “Weak gravitational lensing,” in *Saas-Fee Advanced Courses*, pp. 269–451, Springer Berlin Heidelberg, 2006.

[75] H. Hoekstra and B. Jain, “Weak Gravitational Lensing and Its Cosmological Applications,” *Annual Review of Nuclear and Particle Science*, vol. 58, pp. 99–123, Nov. 2008.

[76] R. Mandelbaum, “Weak lensing for precision cosmology,” *Annual Review of Astronomy and Astrophysics*, vol. 56, no. 1, pp. 393–433, 2018.

[77] A. Congdon and C. Keeton, *Principles of Gravitational Lensing: Light Deflection as a Probe of Astrophysics and Cosmology*. Springer Praxis Books, Springer International Publishing, 2018.

[78] C. Seitz and P. Schneider, “Steps towards nonlinear cluster inversion through gravitational distortions: Iii. including a redshift distribution of the sources,” 1996.

[79] M. Kilbinger, “Cosmology with cosmic shear observations: a review,” *Reports on Progress in Physics*, vol. 78, p. 086901, July 2015.

[80] C. Heymans, L. Van Waerbeke, D. Bacon, *et al.*, “The Shear Testing Programme – I. Weak lensing analysis of simulated ground-based observations,” *Monthly Notices of the Royal Astronomical Society*, vol. 368, pp. 1323–1339, 04 2006.

[81] C. Heymans, B. Rowe, H. Hoekstra, *et al.*, “The impact of high spatial frequency atmospheric distortions on weak-lensing measurements: Atmospheric distortions,” *Monthly Notices of the Royal Astronomical Society*, pp. no–no, Feb. 2012.

[82] W. Hu, “Power spectrum tomography with weak lensing,” *The Astrophysical Journal*, vol. 522, p. L21–L24, Sept. 1999.

[83] M. Troxel and M. Ishak, “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, Feb. 2015.

[84] M. Cacciato, F. C. Van Den Bosch, S. More, *et al.*, “Galaxy clustering and galaxy–galaxy lensing: a promising union to constrain cosmological parameters,” *Monthly Notices of the Royal Astronomical Society*, vol. 394, pp. 929–946, 03 2009.

[85] Ilbert, O., Arnouts, S., McCracken, H. J., *et al.*, “Accurate photometric redshifts for the cfht legacy survey calibrated using the vimos vlt deep survey,” *A&A*, vol. 457, no. 3, pp. 841–856, 2006.

[86] C. Laigle, H. J. McCracken, O. Ilbert, *et al.*, “The cosmos2015 catalog: Exploring the $1 < z < 6$ universe with half a million galaxies,” *The Astrophysical Journal Supplement Series*, vol. 224, p. 24, Jun 2016.

[87] M. Eriksen, A. Alarcon, E. Gaztanaga, *et al.*, "The PAU Survey: early demonstration of photometric redshift performance in the COSMOS field," MNRAS, vol. 484, pp. 4200–4215, Apr. 2019.

[88] H. Hildebrandt, S. Arnouts, P. Capak, *et al.*, "PHAT: PHoto-z Accuracy Testing," A&A, vol. 523, p. A31, Nov. 2010.

[89] C. Sánchez, M. Carrasco Kind, H. Lin, *et al.*, "Photometric redshift analysis in the Dark Energy Survey Science Verification data," *Monthly Notices of the Royal Astronomical Society*, vol. 445, pp. 1482–1506, 10 2014.

[90] M. Tanaka, J. Coupon, B.-C. Hsieh, *et al.*, "Photometric redshifts for Hyper Suprime-Cam Subaru Strategic Program Data Release 1," *Publications of the Astronomical Society of Japan*, vol. 70, p. S9, 10 2017.

[91] M. Salvato, O. Ilbert, and B. Hoyle, "The many flavours of photometric redshifts," 2018.

[92] S. J. Schmidt, A. I. Malz, J. Y. H. Soo, *et al.*, "Evaluation of probabilistic photometric redshift estimation approaches for The Rubin Observatory Legacy Survey of Space and Time (LSST)," *Monthly Notices of the Royal Astronomical Society*, vol. 499, pp. 1587–1606, 09 2020.

[93] M. Brescia, S. Cavuoti, O. Razim, *et al.*, "Photometric redshifts with machine learning, lights and shadows on a complex data science use case," *Frontiers in Astronomy and Space Sciences*, vol. 8, 2021.

[94] S. Arnouts, S. Cristiani, L. Moscardini, *et al.*, "Measuring and modelling the redshift evolution of clustering: the Hubble Deep Field North," MNRAS, vol. 310, pp. 540–556, Dec. 1999.

[95] G. B. Brammer, P. G. van Dokkum, and P. Coppi, "EAZY: A Fast, Public Photometric Redshift Code," ApJ, vol. 686, pp. 1503–1513, Oct. 2008.

[96] N. Benítez, "Bayesian Photometric Redshift Estimation," ApJ, vol. 536, pp. 571–583, June 2000.

[97] R. Feldmann, C. M. Carollo, C. Porciani, *et al.*, "The Zurich Extragalactic Bayesian Redshift Analyzer and its first application: COSMOS," *Monthly Notices of the Royal Astronomical Society*, vol. 372, pp. 565–577, 09 2006.

[98] M. Tanaka, "PHOTOMETRIC REDSHIFT WITH BAYESIAN PRIORS ON PHYSICAL PROPERTIES OF GALAXIES," *The Astrophysical Journal*, vol. 801, p. 20, feb 2015.

[99] A. A. Collister and O. Lahav, "ANNz: Estimating Photometric Redshifts Using Artificial Neural Networks," PASP, vol. 116, pp. 345–351, Apr. 2004.

[100] I. Sadeh, F. B. Abdalla, and O. Lahav, "ANNz2: Photometric redshift and probability distribution function estimation using machine learning," *Publications of the Astronomical Society of the Pacific*, vol. 128, p. 104502, aug 2016.

[101] J. D. Vicente, E. Sá nchez, and I. Sevilla-Noarbe, "DNF – galaxy photometric redshift by directional neighbourhood fitting," *Monthly Notices of the Royal Astronomical Society*, vol. 459, pp. 3078–3088, apr 2016.

[102] Z. Gomes, M. J. Jarvis, I. A. Almosallam, and S. J. Roberts, "Improving photometric redshift estimation using GPz: size information, post processing, and improved photometry," *Monthly Notices of the Royal Astronomical Society*, vol. 475, pp. 331–342, 12 2017.

[103] M. C. Kind and R. J. Brunner, "TPZ: photometric redshift PDFs and ancillary information by using prediction trees and random forests," *Monthly Notices of the Royal Astronomical Society*, vol. 432, pp. 1483–1501, may 2013.

[104] S. Cavuoti, M. Brescia, G. Longo, and A. Mercurio, "Photometric redshifts with the quasi newton algorithm (MLPQNA) results in the PHAT1 contest," *Astronomy & Astrophysics*, vol. 546, p. A13, sep 2012.

[105] S. Cavuoti, V. Amaro, M. Brescia, *et al.*, "METAPHOR: a machine-learning-based method for the probability density estimation of photometric redshifts," *Monthly Notices of the Royal Astronomical Society*, vol. 465, pp. 1959–1973, 11 2016.

[106] B. C. Hsieh and H. K. C. Yee, “Estimating luminosities and stellar masses of galaxies photometrically without determining redshifts,” *The Astrophysical Journal*, vol. 792, p. 102, aug 2014.

[107] M. Lima, C. E. Cunha, H. Oyaizu, *et al.*, “Estimating the redshift distribution of photometric galaxy samples,” MNRAS, vol. 390, pp. 118–130, Oct. 2008.

[108] T. Kohonen, “Self-organized formation of topologically correct feature maps,” *Biological Cybernetics*, vol. 43, pp. 59–69, Jan. 1982.

[109] D. Masters, P. Capak, D. Stern, *et al.*, “Mapping the galaxy color–redshift relation: Optimal photometric redshift calibration strategies for cosmology surveys,” *The Astrophysical Journal*, vol. 813, p. 53, oct 2015.

[110] R. Buchs, C. Davis, D. Gruen, *et al.*, “Phenotypic redshifts with self-organizing maps: A novel method to characterize redshift distributions of source galaxies for weak lensing,” *Monthly Notices of the Royal Astronomical Society*, vol. 489, pp. 820–841, aug 2019.

[111] A. H. Wright, H. Hildebrandt, J. L. van den Busch, *et al.*, “KiDS+VIKING-450: Improved cosmological parameter constraints from redshift calibration with self-organising maps,” *Astronomy & Astrophysics*, vol. 640, p. L14, aug 2020.

[112] J. Myles, A. Alarcon, A. Amon, *et al.*, “Dark Energy Survey Year 3 results: redshift calibration of the weak lensing source galaxies,” *Monthly Notices of the Royal Astronomical Society*, vol. 505, pp. 4249–4277, 05 2021.

[113] J. L. van den Busch, A. H. Wright, H. Hildebrandt, *et al.*, “KiDS-1000: Cosmic shear with enhanced redshift calibration,” *Astronomy & Astrophysics*, vol. 664, p. A170, aug 2022.

[114] J. A. Newman, “Calibrating Redshift Distributions beyond Spectroscopic Limits with Cross-Correlations,” ApJ, vol. 684, pp. 88–101, Sept. 2008.

[115] M. McQuinn and M. White, “On using angular cross-correlations to determine source redshift distributions,” MNRAS, vol. 433, pp. 2857–2883, Aug. 2013.

[116] B. Ménard, R. Scranton, S. Schmidt, *et al.*, “Clustering-based redshift estimation: method and application to data,” 2014.

[117] C. B. Morrison, H. Hildebrandt, S. J. Schmidt, *et al.*, “the-wizz: clustering redshift estimation for everyone,” *Monthly Notices of the Royal Astronomical Society*, vol. 467, pp. 3576–3589, 02 2017.

[118] B. Hoyle and M. M. Rau, “Self-consistent redshift estimation using correlation functions without a spectroscopic reference sample,” *Monthly Notices of the Royal Astronomical Society*, vol. 485, pp. 3642–3660, 02 2019.

[119] M. Gatti, G. Giannini, G. M. Bernstein, *et al.*, “Dark energy survey year 3 results: clustering redshifts – calibration of the weak lensing source redshift distributions with redmagic and BOSS/eBOSS,” *Monthly Notices of the Royal Astronomical Society*, vol. 510, pp. 1223–1247, nov 2021.

[120] J. Herbel, T. Kacprzak, A. Amara, *et al.*, “The redshift distribution of cosmological samples: a forward modeling approach,” J. Cosmology Astropart. Phys., vol. 2017, p. 035, Aug. 2017.

[121] J. Alsing, H. Peiris, D. Mortlock, *et al.*, “Forward modeling of galaxy populations for cosmological redshift distribution inference,” *The Astrophysical Journal Supplement Series*, vol. 264, p. 29, Jan. 2023.

[122] B. Leistedt, D. J. Mortlock, and H. V. Peiris, “Hierarchical Bayesian inference of galaxy redshift distributions from photometric surveys,” MNRAS, vol. 460, pp. 4258–4267, Aug. 2016.

[123] B. Leistedt, D. W. Hogg, R. H. Wechsler, and J. DeRose, “Hierarchical Modeling and Statistical Calibration for Photometric Redshifts,” ApJ, vol. 881, p. 80, Aug. 2019.

[124] C. Sánchez and G. M. Bernstein, “Redshift inference from the combination of galaxy colours and clustering in a hierarchical Bayesian model,” *Monthly Notices of the Royal Astronomical Society*, vol. 483, pp. 2801–2813, 11 2018.

[125] A. Alarcon, C. Sánchez, G. M. Bernstein, and E. Gaztañaga, "Redshift inference from the combination of galaxy colours and clustering in a hierarchical Bayesian model - Application to realistic N-body simulations," MNRAS, vol. 498, pp. 2614–2631, Oct. 2020.

[126] B. Leistedt, J. Alsing, H. Peiris, *et al.*, "Hierarchical bayesian inference of photometric redshifts with stellar population synthesis models," *The Astrophysical Journal Supplement Series*, vol. 264, p. 23, Jan. 2023.

[127] T. Kacprzak, J. Herbel, A. Nicola, *et al.*, "Monte carlo control loops for cosmic shear cosmology with des year 1 data," *Physical Review D*, vol. 101, Apr 2020.

[128] S. Dodelson, *Modern Cosmology.* Academic Press, Elsevier Science, 2003.

[129] S. Nadkarni-Ghosh and A. Refregier, "The einstein–boltzmann equations revisited," *Monthly Notices of the Royal Astronomical Society*, vol. 471, p. 2391–2430, Jul 2017.

[130] E. Bertschinger, "COSMICS: Cosmological Initial Conditions and Microwave Anisotropy Codes," *arXiv e-prints*, pp. astro–ph/9506070, June 1995.

[131] U. Seljak and M. Zaldarriaga, "A line-of-sight integration approach to cosmic microwave background anisotropies," *The Astrophysical Journal*, vol. 469, p. 437, Oct 1996.

[132] M. Doran, "Cmbeasy: an object oriented code for the cosmic microwave background," *Journal of Cosmology and Astroparticle Physics*, vol. 2005, p. 011–011, Oct 2005.

[133] A. Lewis, A. Challinor, and A. Lasenby, "Efficient computation of cosmic microwave background anisotropies in closed friedmann-robertson-walker models," *The Astrophysical Journal*, 2000.

[134] J. Lesgourgues, "The cosmic linear anisotropy solving system (class) i: Overview," tech. rep., CERN, 2011.

[135] M. Zumalacárregui, E. Bellini, I. Sawicki, *et al.*, "hi_class: Horndeski in the Cosmic Linear Anisotropy Solving System," *JCAP*, vol. 08, p. 019, 2017.

[136] B. Hu, M. Raveri, N. Frusciante, and A. Silvestri, “Effective Field Theory of Cosmic Acceleration: an implementation in CAMB,” *Phys. Rev. D*, vol. 89, no. 10, p. 103530, 2014.

[137] J. C. Hill, E. McDonough, M. W. Toomey, and S. Alexander, “Early dark energy does not restore cosmological concordance,” *Phys. Rev. D*, vol. 102, no. 4, p. 043507, 2020.

[138] E. D. Dio, F. Montanari, J. Lesgourgues, and R. Durrer, “The classgal code for relativistic cosmological large scale structure,” *Journal of Cosmology and Astroparticle Physics*, vol. 2013, p. 044–044, Nov 2013.

[139] F. Tarsitano, U. Schmitt, A. Refregier, *et al.*, “Predicting cosmological observables with pycosmo,” *Astronomy and Computing*, vol. 36, p. 100484, 2021.

[140] N. E. Chisari *et al.*, “Core Cosmology Library: Precision Cosmological Predictions for LSST,” *Astrophys. J. Suppl.*, vol. 242, no. 1, p. 2, 2019.

[141] U. Schmitt, B. Moser, C. S. Lorenz, and A. Réfrégier, “sympy2c: From symbolic expressions to fast C/C++ functions and ODE solvers in Python,” *Astronomy and Computing*, vol. 42, p. 100666, Jan. 2023.

[142] Petzold, “Automatic selection of methods for solving stiff and nonstiff systems of ordinary differential equations,” *SIAM Journal on Scientific and Statistical Computing*, vol. 4, no. 1, pp. 136–148, 1983.

[143] D. J. Eisenstein and W. Hu, “Baryonic features in the matter transfer function,” *The Astrophysical Journal*, vol. 496, p. 605–614, Apr 1998.

[144] J. A. Peacock, “The evolution of galaxy clustering,” *Monthly Notices of the Royal Astronomical Society*, vol. 284, p. 885–898, Feb 1997.

[145] A. Refregier, A. Amara, T. D. Kitching, and A. Rassat, “icosmo: an interactive cosmology package,” *Astronomy & Astrophysics*, vol. 528, p. A33, Feb 2011.

[146] J. Lesgourgues, “The cosmic linear anisotropy solving system (class) iii: Comparision with camb for lambdacdm,” 2011.

[147] J. Lesgourgues and T. Tram, “The cosmic linear anisotropy solving system (class) iv: efficient implementation of non-cold relics,” *Journal of Cosmology and Astroparticle Physics*, vol. 2011, p. 032–032, Sep 2011.

[148] G. Ballesteros and J. Lesgourgues, “Dark energy with non-adiabatic sound speed: initial conditions and detectability,” *JCAP*, vol. 10, p. 014, 2010.

[149] A. Tripathi, A. Sangwan, and H. K. Jassal, “Dark energy equation of state parameter and its evolution at low redshift,” *JCAP*, vol. 06, p. 012, 2017.

[150] D. M. Scolnic *et al.*, “The Complete Light-curve Sample of Spectroscopically Confirmed SNe Ia from Pan-STARRS1 and Cosmological Constraints from the Combined Pantheon Sample,” *Astrophys. J.*, vol. 859, no. 2, p. 101, 2018.

[151] N. Aghanim *et al.*, “Planck 2018 results. VI. Cosmological parameters,” *Astron. Astrophys.*, vol. 641, p. A6, 2020.

[152] T. M. C. Abbott *et al.*, “Dark Energy Survey Year 3 Results: Cosmological Constraints from Galaxy Clustering and Weak Lensing,” 5 2021.

[153] Y. Fukuda *et al.*, “Evidence for oscillation of atmospheric neutrinos,” *Phys. Rev. Lett.*, vol. 81, pp. 1562–1567, Aug 1998.

[154] Q. R. Ahmad *et al.*, “Direct evidence for neutrino flavor transformation from neutral-current interactions in the sudbury neutrino observatory,” *Phys. Rev. Lett.*, vol. 89, p. 011301, Jun 2002.

[155] K. Eguchi *et al.*, “First results from kamland: Evidence for reactor antineutrino disappearance,” *Phys. Rev. Lett.*, vol. 90, p. 021802, Jan 2003.

[156] P. F. de Salas, D. V. Forero, S. Gariazzo, *et al.*, “2020 global reassessment of the neutrino oscillation picture,” *Journal of High Energy Physics*, vol. 2021, Feb 2021.

[157] J. Lesgourgues and S. Pastor, “Massive neutrinos and cosmology,” *Phys. Rept.*, vol. 429, pp. 307–379, 2006.

[158] J. Lesgourgues, G. Mangano, G. Miele, and S. Pastor, *Neutrino Cosmology.* Cambridge University Press, 2013.

[159] K. Ichikawa, M. Fukugita, and M. Kawasaki, "Constraining neutrino masses by CMB experiments alone," *Phys. Rev. D*, vol. 71, p. 043001, 2005.

[160] J. Yoo, A. L. Fitzpatrick, and M. Zaldarriaga, "New perspective on galaxy clustering as a cosmological probe: General relativistic effects," *Physical Review D*, vol. 80, Oct 2009.

[161] J. Yoo, "General relativistic description of the observed galaxy power spectrum: Do we understand what we measure?," *Physical Review D*, vol. 82, Oct 2010.

[162] C. Bonvin and R. Durrer, "What galaxy surveys really measure," *Physical Review D*, vol. 84, Sep 2011.

[163] A. Challinor and A. Lewis, "Linear power spectrum of observed source number counts," *Physical Review D*, vol. 84, Aug 2011.

[164] M. Doran, "Speeding up cosmological Boltzmann codes," *JCAP*, vol. 06, p. 011, 2005.

[165] V. Poulin, T. L. Smith, T. Karwal, and M. Kamionkowski, "Early Dark Energy Can Resolve The Hubble Tension," *Phys. Rev. Lett.*, vol. 122, no. 22, p. 221301, 2019.

[166] C. Pitrou, T. S. Pereira, and J. Lesgourgues, "Optimal Boltzmann hierarchies with nonvanishing spatial curvature," *Phys. Rev. D*, vol. 102, no. 2, p. 023511, 2020.

[167] L. Hui, "Wave Dark Matter," 1 2021.

[168] F.-Y. Cyr-Racine and K. Sigurdson, "Photons and baryons before atoms: Improving the tight-coupling approximation," *Physical Review D*, vol. 83, may 2011.

[169] M. Bucher, K. Moodley, and N. Turok, "General primordial cosmic perturbation," *Physical Review D*, vol. 62, sep 2000.

[170] T. Abbott, M. Aguena, A. Alarcon, *et al.*, "Dark energy survey year 3 results: Cosmological constraints from galaxy clustering and weak lensing," *Physical Review D*, vol. 105, Jan. 2022.

[171] C. Heymans, T. Tröster, M. Asgari, *et al.*, “KiDS-1000 cosmology: Multi-probe weak gravitational lensing and spectroscopic galaxy clustering constraints,” *Astronomy & Astrophysics*, vol. 646, p. A140, feb 2021.

[172] S. More, S. Sugiyama, H. Miyatake, *et al.*, “Hyper suprime-cam year 3 results: Measurements of clustering of sdss-boss galaxies, galaxy-galaxy lensing and cosmic shear,” 2023.

[173] J. A. Newman and D. Gruen, “Photometric redshifts for next-generation surveys,” *Annual Review of Astronomy and Astrophysics*, vol. 60, pp. 363–414, aug 2022.

[174] M. Salvato, O. Ilbert, and B. Hoyle, “The many flavours of photometric redshifts,” *Nature Astronomy*, vol. 3, pp. 212–222, June 2019.

[175] D. Huterer, M. Takada, G. Bernstein, and B. Jain, “Systematic errors in future weak-lensing surveys: requirements and prospects for self-calibration,” *Monthly Notices of the Royal Astronomical Society*, vol. 366, pp. 101–114, 02 2006.

[176] C. E. Cunha, D. Huterer, M. T. Busha, and R. H. Wechsler, “Sample variance in photometric redshift calibration: cosmological biases and survey requirements,” *Monthly Notices of the Royal Astronomical Society*, vol. 423, pp. 909–924, 05 2012.

[177] D. Huterer, C. E. Cunha, and W. Fang, “Calibration errors unleashed: effects on cosmological parameters and requirements for large-scale structure surveys,” *Monthly Notices of the Royal Astronomical Society*, vol. 432, pp. 2945–2961, 05 2013.

[178] S. Joudaki, C. Blake, C. Heymans, *et al.*, “CFHTLenS revisited: assessing concordance with Planck including astrophysical systematics,” *Monthly Notices of the Royal Astronomical Society*, vol. 465, pp. 2033–2052, 12 2016.

[179] B. Hoyle, D. Gruen, G. M. Bernstein, *et al.*, “Dark Energy Survey Year 1 Results: redshift distributions of the weak-lensing source galaxies,” MNRAS, vol. 478, pp. 592–610, July 2018.

[180] Joudaki, S., Hildebrandt, H., Traykova, D., *et al.*, “Kids+viking-450 and des-y1 combined: Cosmology with cosmic shear,” *A&A*, vol. 638, p. L1, 2020.

[181] S. Fischbacher, T. Kacprzak, J. Blazek, and A. Refregier, “Redshift requirements for cosmic shear with intrinsic alignment,” *arXiv e-prints*, p. arXiv:2207.01627, July 2022.

[182] M. Carrasco Kind and R. J. Brunner, “Exhausting the information: novel Bayesian combination of photometric redshift PDFs,” MNRAS, vol. 442, pp. 3380–3399, Aug. 2014.

[183] H. Hildebrandt, F. Köhlinger, J. L. van den Busch, *et al.*, “KiDS+VIKING-450: Cosmic shear tomography with optical and infrared data,” A&A, vol. 633, p. A69, Jan. 2020.

[184] M. Gatti, G. Giannini, G. M. Bernstein, *et al.*, “Dark Energy Survey Year 3 Results: clustering redshifts - calibration of the weak lensing source redshift distributions with redMaGiC and BOSS/eBOSS,” MNRAS, vol. 510, pp. 1223–1247, Feb. 2022.

[185] M. M. Rau, R. Dalal, T. Zhang, *et al.*, “Weak lensing tomographic redshift distribution inference for the hyper suprime-cam subaru strategic program three-year shape catalogue,” 2022.

[186] L. Tortorelli, M. Fagioli, J. Herbel, *et al.*, “Measurement of the B-band galaxy Luminosity Function with Approximate Bayesian Computation,” J. Cosmology Astropart. Phys., vol. 2020, p. 048, Sept. 2020.

[187] L. Tortorelli, M. Siudek, B. Moser, *et al.*, “The PAU survey: measurement of narrow-band galaxy properties with approximate bayesian computation,” J. Cosmology Astropart. Phys., vol. 2021, p. 013, Dec. 2021.

[188] J. Bergé, L. Gamper, A. Réfrégier, and A. Amara, “An Ultra Fast Image Generator (UFIG) for wide-field astronomy,” *Astronomy and Computing*, vol. 1, pp. 23–32, Feb. 2013.

[189] E. Bertin and S. Arnouts, “SExtractor: Software for source extraction.,” A&AS, vol. 117, pp. 393–404, June 1996.

[190] C. Bruderer, C. Chang, A. Refregier, *et al.*, “Calibrated Ultra Fast Image Simulations for the Dark Energy Survey,” ApJ, vol. 817, p. 25, Jan. 2016.

[191] C. Bruderer, A. Nicola, A. Amara, *et al.*, “Cosmic shear calibration with forward modeling,” *Journal of Cosmology and Astroparticle Physics*, vol. 2018, pp. 007–007, aug 2018.

[192] J.-C. J. Cuillandre, K. Withington, P. Hudelot, *et al.*, “Introduction to the CFHT Legacy Survey final release (CFHTLS T0007),” in *Observatory Operations: Strategies, Processes, and Systems IV* (A. B. Peck, R. L. Seaman, and F. Comeron, eds.), vol. 8448 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, p. 84480M, Sept. 2012.

[193] P. Martí, R. Miquel, F. J. Castander, *et al.*, “Precise photometric redshifts with a narrow-band filter set: the PAU survey at the William Herschel Telescope,” MNRAS, vol. 442, pp. 92–109, July 2014.

[194] L. Tortorelli, L. Della Bruna, J. Herbel, *et al.*, “The PAU Survey: a forward modeling approach for narrow-band imaging,” J. Cosmology Astropart. Phys., vol. 2018, p. 035, Nov. 2018.

[195] M. R. Blanton, M. A. Bershady, B. Abolfathi, *et al.*, “Sloan Digital Sky Survey IV: Mapping the Milky Way, Nearby Galaxies, and the Distant Universe,” AJ, vol. 154, p. 28, July 2017.

[196] M. Fagioli, J. Riebartsch, A. Nicola, *et al.*, “Forward modeling of spectroscopic galaxy surveys: application to SDSS,” J. Cosmology Astropart. Phys., vol. 2018, p. 015, Nov. 2018.

[197] M. Fagioli, L. Tortorelli, J. Herbel, *et al.*, “Spectro-imaging forward model of red and blue galaxies,” J. Cosmology Astropart. Phys., vol. 2020, p. 050, June 2020.

[198] P. Sudek, L. F. de la Bella, A. Amara, and W. G. Hartley, “The sensitivity of the redshift distribution to galaxy demographics,” MNRAS, vol. 516, pp. 1670–1684, Oct. 2022.

[199] A. C. Robin, C. Reylé, S. Derrière, and S. Picaud, “A synthetic view on structure and evolution of the milky way,” *Astronomy & Astrophysics*, vol. 416, p. 157–157, Feb 2004.

[200] and T. Prusti, J. H. J. de Bruijne, A. G. A. Brown, *et al.*, “The gaia mission,” *Astronomy & Astrophysics*, vol. 595, p. A1, nov 2016.

[201] G. Collaboration, A. Vallenari, A. G. A. Brown, *et al.*, “Gaia data release 3: Summary of the content and survey properties,” 2022.

[202] M. R. Blanton, D. J. Schlegel, M. A. Strauss, *et al.*, “New york university value-added galaxy catalog: A galaxy catalog based on new public surveys,” *The Astronomical Journal*, vol. 129, pp. 2562–2578, jun 2005.

[203] F. Tarsitano, W. G. Hartley, A. Amara, *et al.*, “A catalogue of structural and morphological measurements for DES Y1,” MNRAS, vol. 481, pp. 2018–2040, Dec. 2018.

[204] J. Herbel, T. Kacprzak, A. Amara, *et al.*, “Fast point spread function modeling with deep learning,” J. Cosmology Astropart. Phys., vol. 2018, p. 054, July 2018.

[205] L. Bradley, B. Sipőcz, T. Robitaille, *et al.*, “astropy/photutils: 1.5.0,” July 2022.

[206] N. Scoville, H. Aussel, M. Brusa, *et al.*, “The cosmic evolution survey (COSMOS): Overview,” *The Astrophysical Journal Supplement Series*, vol. 172, pp. 1–8, sep 2007.

[207] A. Gretton, K. M. Borgwardt, M. J. Rasch, *et al.*, “A kernel two-sample test,” *J. Mach. Learn. Res.*, vol. 13, p. 723–773, Mar. 2012.

[208] M. R. Blanton and S. Roweis, “K-Corrections and Filter Transformations in the Ultraviolet, Optical, and Near-Infrared,” AJ, vol. 133, pp. 734–754, Feb. 2007.

[209] C. J. Conselice, “The evolution of galaxy structure over cosmic time,” *Annual Review of Astronomy and Astrophysics*, vol. 52, pp. 291–337, aug 2014.

[210] Euclid Collaboration, Scaramella, R., Amiaux, J., *et al.*, “Euclid preparation - i. the euclid wide survey,” *A&A*, vol. 662, p. A112, 2022.

[211] I. K. Baldry, K. Glazebrook, J. Brinkmann, *et al.*, “Quantifying the Bimodal Color-Magnitude Distribution of Galaxies,” ApJ, vol. 600, pp. 681–694, Jan. 2004.

[212] J. Alsing, H. Peiris, J. Leja, *et al.*, “SPECULATOR: Emulating stellar population synthesis for fast and accurate galaxy spectra and photometry,” *The Astrophysical Journal Supplement Series*, vol. 249, p. 5, jun 2020.

[213] A. P. Hearin, J. Chaves-Montero, A. Alarcon, *et al.*, “DSPS: Differentiable stellar population synthesis,” *Monthly Notices of the Royal Astronomical Society*, vol. 521, pp. 1741–1756, feb 2023.

[214] K. J. Kwon, C. Hahn, and J. Alsing, “Neural stellar population synthesis emulator for the DESI PROVABGS,” *The Astrophysical Journal Supplement Series*, vol. 265, p. 23, mar 2023.

[215] P. Berner, A. Refregier, B. Moser, *et al.*, “Fast Forward Modelling of Galaxy Spatial and Statistical Distributions,” *arXiv e-prints*, p. arXiv:2310.15223, Oct. 2023.

[216] Q. Wang, S. R. Kulkarni, and S. Verdu, “Divergence estimation for multidimensional densities via k-nearest-neighbor distances,” *IEEE Transactions on Information Theory*, vol. 55, no. 5, pp. 2392–2405, 2009.

[217] J. Amiaux, R. Scaramella, Y. Mellier, *et al.*, “Euclidmission: building of a reference survey,” *Space Telescopes and Instrumentation 2012: Optical, Infrared, and Millimeter Wave*, Aug 2012.

[218] Z. Ma, W. Hu, and D. Huterer, “Effects of Photometric Redshift Uncertainties on Weak-Lensing Tomography,” ApJ, vol. 636, pp. 21–29, Jan. 2006.

[219] M. Salvato, O. Ilbert, and B. Hoyle, “The many flavours of photometric redshifts,” *Nature Astronomy*, vol. 3, pp. 212–222, June 2019.

[220] S. Fotopoulou and S. Paltani, “CPz: Classification-aided photometric-redshift estimation,” A&A, vol. 619, p. A14, Oct. 2018.

[221] P. Graff, F. Feroz, M. P. Hobson, and A. Lasenby, “SkyNet: an efficient and robust neural network training tool for machine learning in astronomy,” *Monthly Notices of the Royal Astronomical Society*, vol. 441, pp. 1741–1759, 05 2014.

[222] H. Hildebrandt, M. Viola, C. Heymans, *et al.*, “KiDS-450: cosmological parameter constraints from tomographic weak gravitational lensing,” MNRAS, vol. 465, pp. 1454–1498, Feb. 2017.

[223] T. Kohonen, M. R. Schroeder, and T. S. Huang, *Self-Organizing Maps*. Berlin, Heidelberg: Springer-Verlag, 3rd ed., 2001.

[224] A. Johnson, C. Blake, A. Amon, *et al.*, “2dFLenS and KiDS: determining source redshift distributions with cross-correlations,” MNRAS, vol. 465, pp. 4118–4132, Mar. 2017.

[225] M. Gatti, G. Giannini, G. M. Bernstein, *et al.*, “Dark energy survey year 3 results: clustering redshifts – calibration of the weak lensing source redshift distributions with redmagic and boss/eboss,” *Monthly Notices of the Royal Astronomical Society*, vol. 510, p. 1223–1247, Nov. 2021.

[226] T. M. C. Abbott, F. B. Abdalla, S. Allam, *et al.*, “The dark energy survey: Data release 1,” *The Astrophysical Journal Supplement Series*, vol. 239, p. 18, Nov 2018.

[227] I. Sevilla-Noarbe, K. Bechtol, M. Carrasco Kind, *et al.*, “Dark energy survey year 3 results: Photometric data set for cosmology,” *The Astrophysical Journal Supplement Series*, vol. 254, p. 24, May 2021.

[228] M. Jarvis, G. M. Bernstein, A. Amon, *et al.*, “Dark energy survey year 3 results: point spread function modelling,” *Monthly Notices of the Royal Astronomical Society*, vol. 501, p. 1282–1299, Nov. 2020.

[229] E. S. Sheldon and E. M. Huff, “Practical weak-lensing shear measurement with metacalibration,” *The Astrophysical Journal*, vol. 841, p. 24, May 2017.

[230] I. Sevilla-Noarbe, K. Bechtol, M. Carrasco Kind, *et al.*, “Dark Energy Survey Year 3 Results: Photometric Data Set for Cosmology,” ApJS, vol. 254, p. 24, June 2021.

[231] J. R. Myers, C. B. Sande, A. C. Miller, *et al.*, “VizieR Online Data Catalog: SKY2000 Catalog, Version 4 (Myers+ 2002),” *VizieR Online Data Catalog*, p. V/109, Sept. 2001.

[232] W. G. Hartley, A. Choi, A. Amon, *et al.*, “Dark energy survey year 3 results: Deep field optical + near-infrared images and catalogue,” *Monthly Notices of the Royal Astronomical Society*, vol. 509, p. 3547–3579, Oct. 2021.

[233] S. Everett, B. Yanny, N. Kuropatkin, *et al.*, “Dark energy survey year 3 results: Measuring the survey transfer function with balrog,” *The Astrophysical Journal Supplement Series*, vol. 258, p. 15, Jan. 2022.

[234] S. J. Lilly, V. Le Brun, C. Maier, *et al.*, “The zCOSMOS 10k-Bright Spectroscopic Sample,” ApJS, vol. 184, pp. 218–229, Oct. 2009.

[235] D. C. Masters, D. K. Stern, J. G. Cohen, *et al.*, “The Complete Calibration of the Color-Redshift Relation (C3R2) Survey: Analysis and Data Release 2,” ApJ, vol. 877, p. 81, June 2019.

[236] O. Le Fèvre, P. Cassata, O. Cucciati, *et al.*, “The VIMOS VLT Deep Survey final data release: a spectroscopic sample of 35 016 galaxies and AGN out to z ~6.7 selected with $17.5 \le i_{AB} \le 24.75$,” A&A, vol. 559, p. A14, Nov. 2013.

[237] M. Scodeggio, L. Guzzo, B. Garilli, *et al.*, “The VIMOS Public Extragalactic Redshift Survey (VIPERS). Full spectroscopic data and auxiliary information release (PDR-2),” A&A, vol. 609, p. A84, Jan. 2018.

[238] A. Alarcon, E. Gaztanaga, M. Eriksen, *et al.*, “The pau survey: an improved photo-z sample in the cosmos field,” *Monthly Notices of the Royal Astronomical Society*, vol. 501, p. 6103–6122, Jan. 2021.

[239] E. Suchyta, E. M. Huff, J. Aleksić, *et al.*, “No galaxy left behind: accurate measurements with the faintest objects in the dark energy survey,” *Monthly Notices of the Royal Astronomical Society*, vol. 457, p. 786–808, Jan 2016.

[240] G. Bruzual and S. Charlot, “Stellar population synthesis at the resolution of 2003,” *Monthly Notices of the Royal Astronomical Society*, vol. 344, pp. 1000–1028, 10 2003.

[241] A. G. A. Brown, A. Vallenari, T. Prusti, *et al.*, “Gaia data release 2,” *Astronomy & Astrophysics*, vol. 616, p. A1, Aug 2018.

[242] J. Akeret, A. Refregier, A. Amara, *et al.*, “Approximate bayesian computation for forward modeling in cosmology,” *Journal of Cosmology and Astroparticle Physics*, vol. 2015, p. 043–043, Aug 2015.

[243] T. Kacprzak, J. Herbel, A. Amara, and A. Réfrégier, “Accelerating Approximate Bayesian Computation with Quantile Regression: application to cosmological redshift distributions,” J. Cosmology Astropart. Phys., vol. 2018, p. 042, Feb. 2018.

[244] D. Gruen and F. Brimioulle, “Selection biases in empirical p(z) methods for weak lensing,” MNRAS, vol. 468, pp. 769–782, June 2017.

[245] C. Sánchez, M. Raveri, A. Alarcon, and G. M. Bernstein, “Propagating sample variance uncertainties in redshift calibration: simulations, theory, and application to the COSMOS2015 data,” MNRAS, vol. 498, pp. 2984–2999, Oct. 2020.

[246] E. Morganson, R. A. Gruendl, F. Menanteau, *et al.*, “The Dark Energy Survey Image Processing Pipeline,” PASP, vol. 130, p. 074501, July 2018.

[247] J. DeRose, R. Wechsler, M. Becker, *et al.*, “Dark energy survey year 3 results: Cosmology from combined galaxy clustering and lensing validation on cosmological simulations,” *Physical Review D*, vol. 105, June 2022.

[248] F. Pedregosa, G. Varoquaux, A. Gramfort, *et al.*, “Scikit-learn: Machine learning in Python,” *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

[249] R. H. Lupton, J. E. Gunn, and A. S. Szalay, “A Modified Magnitude System that Produces Well-Behaved Magnitudes, Colors, and Errors Even for Low Signal-to-Noise Ratio Measurements,” AJ, vol. 118, pp. 1406–1410, Sept. 1999.

[250] O. Troyanskaya, M. Cantor, G. Sherlock, *et al.*, “Missing value estimation methods for DNA microarrays ,” *Bioinformatics*, vol. 17, pp. 520–525, 06 2001.

[251] J. DeRose, R. H. Wechsler, M. R. Becker, *et al.*, “The Buzzard Flock: Dark Energy Survey Synthetic Sky Catalogs,” *arXiv e-prints*, p. arXiv:1901.02401, Jan. 2019.

[252] K. Kuijken, C. Heymans, A. Dvornik, *et al.*, “The fourth data release of the kilo-degree survey: ugri imaging and nine-band optical-ir photometry over 1000 square degrees,” *Astronomy & Astrophysics*, vol. 625, p. A2, Apr. 2019.

[253] J. R. Weaver, I. Davidzon, S. Toft, *et al.*, “Cosmos2020: The galaxy stellar mass function: The assembly and star formation cessation of galaxies at $0.2 \leq z \leq 7.5$,” *Astronomy & Astrophysics*, vol. 677, p. A184, Sept. 2023.

[254] J. Alsing, S. Thorp, S. Deger, *et al.*, “pop-cosmos: A comprehensive picture of the galaxy population from cosmos data,” 2024.

[255] D. C. Masters, D. K. Stern, J. G. Cohen, *et al.*, “The Complete Calibration of the Color-Redshift Relation (C3R2) Survey: Survey Overview and Data Release 1,” ApJ, vol. 841, p. 111, June 2017.

[256] D. C. Masters, D. K. Stern, J. G. Cohen, *et al.*, “The complete calibration of the color–redshift relation (c3r2) survey: Analysis and data release 2,” *The Astrophysical Journal*, vol. 877, p. 81, may 2019.

[257] Euclid Collaboration, Guglielmo, V., Saglia, R., *et al.*, “Euclid preparation - viii. the complete calibration of the colour–redshift relation survey: Vlt/kmos observations and data release,” *A&A*, vol. 642, p. A192, 2020.

[258] J. McCullough, D. Gruen, A. Amon, *et al.*, “Desi complete calibration of the color-redshift relation (dc3r2): Results from early desi data,” 2023.

[259] S. P. Driver, J. Liske, L. J. M. Davies, *et al.*, “4MOST Consortium Survey 7: Wide-Area VISTA Extragalactic Survey (WAVES),” *The Messenger*, vol. 175, pp. 46–49, Mar. 2019.

[260] D. Gruen, J. McCullough, A. Amon, *et al.*, “4MOST Complete Calibration of the Colour-Redshift Relation (4C3R2),” *The Messenger*, vol. 190, pp. 28–30, Mar. 2023.

[261] C. R. Harris, K. J. Millman, S. J. van der Walt, *et al.*, “Array programming with NumPy,” *Nature*, vol. 585, pp. 357–362, sep 2020.

[262] P. Virtanen, R. Gommers, T. E. Oliphant, *et al.*, "SciPy 1.0: fundamental algorithms for scientific computing in python," *Nature Methods*, vol. 17, pp. 261–272, feb 2020.

[263] J. D. Hunter, "Matplotlib: A 2d graphics environment," *Computing in Science & Engineering*, vol. 9, no. 3, pp. 90–95, 2007.

[264] A. Zonca, L. Singer, D. Lenz, *et al.*, "healpy: equal area pixelization and spherical harmonics transforms for data on the sphere in Python," *The Journal of Open Source Software*, vol. 4, p. 1298, Mar. 2019.

[265] M. Abadi *et al.*, "TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems," 3 2016.

[266] A. Lewis, "Getdist: a python package for analysing monte carlo samples," 2019.

[267] F. Perez and B. E. Granger, "IPython: A System for Interactive Scientific Computing," *Computing in Science and Engineering*, vol. 9, pp. 21–29, Jan. 2007.

[268] T. Kluyver, B. Ragan-Kelley, F. Pérez, *et al.*, "Jupyter Notebooks—a publishing format for reproducible computational workflows," in *IOS Press*, pp. 87–90, 2016.

[269] T. A. Collaboration, A. M. Price-Whelan, B. M. Sipőcz, *et al.*, "The astropy project: Building an open-science project and status of the v2.0 core package*," *The Astronomical Journal*, vol. 156, p. 123, aug 2018.

[270] S. K. Lam, A. Pitrou, and S. Seibert, "Numba: a LLVM-based Python JIT compiler," in *International Conference for High Performance Computing, Networking, Storage and Analysis: HPC Transforms*, 2015.

[271] J. V. D. B. e. a. J. Reback, W. McKinney, "pandas-dev/pandas: Pandas 1.1. 1," in *Zenodo*, 2020.

[272] F. Pedregosa, G. Varoquaux, A. Gramfort, *et al.*, "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

[273] S. Behnel, R. Bradshaw, C. Citro, *et al.*, "Cython: The best of both worlds," *Computing in Science & Engineering*, vol. 13, pp. 31 – 39, 05 2011.

[274] A. Meurer, C. P. Smith, M. Paprocki, *et al.*, “Sympy: symbolic computing in python,” *PeerJ Computer Science*, vol. 3, p. e103, Jan. 2017.

[275] T. Kacprzak and J. Fluri, “Deeplss: breaking parameter degeneracies in large scale structure with deep learning analysis of combined probes,” 2022.

[276] D. Zürcher, J. Fluri, R. Sgier, *et al.*, “Cosmological forecast for non-gaussian statistics in large-scale weak lensing surveys,” *Journal of Cosmology and Astroparticle Physics*, vol. 2021, pp. 028–028, jan 2021.

[277] D. Zürcher, J. Fluri, V. Ajani, *et al.*, “Towards a full wCDM map-based analysis for weak lensing surveys,” MNRAS, vol. 525, pp. 761–784, Oct. 2023.

ACKNOWLEDGEMENTS

Gli anni passano e non lasciano niente
No non è vero qui c'è un sacco di gente
Un sacco di gente, noi fradici sotto
E chi se ne frega se piove a dirotto

— EUGENIO IN VIA DI GIOIA, Tornano

I feel at the same time as if I just started and have been here for my entire life. Four and a half years can feel like an eternity, particularly when you spend half of that time wondering whether you will see your office and your colleagues ever again. During a PhD, the passage of time is marked by people who join and leave the corridor. I have met many people, and would not be here if it was not for some of them.

The first thank you is for my supervisor, Alexandre Refregier. You gave me a great opportunity, helped me widen my horizons and allowed me to grow and have many experiences, that I could not have imagined the day I decided to join your group. We had our disagreements but I admire your enthusiasm and the fact that you listened and tried to improve the group for all of us. Secondly, I want to thank my second advisor, Tomasz Kacprzak. You encouraged me to try out new things, provided ideas and taught me a lot in terms of practical skills and tackling scientific problems. Our approaches were sometimes quite different, but this added value to the final result. I would also like to thank Daniel Grün and not only for being my external co-examiner.

You offered your help when I most needed it, despite not having to. You later welcomed me in Munich, and provided very valuable advice.

I also owe a big thank you to the DES collaboration, in particular to the conveners and members of the redshift and weak lensing working groups, for always being available for answering my questions, for the telecons, the collaboration meetings and for giving me great suggestions on my project.

I would like to thank Dominik for being my PhD big brother and an amazing office mate. You introduced me to the secrets of the Cosmology group and became my friend during our trip to Stanford. This is when I started enjoying my PhD. You taught me a lot, both in terms of making my code less hacky and showing me how Swiss people can have a heart too. Thank you Pascale, for being there from (and before) the beginning. You have been a loyal friend and a valuable colleague. I have enjoyed our endless chats about galaxy population models and life and many trips together. I would also like to thank Martina. We only overlapped as colleagues for a short period of time but that was enough to make you one of my dearest friends. I hope I succeeded at keeping your legacy and that of Italian women in the group. Thank you Agne for being my PhD twin sister. Meeting you in Leiden and then Tonale was like a sign of fate: we were always destined to be friends. Thanks for always being there for me, even without ever living in the same city.

The whole PhD journey would have been unbearable without many other travel companions. I would like to thank Al for our office chats and for being the best co-organizer for scientific (and less scientific) events such as Journal Club, Cosmo Christmas, Friday beers and pub quizzes. Thank you Silvan for bearing with me. Noone in the office was as exposed to my complaints and frustrations as were you (my Arne) and you did really well in keeping me sane. Having a second pair of eyes on my code was super helpful, as were your code optimizations (the Earth is thankful for the reduced CO_2 emissions). You have both been amazing friends: I believe it is people like you who make this world a better place. Thank you Arne for making me feel a bit of a mama bear when you started your PhD and opening up with me. I would like to thank all the PhDs of the Cosmology group for the incredible memories we made together in South Africa, Greece, France, Italy and Switzerland. Thank you Pascalito, Joël, Jennifer and Veronika.

Thank you also to the previous generation of PhD students. Thank you Raphaël and Jörg, for providing guidance at the beginning of my PhD and for the beers enjoyed together. Thank you Luca for many scientific chats, for having hosted me in Munich and being a friend. Thank you to all other members of the Cosmology group: Milena, Devin, Tilman, Geri, John, Jozef, Luis and Pierre. Thank you Christiane for contributing to our project together. And thank you Uwe for teaching me so much about best code practices, Python and also for many good laughs. And finally a huge *grazie* to Marta and Virgi. It was such *una boccata d'aria fresca* to have you here in Zürich. You reminded me of how much I missed spontaneity and oversharing. I could be truly myself with you and always felt understood. You have been amazing both scientifically and personally, and I already miss you incredibly much.

I would also like to thank my past and present office mates Sofia, Saeed, Dominik and Joël for making my time at the office lighter and more enjoyable. Thank you to Christian Herzog for the IT coffee chats, both in the first and last semester of my PhD: I truly enjoyed talking to you and learning something about IT. Thank you to my tandem partners, Sina and Lukas, for many lunches together learning German and Italian.

I would like to thank Adrian Glauser for giving me the opportunity of being an Astrowoche assistant and for many enjoyable conversations. Thank you also to the other Astrowoche assistants, it has been an incredibly fun adventure: Adrian, Polychronis, Gabriele, Jie, Silvan, Janina and Markus. And thank you also to my students who managed to survive and take some measurements despite my inexperienced guidance. Thank you to all the members of the other astrophysics groups from the HIT J floor: it has been fun to get to know you. Particular mention to Larry, our game master for many sessions of DnD and a friend.

Thank you to the students I had the pleasure to supervise: Michael, Dominic, Jöel, Hao and Alessandro. And also thank you to the Euler Cluster support for answering all of my (not always smart) questions and making my work possible through the maintenance of the Euler cluster.

I would also like to thank the beautiful people who became my friends in the last 7 years in Zürich. Thank you to Alex for making me feel less alone, for the many chats about life, Switzerland and the world. Thank you Gabriela for listening to me and being my number 1 supporter. Spending time with you always makes me feel happier. Thank you Chigusa for all the planned activities, the Dietikon runs (and pizzas) and being a great friend. Thank you Nikhil and Richard for staying in touch after moving abroad, for the hikes, the laughs and for the PhD day count. I defend my PhD on PhD day 1651. Thank you to all my other flatmates for the chats, the dinners and making me feel at home: Esteban, Lorenz, Milena, Merel, Kalle, Yannis, Alessia and all the others. Thank you to the Master thesis gang Deniz, Simon, Jasper and Jules for being completely different from me and still becoming surprisingly close. Thank you Gitanjali and Kaustuv, because while we meet and chat very little, I always appreciate when we do.

Ci sono poi tante persone da ringraziare che mi conoscono da ben prima che mi trasferissi a Zurigo e che mi hanno accompagnata anche in questa avventura. Voglio ringraziare le mie amiche di sempre, quelle senza le quali io non sarei *Io*. Grazie Megghi per avermi sopportata fin da quando ti rubavo i giocattoli di mano e ti inondavo con ore e ore di chiacchiere. E grazie che sei ancora con me. Grazie a Madda e Marti per essere le mie amiche del cuore, per aver accompagnato la mia vita fin qui e spero fino alla fine, grazie per i *se sei felice tu*. Grazie ai miei amici della triennale Fede, Sangio, Clara, Alvise, Giacomino, George e Arianna per aver accompagnato i miei primi passi nella fisica. E soprattutto grazie Zac: per quanto le nostre strade si siano divise, rimani una delle persone più importanti della mia vita e sei stato essenziale durante i primi anni del dottorato. Grazie alle mie amiche di pallavolo Anna, Ika ed Eli, le pallavoliste mancate, per essere rimaste un punto fermo del mio panorama trentino. Grazie a Lucrezia, per avermi fatto vedere delle parti di me che non conoscevo e avermi accompagnata nell'ultimo anno e mezzo di PhD, e a Giuly e Mareike per avermi riaggiustata quando ho perso pezzi per strada.

Grazie alla mia Famiglia. Prima di tutto grazie a mamuskino e patch per avermi guidata, sostenuta e aver sempre creduto in me. Senza i vostri *Cice capace* e la passione e l'impegno nel lavoro che mi avete insegnato non sarei qui a scrivere questa tesi. E grazie per aver lasciato volare il pulcino dal nido. Non era scontato e mi siete sempre stati vicini nonostante la distanza fisica che ci separa. Grazie ai miei meravigliosi nonni. Grazie alla nonna Angioletta, ai suoi ovetti sbattuti e alle sue confidenze, al nonno Guido e alle storie dei *suoi tempi*, grazie al nonno Renato per la sua bontà silenziosa e alla nonna Anna per la fierezza con cui parla di me. Grazie a tutti i miei cugini, in particolare a Michi, che mi ha accompagnata dai pomeriggi a giocare a Playmobil, al liceo Prati fino allo *spazio* e alla mia *cuginetta* Madda, che è sempre stata splendida, da quando l'ho presa in braccio per la prima volta ad adesso che sta per iniziare l'Università. Grazie anche a tutti i miei zii, ma in particolare a Claudio e Marianna, tra i più grandi sostenitori del mio percorso accademico. Grazie Claudio, sono gli scienziati come te che mi ricordano perché ho iniziato questo percorso.

Finally, I would like to thank Janis for being my partner and my friend. You have been special from the very beginning: when you were my colleague, you have always been ready to listen and help. And in the past years, and months especially, you have been incredibly supportive: you waited for me when I was late, you cooked my dinners and listened to my *unnecessarily long* stories. I am truly grateful and happy to have you by my side.

This thesis was supported in part by grant 200021_192243 from the Swiss National Science Foundation. Chapter 3 is based on data collected at the Subaru Telescope and retrieved from the HSC data archive system, which is operated by Subaru Telescope and Astronomy Data Centre (ADC) at NAOJ. COSMOS2020 is based on observations collected at the European Southern Observatory under ESO programme ID 179.A-2005 and on data products produced by CALET and the Cambridge Astronomy Survey Unit on behalf of the UltraVISTA consortium. This work has made use of data from the European Space Agency (ESA) mission *Gaia* (https://www.cosmos.esa.int/gaia), processed by the *Gaia* Data Processing and Analysis Consortium (DPAC, https://www.cosmos.esa.int/web/gaia/dpac/consortium). Funding for the DPAC has been provided by national institutions, in particular the institutions participating in the *Gaia* Multilateral Agreement.

Funding for the DES Projects has been provided by the U.S. Department of Energy, the U.S. National Science Foundation, the Ministry of Science and Education of Spain,

the Science and Technology Facilities Council of the United Kingdom, the Higher Education Funding Council for England, the National Center for Supercomputing Applications at the University of Illinois at Urbana-Champaign, the Kavli Institute of Cosmological Physics at the University of Chicago, the Center for Cosmology and Astro-Particle Physics at the Ohio State University, the Mitchell Institute for Fundamental Physics and Astronomy at Texas A&M University, Financiadora de Estudos e Projetos, Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro, Conselho Nacional de Desenvolvimento Científico e Tecnológico and the Ministério da Ciência, Tecnologia e Inovação, the Deutsche Forschungsgemeinschaft and the Collaborating Institutions in the Dark Energy Survey. The Collaborating Institutions are Argonne National Laboratory, the University of California at Santa Cruz, the University of Cambridge, Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas-Madrid, the University of Chicago, University College London, the DES-Brazil Consortium, the University of Edinburgh, the Eidgenössische Technische Hochschule (ETH) Zürich, Fermi National Accelerator Laboratory, the University of Illinois at Urbana-Champaign, the Institut de Ciències de l'Espai (IEEC/CSIC), the Institut de Física d'Altes Energies, Lawrence Berkeley National Laboratory, the Ludwig-Maximilians Universität München and the associated Excellence Cluster Universe, the University of Michigan, NSF's NOIRLab, the University of Nottingham, The Ohio State University, the University of Pennsylvania, the University of Portsmouth, SLAC National Accelerator Laboratory, Stanford University, the University of Sussex, Texas A&M University, and the OzDES Membership Consortium. Based in part on observations at Cerro Tololo Inter-American Observatory at NSF's NOIRLab (NOIRLab Prop. ID 2012B-0001; PI: J. Frieman), which is managed by the Association of Universities for Research in Astronomy (AURA) under a cooperative agreement with the National Science Foundation. The DES data management system is supported by the National Science Foundation under Grant Numbers AST-1138766 and AST-1536171. The DES participants from Spanish institutions are partially supported by MICINN under grants ESP2017-89838, PGC2018-094773, PGC2018-102021, SEV-2016-0588, SEV-2016-0597, and MDM-2015-0509, some of which include ERDF funds from the European Union. IFAE is partially funded by the CERCA program of the Generalitat de Catalunya. Research leading to these results has received funding from the European Research

Council under the European Union's Seventh Framework Program (FP7/2007-2013) including ERC grant agreements 240672, 291329, and 306478. We acknowledge support from the Brazilian Instituto Nacional de Ciência e Tecnologia (INCT) do e-Universo (CNPq grant 465376/2014-2). This manuscript has been authored by Fermi Research Alliance, LLC under Contract No. DE-AC02-07CH11359 with the U.S. Department of Energy, Office of Science, Office of High Energy Physics.

This work benefited from open-source software and we would like to thank the community for developing and maintaining the used packages. We used functionalities from `numpy` [261], `scipy` [262], `matplotlib` [263], `HEALPix` and `healpy` [264], `tensorflow` [265], `GetDist` [266], `ipython` [267], `jupyter` [268], `astropy` [269], `Numba` [270], `pandas` [271], `scikit-learn` [272], `Cython` [273] and `sympy` [274]. Corner plots were created with `trianglechain` [181, 275]. We submitted jobarrays to the Euler cluster using `esub-epipe` [276, 277].